AWKA JOURNAL OF ENGLISH LANGUAGE AND LITERARY STUDIES (AJELLS)

Volume 13 Number 3
September, 2026

Orthographic Architecture of an Unwritten Language: A Descriptive-Linguistic Framework for the Codification of Zafan

Mufida Isyaku Bashir
Department of English
Federal University Gusau
Mufidartuw@gmail.com
ORCHID ID: 0009 0006 1791 7001

&
Shehu Sidi Ibrahim
Department of English
Zamfara State University Talata-Mafara
ccd379@yahoo.com

Abstract
Writing plays an important role in preserving language and culture, but at this point in time, many Nigerian Languages are yet to have a standard written form. This has left their native speakers not to use them in school or digital spaces even if they want to. This paper looks at this problem through the lens of Zafan, an imaginary unwritten Nigerian Language, demonstrating how the learners of the community abandon its use in both mobile and online communication which leads to inconsistent spelling even when they try to formally use it. The work was underpinned by Mother Tongue Based Multilingual Education (MTB-MLE) framework adapting its idea of building a bridge from the known (Zafan) to the unknown (English or Hausa) leveraging on the frameworks' emphasis on preserving cultural identity even though it was originally designed for classroom teaching rather than for creating a writing system. The work follows a comprehensive process of phonetic documentation, phonemic analysis, morphological mapping, script selection, grapheme assignment as well as community testing and standardization which were all guided by universal patterns across human languages including,

but not limited to the principle of accuracy, familiarity as well as ease of typing. The result of a potential Zafan alphabet was seen based on Latin script, familiar vowel and consonant symbols alongside clear pronoun system. The paper concludes by reaffirming that a simple, easy to type, community approved writing system offers the Zafan community speaker's the ability to read, write and preserve their language for their generation unborn.
Keywords: Orthography design, phonemic inventory, Language codification, Language Documentation, Grapheme-Phoneme

Introduction

Writing is a basic tool of civilization, without it, the world as we know it could not exist (Fromkin et.al 2014). Transition of a language from a purely oral tradition to a codified written form is one of the most complex undertakings in applied linguistics. Beyond the mere assignment of symbols to sounds, the development of an orthography represents a profound act of linguistic documentation and preservation. As noted by Stephen May (2017), the "observability" of a languages internal logic and cognitive value is often dependent upon its presence in a written, standardized format. Following Greenbergs (1963) typology, we assume that Zafan (an imaginary name given to the language), though previously unwritten, must conform to the universal constraints of human speech.

For instance, the absolute universal that all languages distinguish between consonants and vowels, all languages have nouns and verbs, all languages use a set of phonemes or vowels, and again, all languages have universal tendencies that in most languages; the subject precedes the object (SOV or SVO) allows us to categorize the Zafan sound system with confidence, ensuring the script is

compatible with the universal mental grammar shared by all humans. This is because, without a script, the intricate morphosyntactic structures as well as phonological nuances of a speech community remain largely invisible to the global academic community and the formal educational system.

This work reduces a previously unwritten language Zafan to writing. By applying the criteria of Smalley (1964) and the foundational principles of Pike (1947), the work seeks to provide a scientifically robust and psychologically accessible script. It is worth noting that the process requires a systematic progression through the levels of linguistic analysis beginning with the raw acoustic data of phonetics and moving towards the mental representations of phonemics to the last stage of testing the proposed orthography.

In the succeeding sections presented, statement of the problem, followed by theoretical framework, some contemporary literature review would follow and the rigorous stages required to architect this writing system shall be outlined. The stages include, the phonetic documentation, phonemic analysis, morphological mapping, script selection, grapheme assignment and community testing and standardization that satisfies the competing demands of linguistic accuracy and socio-technical viability.

Statement of the Problem

Language is not merely a tool for communication; it is the primary vehicle for cognitive development and the preservation of cultural heritage. For many indigenous communities in Nigeria, their linguistic identity remains confined to oral traditions. In Nigeria, the exact number of unwritten minority languages are difficult to

pin-point with a single figure because unwritten is a moving target now as many languages are currently in the process of developing orthographies. However, there are about 400-450 languages that are either completely unwritten or lack a standardized, widely accepted orthography. And again, there are about 165 languages in Nigeria that are currently endangered. Among the endangered languages are Ajawa, formerly spoken in Bauchi state but speakers shifted to Hausa between 1920 and 1940.

Auyokawa & Teshenawa, Afro Asiatic language spoken in Jigawa, Basa- Gumna, a Kainji Language from Niger and Nasarawa, Kpati, a grassfields language from Taraba reported extinct by 1984, Kubi (Kuba/Kabawa), formerly spoken in Bauchi but the ethnic group now speaks Hausa, Jalaa (Centum), an unclassified language of Gombe is also extinct to mention but a few. Unwritten languages face an elevated risk of extinction when younger generations cease speaking them. This generational decline drives a process known as language shift, which can ultimately relegate a heritage language to the status of a bedroom language one restricted entirely to private, domestic use. When a community is losing its native language, (like a minority language in Nigeria being replaced by English or Hausa), for example, the native language often retreats to the bedroom before disappearing entirely. As the world moves toward a digital-first global economy, unwritten languages face a dual threat: linguistic extinction (language shift) and educational exclusion. Without a script, a language cannot be used as a medium of instruction, a tool for literacy, or a medium for digital expression, effectively silencing its speakers in the global literacy landscape (Jones & Mooney, 2020).

The work seeks to address the persistent lack of a standardized orthography for Zafan Language, which has led to three critical crises: first, t**he Pedagogical Gap:** students who are native speakers of Zafan Community will enter the school environments without the benefit of L1-to-L2 Transfer. Because they have never

seen their own language written, they lack the cognitive scaffolding necessary to map phonemes to graphemes, making the acquisition of English literacy significantly more difficult (Mahlangu, 2024).

Secondly, **the Digital Divide:** in an era of mobile communication, the absence of a Unicode-compatible script prevents the youth from using their mother tongue in digital spaces. This leads to domain loss, where the language is abandoned for social media and texting in favour of English or Hausa, accelerating the rate of language endangerment (Berez-Kroeker et al., 2023). And, thirdly, **Orthographic Inconsistency:** existing attempts at writing the language by community members remain fragmented and inconsistent, often relying on English or Hausa spelling conventions that fail to capture the unique tonal and phonological nuances of the language. This lack of standardization creates confusion in literacy materials and hinders the production of formal literature. Therefore, the work recommends an orthography for the Zafan imaginary language.

Theoretical Framework

This work adopts the Mother Tongue-Based Multilingual Education (MTB-MLE) framework. While the original framework focuses on pedagogical delivery and classroom instruction, this study adopts its core principles to the pre-pedagogical stage: the linguistic codification and orthography design of the Zafan language. We are not merely adopting a teaching method but rather, modifying the framework to serve as a blueprint for technical linguistic development. The choice of MTB-MLE is justified by the sociolinguistic reality of Northern Nigeria. In an environment dominated by "vehicular languages" like Hausa and English, minority languages often face "linguistic displacement." Using MTB-MLE provides a rights-based justification for

orthography development, arguing that literacy is most effective when it begins in the learner's primary linguistic environment. It justifies the creation of the Zafan script as a linguistic exercise.

The MTB-MLE framework has useful components that this work finds imperative for inclusion such as, The Bridge Principle**,** which posits that literacy should move from the known (Zafan) to the unknown (English/Hausa). This justifies the design of a phonemic orthography that is shallow or transparent, making it easy for Zafan speakers to decode symbols based on their intuitive phonetic knowledge. Cultural Identity Preservation, this component is vital because it informs the choice of graphemes. For the Zafan language, we have selected symbols that respect the cultural nuances of the speakers, ensuring that the script feels theirs rather than an imposition.

Literature Review: Thematic Review

A Transition from extractive Research Methods to Co-created Grassroots Initiatives: in the last five years, expert linguistic model has moved towards participatory research in linguistics (PRL). Jones and Mooney (2020) argue that orthography is a key component of language revitalization that must be negotiated rather than imposed. They are added by highlighting that the communitys sense of identity often outweighs scientific phonemic mapping. Consequently, Sackett (2017) provides a case study of Tashkur community, demonstrating that Goal-centered orthography where the community defines what they want to use the writing for (e.g WhatsApp vs. formal textbooks), should dictate the scripts design. He however also emphasises that orthography development is a dynamic, ongoing process rather than a single

event, requiring continued adaptation and community engagement. Similarly, Kutsch Lojenga (as cited in Yoder, 2000) continues to influence the field with a six-step participatory process where native speakers, not outside linguists, make the final decisions on grapheme selection. Her work focuses on empowering native speakers to make decisions about their own writing systems.

Orthographic Depth and Cognitive Scaffolding: research has reconsidered the Orthographic Depth Hypothesis, specifically in the context of multilingual education and translanguaging. Snider (2025) synthesizes the tension between shallow (one-to-one phoneme/grapheme) and deep (representing underlying morphology) orthographies. He suggests that for new readers, a shallow system is easier to learn, but a deeper system may better serve long-term fluency and morphological awareness. Consequently, Mahlangu (2024) introduces a critical perspective on the monopoly of dominant languages like English. He posits that indigenous orthographies stand to gain from a terminological association with multinational languages. This suggests that designing an orthography that shares some conventions with a prestige language like English can act as a cognitive scaffold, making it easier for students to transfer literacy skills between their mother tongue and English Medium Instruction (EMI).

Sociolinguistic Wars and Ideological Distance: The concept of orthography war remain a central theme in recent years, focusing on the political nature of spelling. Scholars such as, Hinton (2020) and Leggio & Matras (2021) observe that in the digital age, informal bottom- up orthographies often emerge on social media such as, Youtube and Facebook before linguists finish their formal

proposals. Conversely, Hull (2021) discusses ideological distance, where communities might intentionally choose symbols that look different from a colonial or neighbouring language to assert a distinct ethnic identity, even if those symbols are harder to type on a standard keyboard.

Technical and Digital Graphization: literature has turned sharply recently toward Language Technology (LT). This has been buttressed by Berez-Kroeker et al. (2023) who emphasize that a modern orthography is useless if it lacks Unicode compatibility. If a language unique tonal markers or special vowels cannot be rendered on a basic smartphone, the orthography is likely to fail. From a different perspective, Reinders and Chong (2023) explore the role of Computer Assisted Language Learning (CALL), emphasising its contributions to language learning.

This work takes the stance that the orthographic development of Zafan must go beyond traditional phonemic conservatism in favour of a socio-technological and pedagogical pragmatism. Agreeing with the community first framework of Jones and Mooney (2020) as well as Sackett (2021), this work maintains that the ultimate validity of a script lies not in its phonetic perfection, but in its social buy-in and digital utility. In addition, this work asserts that in the context of Nigerian higher education, an orthography cannot exist in a vacuum; it must serve as a cognitive scaffold (Mahlangu, 2024) that facilitates the transfer of literacy skills to English. For that, the work advocates for a shallow yet digitally compatible orthography that prioritizes Unicode accessibility and pedagogical ease, to ensure the language is both teachable in classrooms and typable on keyboard in the digital age we are.

Stages of Orthography Development

Phonetic Documentation- the Pre-Orthographic Stage

Before writing commences, listening must come first. In phonetic documentation, the Swadesh List (1950s), developed by the American linguist Morris Swadesh is utilized to collect stable, basic vocabulary. The list essentially consists of basic vocabularies found in almost every human culture. Basic words such as water, fire, food, utensils, you, mother, I, etc., are stable and less likely to be borrowed from other languages, hence the rationale for collecting them. Phonetic documentation distributes further into broad transcription, another word for this is phonemic transcription which uses a minimal set of symbols to represent the contrastive sound units (phonemes) of a language. Broad transcription is utilised for Zafan Language. Phonemic analysis comes next where the orthographic journey of Zafan transits from physical sound to phonemic representation.

And in order to establish a functional script, we move beyond raw phonetic data to identify the functional integration of the language. Zafan phonemes are identified primarily through minimal pair test through minimal pair test which is also known as gold standard of phonological analysis. This is the process where identical word pairs are located in every phonetic environment except for a single sound in the same position. And if substituting a sound for another results in a change of meaning, those sounds are said to be in contrastive distribution hence recognized as separate mental categories (phonemes). Phonemic contrast /p/ vs /b/ in the word /pæt/ (pat) and /bæt/ (bat) for example, the initial sounds are the only difference but since the change alters the words meaning, /p/ and /b/ are said to be distinct phonemes. According to the principle

of accuracy as defined by Williams (1984), sounds that pass minimal pair test must be represented by unique symbols in the Zafan orthography because these sounds are mutually exclusive and contrastive which creates unacceptable ambiguity. In Zafan, for example, /p/ vs /b/: /paka/ (to jump) vs /baka) to sit respectively. Furthermore, since changing /p/ to /b/ changes the meaning, /p/ and /b/ are distinct phonemes. Also, /i/ vs /u/ : /tina/ (sun) vs /tuna/ (moon) also establishes phonemes /i/ and /u/.

Complementary Distribution Test: according to Pikes (1947) principle of One Grapheme per Phoneme, an efficient orthography avoids over-differentiation. Complementary Distribution is the primary test used to identify allophones predictable phonetic variants of a single phoneme. Two sounds are in complementary distribution if they never occur in the same phonetic environment; where one appears, the other is strictly excluded. Because their appearance is governed by a predictable rule, the native speaker perceives them as the same sound. A classic illustration is found in the English phoneme /l/, which has two principal allophones: The "clear" /l/: an alveolar lateral approximant used at the beginning of syllables (e.g., *leaf* /liːf]/.The "dark" /l/: a velarized lateral approximant used at the end of syllables or before a consonant (e.g., *pool* /puːɬ/). Since the "dark l" only appears in the syllable coda and the "clear L" only appears in the syllable onset, they are in complementary distribution also are phonetic partners rather than rivals.

For the Zafan orthography, sounds in complementary distribution are represented by a single grapheme (letter). As Williamson (1984) and Hyman (1975) suggest that representing allophones with different letters would create over differentiation, making the script harder to learn. By using only one symbol for the underlying phoneme, we respect the mental grammar of the speaker and

maintain orthographic economy. If two sounds never occur in the same environment, they are allophones of the same phoneme. For example, if Zafan uses /k^h/ (aspirated) only at the start of words and /k/ everywhere else, they are partners. We only write 'k'. Example: If vowels in Zafan are always nasalized before a nasal consonant (e.g., /manta/), we do not need a special symbol for /ã/; the nasalization is a predictable phonetic by product.

Free Variation is the third essential test in conducting a phonemic analysis. It occurs when two or more different sounds (phonemes) appear in the same phonetic environment interchangeably without changing the meaning of the word and without being governed by a specific phonological rule. Unlike contrastive distribution, where a sound change creates a new word; free variation is often the result of personal habit, social factors, or regional dialects. For example, in English the final voiceless alveolar plosive /t/ in the word /hat/ provides a clear illustration. Some speakers release the /t/ with a distinct puff of air (aspiration), while others do not release the air at all (unreleased). Because the word remains "hat" regardless of this phonetic difference, these two sounds are in free variation. The final /p/ in /stop/ saying it while withholding the /p/ sound in the end, while others articulate it completely. In English, we don't change the spelling for these variations because the "mental representation" (the phoneme) remains the same despite the different physical pronunciations. Example, words ending in /p, t, or k/ can be pronounced with a "burst" of air or without one.

In certain positions, especially in British or some American dialects, the phoneme /t/ can be swapped for a glottal stop. For example, the word /button/ can be realised as [bʌtən] (with an explicit alveolar stop [t] or as [bʌn], where the\ t\ undergo glottal replacement. Both pronunciations correspond to the same underlying phonemic representation. Some words in English allow for two different phonemes to be used interchangeably by the same

speaker or within the same community without creating a "minimal pair" effect. Example, the word “Economics” could be pronounced in two ways, [ˌɛkəˈnɒmɪks] (starts with /i/). [ˌikəˈnɒmɪks] (starts with "e"). given that the meaning does not change, the speaker is free to vary the sound. The word "Either", pronounced; /ˈaɪðər/ (starts with "aɪ") or /ˈiðər/ (starts with "eɪ").

This is one of the most famous examples of free variation in the English phonology. An example of free variation in Zafan where a speaker might say /zafan/ or /safan/ to mean the same thing, because the meaning remains "Zafan," the fricatives /z/ and /s/ are in free variation here. We would choose one standard grapheme (e.g., "z") for the script. The Zafan orthography would ignore free variation to keep the script simple and user-friendly. Adhering to the theories of Daniel Jones and Nikolai Trubetzkoy, this process ensures the script is neither too complex for native speakers nor too simple to capture the language's core meanings.

Morphological Mapping (Nouns and Pronouns):
Once the sounds are set, the next step is documentation of how words are built. For an unwritten language, the grammar must be extracted from discourse. The Noun Categorization System must determine if the language uses Noun Classes (common in Niger-Congo languages), gender systems or the concordial method where you observe if the noun triggers changes in adjectives or verbs. If "The big man" and "The big tree" use different words for "big," you have discovered a noun class system. The orthography must then decide how to mark these classes (usually through prefixes). In a nut shell, the linguist must consider the following; noun categorisation This is where noun classes or genders are identified

and also ensure there are prefixes or suffixes that mark plurality. The pronoun system. This is where mapping out of the person (1st, 2nd, 3rd), number (singular, dual, plural), and clusivity (inclusive "we" vs. exclusive "we") is done. In morpheme boundary marking, researchers must decide whether bound morphemes (prefixes and suffixes) should be attached directly to the root word or separated by hyphens for clarity.

Script Selection and Grapheme Assignment: the semi-final stage of development involves selection of actual symbols (graphemes). Following the Harmonization Principle advocated by Williamson (1984), this work prioritizes symbols that align with the Pan-Nigerian alphabet. Following the principle of similarity to facilitate transition literacy, we use symbols that are familiar to speakers already literate in Nigeria's major languages (Hausa, Igbo, or Yoruba). Regarding diacritic assignment, rather than inventing novel symbols for Zafan's unique sounds, sub-dots (ẹ, ọ) were utilised. This ensures that the script is technically convenient for modern keyboards and aligns with the national orthographic standards of Nigeria. This approach will reduce the learning curve for the community and as well ensures the language is perceived as having the same prestige as the national languages. Also**,** the linguist should choose the actual symbols (graphemes).

Usually, for academic and practical reasons, a modified Latin script is used where the One-to-One Principle is applied**:** ideally, one phoneme should be represented by one grapheme (e.g., using 'sh' for /ʃ/ is a digraph; using 'š' is a single grapheme). In addition**,** choosing between a shallow orthography (phonemic write exactly

as it sounds) and a deep orthography (morphophonemic preserving the look of the root word even when the sound changes).

Community Testing and Standardization

An orthography is useless if the community cannot read it. This is what brings us to Triability. This stage provides a practical testing or trial and run of the newly developed writing system within a community. Here, a short text is presented to native speakers to see if they can decode the symbols without confusion.

The Role of Linguistic Universals in Codification

In developing an orthography for a previously unwritten language, it is essential to ground the work in the theory of linguistic universals as proposed by Greenberg (1963) and later expanded by Chomsky's (1965) theory of Universal Grammar, these are structural features or patterns that are found in all human languages, regardless of their geographical or genetic origin. In this work, we prioritize absolute universal features that occur without exception. In reducing Zafan to writing, we recognize the universal constraint that all languages have at least three vowels /i/, /u/, and /a/, which we adopt Zafan as core, then proceed to more complex nasalized or long vowels.

Hockett (1958) notes that the presence of a consonant vowel (CV) syllable structure is an absolute universal; therefore, the Zafan script prioritizes the CV syllable as its foundational building block. Every language possesses a category of nouns and verbs. In the sketch of Zafan, we align with the absolute universal that if a language grammaticalizes the category of person(first person 'I/we', second person 'you', third person 'he/she/they'), it will always distinguish the first person from the others. This is a justification for including pronominal matrix in this work because it indicates that we are not just recording random words, but documenting a system that adheres to the universal human mind.

The Grammar Sketch of Zafan: At this stage, a pronominal matrix is constructed to map out the language's entire pronoun system. Many African languages feature complex pronoun systems, including clusivity (the distinction between inclusive 'we' [you and I] from exclusive 'we' [they and I, excluding you]).

Person	Singular	Dual (Two people)	Plural (Three) +
1st (inclusive)	-	'You and I'	'All of us'
1st exclusive	'I'	'He and I' (not you)	'They and I' (not you)
2nd	'You'	'You two'	'You all'
3rd	'He/she/it'	'Those two'	They'

Going by the many unwritten languages nature particularly in Africa are tonal, and to ensure Zafan orthography is not cluttered with unnecessary symbols, we apply the principle of complementary distribution and mutual exclusivity. This will allow the linguist to differentiate between meaning changing sounds and predictable variations. The rule as it is, says when two phonetically similar sounds are in complementary distribution, they are said to be allophones of the same phoneme. If sound /a/ only occurs at the beginning of words and sound/b/ only occurs at the end, they are partners. Following the principle of invariance, we only assign one letter to represent both sounds. This in turn reduces the cognitive load for the learners as Hyman (1975) posits a practical orthography should represent the phoneme (the mental idea), not every phonetic variation. Under the principle of mutual exclusivity, sounds that contrast in the same position to change meaning form minimal pairs. For example, replacing sound /k/ with /b/ changes the word from cat to bat, proving these sounds are distinct phonemes.

Orthography Design Principles: there has been a constant dilemma between linguistic accuracy (representing every sound) and practicality (making it easy to read and print). This process is now a matter of national priority in the Nigerian context. The national policy (NPL 2022) mandates the use of mother tongue as

the primary medium of instruction for six years of primary education. A renowned Nigerian linguist Williamson (1984) argues that an effective orthography must balance linguistic accuracy with harmonization ensuring it looks familiar to speakers of neighbouring or dominant languages to aid literacy transfer. Equally, Emenanjo (2015) emphasizes that codification is the only way to ensure linguistic sovereignty against intrusion of global vehicular languages.

To this end, building an effective system especially for languages with complex phonology like, those in Nigeria, According to Egbokhare (2011) and Ida Ward (1968), four core principles are applied to Zafan, which include the principle of accuracy/ consistency, harmonization, transferability and typability. The principle of typability posits that a writing system loses functionality if an orthography relies on specialised symbols that cannot be easily type on standard keyboards or mobile devices, leading the community to abandon it in favour of a simpler script.

Writing Inventory (The Alphabet)

According tp the phonemic analysis, the Zafan orthography shall use a modified Latin script which contains vowels such as /a,e,i,o,u/ (standard) as well as consonants to include the bilabial /ɓ/ and the velar /ŋ/ written as /ng/ respectively. The Zafan tone is registered, therefore, we will use the acute accent (á) for high tone and leave low tone unmarked. Consequently, in morphophonology of Zafan, nouns will undergo initial consonant mutation (to indicate plurality e.g. ɓano (house) → pano (houses). We write the mutation exactly the way it sounds to assist new learners instead of using a hidden plural marker. Additionally, Zafan exhibits nominative-accusative alignment, where pronoun forms alternate depending on whether the referent functions as the agent (subject) or undergoer (object).

Person	Subject (Actor)	Object (undergoer)
1st person singular	Mo	Ma
2nd person singular	Ze	Za
3rd person singular	Lo	La
First Plural inclusive	Ki-mo	Ki-ma
2nd plural exclusive	Su-mo	Su-ma

Zafan could follow head initial parameter. This means noun comes before modifiers (adjectives and possessives). The structure is as follows:

Noun +Adjective + possessive example**:** *ɓano zángi mó* ɓano (house) + zángi (white) + mó (my) translation 'My white house'. Zafan tense is marked by tone melodies on the verb root than the suffixes. Some examples below:

Verb root: Kara (to speak), present tense: *Ló kára* (He is speaking-with a high tone on the first syllable) and past tense: *Ló karά* (He spoke-with high tone on the last syllable).

Conclusion

Developing an orthography for a previously unwritten language involves far more than merely assigning symbols to speech sound; rather, it provides a community with a vital tool for preserving its cultural heritage. The entire work reduced a hitherto unwritten language hypothetically named Zafan. It has a systematically yet painstaking rigorous process. Underpinned by the Mother Tongue Based Educational Framework, MTB-MLE adopting its ideas of building bridge from known (Zafan) to the unknown (English Hausa), the work equally leveraged on the MTB-MLE frameworks emphasis on preserving cultural identity. The work reviewed contemporary literature in a thematic way and later outlined the comprehensive processes of phonetic documentation, phonemic analysis, morphological mapping, script selection, grapheme assignment and community testing and standardization. These processes were guided by universal patterns globally across human

languages and included principles of accuracy, familiarity as well as typability. Zafans alphabets were created based on Latin scripts adopting familiar vowels and consonant symbols alongside clear pronoun system ensuring that literacy in the mother tongue is strengthened and again, it gave the Zafan community members tools to write their own history in their L1 to ensure survival of their language round the globe.

For teachers and speakers of Zafan, this orthography is a tool to pass the language on the next generation through books, schools and text messages. And for researchers, it is a foundation to better understand and document the languages unique structure. Moving forward, the real work begins because we need linguists and Zafan speakers to team up and actually test these letters in real life classrooms and everyday conversation. By working together, we can make sure this writing system feels natural to the people using it as well as help keep Zafan strong for many years to come.

Reference

Berez-Kroeker, A. L., Gabber,S., & Slayton, A. (2023). Recent advances in technologies for resource creation and mobilization in language documentation. *Annual Review of Linguistics*, 9(1), 413-433.doi.org

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Egbokhare, F. (2011). *The sound of meaning*. Ibadan University Press.

Emenanjo, E. (2015). *A grammar of contemporary Igbo: Constituents, features and processes*. M. and J. Grand Orbit Communications.

Federal Government of Nigeria. (2022). *National Language Policy (NLP)*.

Fromkin, V. Roadman, R., & Hyams, N. (2014). *An introduction to language* (12th ed.). Cengage Learning.

Greenberg, J. H. (1963). *Universals of language*. MIT Press.

Hockett, C. F. (1958). *A course in modern linguistics*. Macmillan.

Hyman, L. M. (1975). *Phonology: Theory and analysis*. Holt, Rinehart and Winston.

Ida, W. C. (1968). *The phonetics of english*. Heffer.

Jones, M. C., & Mooney, D. (Eds.). (2020). *Creating orthographies for endangered languages*. Cambridge University Press.

Leggio, D. V., & Matras, Y. (2021). Orthography development on the internet: Romani on YouTube.

Lugh, Y. (2021). *Extinct languages in Nigeria*.

Mahlangu, K. S. (2024). Development of indigenous language orthographies: Setting up English as the torch bearer. *International Journal of Critical Diversity Studies*.

May, S. (2017). Bilingual education: What research tells us. In O. García, A. M. Y. Lin, & S. May (Eds.), *Bilingual and multilingual education* (3rd ed., pp. 3–15). Springer.

National Council for Arts and Culture (NCAC). (2020). *Reports on endangered and extinct indigenous languages in Nigeria*.

Pike, K. L. (1947). *Phonemics: A technique for reducing languages to writing*. University of Michigan Press.

Sackett, K. D. (2021). Community-driven, goal-centred orthography development: A Tsakhur case study. In M. C. Jones & D. Mooney (Eds.), *Creating orthographies for endangered languages*. Cambridge University Press.

Smalley, W. A. (1964). *Orthography studies: Articles on new writing systems*. United Bible Societies.

Snider, K. (2014). *Orthography and phonological depth*. SIL International.

Snider, K. L. (2025). Orthography design. *The Encyclopedia of Applied Linguistics*.

Trubetzkoy, N. S. (1969). *Principles of phonology*. University of California Press.

Williamson, K. (1984). *Practical orthography in Nigeria*. Heinemann Educational Books.

Yoder, B. E. (2020). *A grammar of Abawiri, a Lakes Plain language of Papua, Indonesia* [Doctoral dissertation, University of California, Santa Barbara]. eScholarship. https://escholarship.org/uc/item/4mj571tr