Journal of West African Languages

Volume 46.2 (2019)

A CORPUS-DRIVEN STUDY OF MULTIWORD EXPRESSIONS IN IGBO

Ochu, Michael Chima
University of Botswana
ochumichaelc@gmail.com

Abstract

In recent years, a number of investigations have been devoted to how computers can facilitate linguistic research. One specific area on the computer frontier which still remains quite open to exploration is corpus linguistics. This paper argues that computer-enabled linguistic approach can contribute to the study of multiword expressions (henceforth MWEs) in Igbo. We extract Igbo MWEs using concordance lines. The objective is to use concordance lines to derive MWE candidates and to reveal meanings and usages that cannot be realized through mere mental prodding. Through in-depth meaning analysis, the study demonstrates how extracted word groups qualify as MWEs. We use the Oxford Wordsmith Tools (henceforth WST) Version 6. The WST software package was developed by Mike Scott and it is published by the Oxford University Press. Data used in this study consists of a four-million-word Igbo corpus. Keywords in context (KWIC) were two parts of the human body. These are *anya* (eye) and *ọnụ* (mouth). In each case of the selected words, concordance lines are presented to illustrate, with examples, the company each query item keeps. Extracted MWEs are presented in tables. The conclusion is that, indeed, concoerdance tool is a productive approach in extracting Igbo MWEs.

Keywords: multiword expressions (MWEs), concordance lines, Igbo, *anya* (eye), and *ọnụ (*mouth)

1. Introduction

Igbo is one the three major national languages in Nigeria. Since Nigeria's independence in 1960, its three main languages have remained Hausa, Igbo, and Yoruba. Igbo is predominantly spoken in Abia, Imo, Enugu, Anambra, and Ebonyi States in the south-east Nigeria. However, Emenanjo (2005) avers that a sizeable population of south-south states such as Delta, Rivers, Cross River, Akwa Ibom, Edo, and Bayelsa also speak Igbo. That is to say that Igbo is spoken either fully or partly in eleven (11) of the Nigeria's thirty-six (36) states. Ugochukwu (2008) indicates that the Igbo-speaking area of Nigeria is vast and is one of the most densely populated of the whole central Africa. This study therefore investigates MWEs in Igbo from a corpus linguistics perspective. In other words, the focus of the current study is on words that frequently co-occur with two Igbo words *ọnụ* (mouth) and *anya* (eye). It uses Oxford WSTs to identify and extract the MWEs.

Kennedy (1998) posits that the field of corpus linguistics is broad, covering areas such as grammatical and lexical studies. Corpus-based studies have not been greatly explored in Igbo. This is not to say that there are no corpus-based studies in Igbo. However, the point is that linguists, researchers, translators, teachers and students working in Igbo have not really investigated what is typical and frequent in the language. A great deal of studies in Igbo have focused on language attitude (Adesoji-Farayibi 2017, Onwuka 2012, Okudo and Ifeagwasi 2013); language ecology (Nge-Meier and Bendor-Samuel 1975, Acholọnụ (2011); Uchechukwu 2004); phonetics and phonology (Nkamigbo 2010), Nkamigbo and Obiamalu 2016); and grammar (Ward 1936, Abraham 1967, Lord 1975, Emenanjo 1976, Ofomata 2011). Most of the afore-stated studies are not corpus-based. This therefore creates a research gap since corpus-based studies in Igbo are in the minority. The objective of this paper is to determine the effectiveness of concordance lines in extracting MWEs in Igbo.

2.1. Literature Review

The use of language involves a combination of sounds, words, phrases, clauses, and sentences. The combination could be in their predictable structured and synergistic forms. As a result, there are essentially two types of combinations. There are free combinations and fixed combinations (Sprenger 2006). Sprenger (2006) argues that fixed expressions are groups of words that have a specific meaning that cannot be expressed in any other way and also cannot at the same time be deduced just by considering the sum meaning of its parts. However, expressions that are

referred to as fixed have levels of fixedness. This means that the feature of fixedness can be fluid. Fixed combinations or fixed expressions, in the language of Moon (1998), are what both Sprenger (2006) and Baldwin (2006) refer to as MWEs. And MWEs are identified as frequent in human language. Xu, Lu, and Li (2006) define MWEs as arbitrary word combinations that are very frequent in natural language. In Xu et al. (2006), MWEs are defined as arbitrary because they do not have to adhere to rules of word patterning.

MWEs, according to Masini (2005), are lexical units that can bear both idiomatic and compositional meanings. On the one hand, when an MWE is said to bear idiomatic meaning, this means that its meaning cannot be derived from the meanings of words that make up such MWE. It is semantically non-compositional. For example, the meaning of the expression *to shoot the breeze* in Hornby (2010) does not have anything to do with a *gun* or *bullet*. Rather, it means to talk with someone about unimportant things for a long time. Therefore, when the meaning of an expression cannot be extracted from the meaning of its constituents, the expression is said to be semantically non-compositional. On the other hand, an MWE is compositional when the overall meaning is derived from the constituting words. For example, *put together* also found in Hornby (2010) means combining people or things to show that one refers to them as a group rather than separately. Clearly, the meaning of the expression is deduced from the constituents *put* and *together*. Sprenger (2006) posits that MWEs are fixed expressions that refer to specific combinations of two or more words that are used to express a specific concept. Such expressions sometimes have distinctive meanings. They are also frequent in everyday language. On the basis of their nature, there is need to extract and develop MWEs through linguistic research.

According to Kim and Baldiwn (2010), the underlying assumption in MWE extraction is that there is evidence in a given corpus that there are potential MWEs to be extracted in some context, without making any claim about whether there also exist simple compositional combinations of those same words. The motivation for MWE extraction is generally for lexicon development or expansion. Lexicon development can be in the form of recognizing such newly-formed MWEs or domain-specific MWEs. In addition, extracting MWEs is relevant to any lexically-driven application, such as grammar development or information extraction (Baldwin 2005).

In literature, MWE extractions are discussed from three different perspectives. These are statistical, linguistic, and hybrid approaches (Moiron and Tidermann 2006). However, the current study discusses only two of the three approaches. The justification is that after a thorough inquiry of the hybrid approach, this study observes that the hybrid approach appears to be a fusion of statistical and linguistic approaches. Therefore, having a sub-section on hybrid extraction approach will simply be a repetition of what has been discussed in either statistical or linguistic method. The remaining part of this section is divided into two sub-sections. Each discusses an MWE extraction method. This begins with the statistical approach to MWE extraction.

2.1.1. Statistical Extraction of MWEs

Statistical approach for extracting MWEs consists of applications of various statistical techniques such as Mutual Information (MI), WordNet, POS taggers, chunkers and parsers. According to Moiron (2005), the commonly used statistical extraction measures are MI, dice co-efficient, and log-likelihood. These statistical techniques perform various functions. To illustrate, Church and Hanks (1990) applied a variant of the MI measure to identify word associations automatically. The underlying idea of this approach is to divide the joint probability of a word pair (x, y) by the probabilities of observing *x* and *y* independently from one another. The expected result is to see a co-occurrence pattern of *x* and *y*. In Pecina (2006), there is a comparison of 84 kinds of association measures for bi-gram collocation extraction in a Czech data. The study concluded that in the Czech data, MI has the best performance. The MI statistical measure works well for word pairs. Therefore, it is hardly applicable for MWEs consisting of more than two words.

In another instance of MWE extraction, Collins (2002) experimented Mutual Expectation (ME) as an MWE extraction approach. According to Collins (2002), ME is a language independent statistically-based system that identifies and extracts multiword lexical units and combines candidate's frequency and its possibility to be a fixed phrase as the inputs of the method. In an earlier study, Dias, Gabriel, Lopes, and Gullimore (1999) compared ME with LocalMax algorithm introduced by Silva (1999). The comparative results show that ME does not only give high precision and extraction rates, but also overcomes the problem of highly frequent words raised by LocalMax and tends to elect longer MWEs. Dias et al (1999), therefore, infer that ME is very suitable for extracting MWEs.

Furthermore, Aggarwal, Wunner, Arcan, Buitelaar, and O'Riain, (2012) applied dice coefficient on the head and modifier syntactical arguments of two terms. Dice coefficient is simply the ratio of common modifiers to all modifiers of two concepts. Aggarwal et al. (2012) used a 17 million word-word of the British National Corpus World Edition (BNC). The results of the study shows that the expressions *financial income* and *net financial income* have 3 modifiers, which are *financial*, *net,* and *net financial*, whereby only *financial* is a common modifier.

Finally, Schone and Jurafsky (2001) evaluated a variety of collocation extraction approaches. The evaluated approaches include frequency, Pointwise Mutual Information (PMI), selectional association, symmetric conditional probability, Dice formula, log-likelihood ratios, Pearson's chi-square, z-score and t-test. They showed that information-like approaches, particularly z-score symmetric conditional probability, and chi-square perform better than the others but in general results. Schone and Jurafsky (2001) also proposed two new approaches namely: joint probabilities and likelihood ratios. In all of the foregoing, it is important to test the efficacy of a given measure in a specific task. The rationale is that there is no guarantee that a statistically reliable measure will yield the best results. Evert (2004: p, 113) sums it thus: "The statistical soundness of log-likelihood does not always translate into better performance. A conclusive answer can therefore only come from a comparative empirical evaluation of extraction measures''. The next sub-section discusses a number of studies that have applied linguistic approaches to MWE extraction.

2.1.2. Linguistic Approaches to Extracting MWE

There are a number of linguistic approaches to the extraction of MWEs. Attia Tounsi, Pecina, van Genabith, and Toral (2010) proposed two complementary approaches to extracting MWEs from available data resources in Arabic. The first approach is correspondence asymmetries reliance between Arabic Wikipedia titles and titles in 21 different languages. This approach is aimed at capturing semantic non-decomposable MWEs. The second approach is the collection of English MWEs from Princeton WordNet (henceforth PWN), translates the collection into Arabic using Google Translate mechanism, and utilizes different search engines to validate the output. PWN is a large lexical database of English nouns, verbs, adjectives, and adverbs grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. The database is provided by Princeton University.

The second approach used by Attia et al. (2010) is translation-based. The approach is bilingual and complements the first approach which is a cross-lingual correspondence approach. This method is partly similar to that of Vintar and Fisher (2008) who automatically extended the Slovene WordNet with nominal MWEs by translating the MWEs in PWN using a technique based on word alignment and lexico-syntactic patterns. However, Attia et al. (2010) departed from Vintar and Fiser (2008)'s extraction approach by not using a parallel corpus to find the translation. Rather, they use Google Translate. In simple terms, Attia et al. (2010) use cross lingual correspondence asymmetries and translation-based extraction.

In addition, Fazly, Cook, and Stevenson (2009) use syntactic fixedness as a means of extracting MWEs. Their study was based on the assumption that semantically idiomatic MWEs undergo levels of syntactic variation, such as passivation or other internal modifications, less readily than simple verb–noun combinations. Similarly,

Venkatapathy and Joshi (2005) use some kind of English translation for extracting Hindi MWEs. They use a classification based approach for extracting noun-verb collocations in Hindi. Their study uses the identity of the verb, semantic type of the object, case marker with the object, and similarity of the verb form of the object with the verb-object pair under consideration as features in a Max (maximum) Entropy classifier. The Max Entropy classifier is a discriminative classifier commonly used in Natural Language Processing, Speech and Information Retrieval. One advantage of the maximum entropy framework is the ability to incorporate linguistic features or clues of MWEs. This is particularly effective when the corpus is tagged. In another approach, Moiron and Tiedemann (2006) used translation ambiguity to extract non-compositional MWEs. They posit that the non-compositional collocations will have more translation candidates on account of more uncertainty in translation. This uncertainty is measured as translational entropy. Furthermore, Taljard and de Schryver (2002) report that the initial stage of the investigation of what constitutes MWEs in Northern Sotho linguistics consisted of a manual excerption of terms from the linguistic texts. Manual excerption implies close scrutiny of a text in order to identify terms which are relevant to a specific subject field; in this case, linguistics. This manual reading and marking was performed by a professional terminologist, and the terms were entered into a preliminary term list.

In summary, this section of the paper has presented literature on the statistical and linguistic approaches to extracting MWEs. From the discussion, it has been seen that statistical extraction techniques, on the one hand, compute frequency counts of words, tokens, N-grams, co-occurrences of words, and features that capture the context of instances for various context types. On the other hand, linguistic extraction techniques deal with linguistic properties of words that make up the corpus.

The section that follows focuses on the theoretical framework that has been adopted for the current study.

2.2. Conceptual Approach of the Study

The current study proposes Corpus Pattern Analysis (henceforth CPA). CPA has its theoretical basis in Theory of Norms and Exploitation (TNE) (Hanks 2013). The main claim of TNE is that 'a language consists of a constantly moving and developing double helix of rules governing linguistic behaviour: normal uses and exploitations of normal use.' (Hanks 2013: 215). CPA is a new procedural technique for mapping meaning into use (Hanks 2002); based on determining syntactic and semantic patterns of usage of verbs. It allows the discovery of how meanings arise from patterns of usage, words in context, rather than treating words as isolated elements in a compositional structure (Popescu 2014). Cinkova, Holub, Rychly, Smejkalova, and Sindlerova (2010) present CPA as a semi-formal lexical description method that consistently captures meanings in patterns of language rather than lexical units in the token-centred lexicographic tradition.

CPA captures normal, that is, reasonably frequent, usages of a given lexical item by sorting them into patterns. A pattern is some sort of construction, a representation of a cluster of similar occurrences, or structures which may 'cover a long stretch of text' (Sinclair 2004: p 24), the elements of which may be more or less fixed. In the view of Maarouf (2014), patterns can as well be considered as ordered propositional units involving syntactic relations. In simple terms, patterns are recurrent textual sequences around a key word.

The analytical procedure in this study consists of searching for two selected parts of the human body *ọnụ* (mouth) and *anya* (eye) in a corpus and selecting a random sample of concordance of lines. By analysing concordance lines extracted from a corpus, CPA guides this study in identifying lexico-syntagmatic patterns of a Key Word in Context (KWIC). This facilitates the association of each pattern of usage to a meaning potential. Each pattern is associated to a prototypical use, normal and conventional, of a word or an exploitation of the normal use, which is related to paraphrase of the pattern (Alonso and Araque 2013).

3.1. Method

This study adopts a corpus linguistics method. The main focus of corpus linguistics is to discover patterns of authentic language use through analysis of actual usage. A corpus linguistics method is based on the study of a corpus. The concept of a corpus has been explained in different ways. For example, According to Taljard (2006: 4), a corpus is an 'electronic collection of authentic texts which can be automatically or semi-automatically analysed by means of certain software'. Houston (2002: 2) opines that the recent use of the term *corpus* has been reserved for 'collections of texts or parts of texts that are stored and accessed electronically', and 'designed for some linguistic purpose'; and this specific purpose 'determines the selection of texts'. Houston's definition highlights three key features of a corpus. These are storage, accessibility, and purpose. This suggests that a corpus has to be stored in such a way that it can be accessed for use and that it is designed for a specific linguistic inquiry. However, this does not mean that collection of texts that are not in electronic form cannot be regarded as corpus or corpora as the case may be.

Additionally, Sinclair (2004: 47) defines a corpus as collection of 'pieces of language texts in electronic form selected to represent' a language for linguistic research. All the sources of the first three definitions above point out the electronic feature in what a corpus is. Also, Cabrio (2010: 9) refers to a corpus as a 'naturally occurring sequences'. This means that a corpus reflects a language use in natural settings. A corpus should be a natural and authentic language (Housen 2002: 77). On the basis of Houston's view, authenticity is embedded in a corpus. Leech (1991: 8) argues that a corpus as a 'sufficiently large body of naturally occurring data of the language to be investigated'. From the foregoing definitons of a corpus, it can be established that it is a collection of running texts usually stored in computers; and it is an authentic linguistic sample of a given language. Mikulov (2019) defines a corpus as a large collection of text produced by real users of the language and used to analyse how words, phrases and language in general are used. This definition sees corpus as an assemblage of texts that is not necessarily computerized. One therefore, sees that a corpus can exist outside the use of electronic gadgets. For the purposes of this study, a corpus consists of a databank of natural texts, compiled through the use of an electronic system from written texts and/or a transcription of recorded speech.

In the light of the definitions of corpus presented in this study, it is pertinent to succinctly state the various types of corpus that are commonly used in linguistic research. Thereafter, the type of corpus used in the current study is stated and justification for such choice is also provided. Some of the types of corpus include monolingual corpus, which contains texts in one language only; parallel corpus, consists of two monolingual corpora; multilingual corpus, contains texts in several languages which are all translations of the same text and are aligned in the same way as parallel corpora; comparable corpus, a set of two or more monolingual corpora whose texts relate to the same topic; learner corpus, a corpus of texts produced by learners of a language; diachronic corpus, a corpus containing texts from different periods and is used to study the development or change in language, among others. In this study, a monolingual corpus has been adopted. The justification is that if offers information about standard language use of a language, in the context of this study, Igbo. It is also an important source of specific expressions, technical terms, and recent borrowings. These enable the extraction of MWEs related to specific words in Igbo.

3.2. Data Collection Procedures

Data for this study was collected in three different ways. The first was by scanning books written in Igbo and saving them as plain text files. The books are largely creative works. Data obtained through this process is classified as *Creative Works*. The second way through which data was collected is by downloading Bible-based magazines and books that are obtainable from *www.jw.org*, the official website of Jehovah's Witnesses. Also, news reports from *www.igboradio.com* a website wherein news reports written in Igbo are posted, though, not so frequently were downloaded. Data elicited through downloading texts from the two websites are classified as *Religious Texts* and *Media Texts* respectively. Finally, the fourth way in

which data for this study was collected is through direct receipt from Adam Kilgariff and Kevin Scannell. Data elicited from Kilgariff and Scannell are a combination of all sorts of materials in Igbo. Data from the two sources is saved as *Miscellaneous Texts*. As highlighted earlier, the *Creative Work* data is in hard copy format. So, it was converted into soft version through scanning. This is to integrate it into the other three data sets.

In all, the corpus for this study comprises 4, 612, 538 tokens or running words in the texts, 4, 475, 607 tokens that are used for wordlist, and 119, 766 distinct words or types. The tokens of a corpus refer to a single word count, that is, the number of running words. Each successive word form is counted as one, whether or not it has occurred earlier. The number of types in a corpus refers to the number of different words in a corpus, that is, each word is counted only once.

3.3. Data Analysis

Concordance tool is used to analyse data in this study. In the view of Hunsen (2002: 47), a concordance is a 'programme that searches a corpus for selected word or phrase and presents every instance of that word in the centre of a computer screen with the words that come before and after it to the left and right.' Also, Kennedy (1998) posits that a concordance is a display of all the occurrences of tokens of a particular type in a corpus. The type is often referred to as a keyword in context (KWIC), target item, or a node. Similarly, Flowerdew (1995: 87) avers that concordance is a means of accessing a 'corpus to show how any given word or phrase in the text is used in the immediate context in which it appears'. In a nutshell, the concordance lines provide an opportunity to see the target item in context so that one can see not only the patterns in which such an item is used, but also the company it keeps in the corpus. The aim, therefore, is to study the way the identified or selected word is used. More particularly, the aim is to find terms which collocate with it. The reason why concordance tool has been chosen for this study is that it facilitates interpretation of meanings by bringing words and their contexts into closer proximity on the page. It, thus, provides a sharper focus for linguistic analysis.

4.1. Results

This section of the study presents the results and the interpretive discussion on the findings. We begin with the presentation of the results of ọnụ (mouth).

4.1.2. Extraction of *ọnụ* MWEs through concordancing

The first word that is analysed using concordance is *ọnụ* (mouth). Figure 1.1 presents the concordance lines for it.

Figure 1.1: Concordance results of the word *ọnụ*

1	Onwe-Ya ọnụ ogu ọnụ du ọnụ uzo ama onoro onochu
2	ọnụ -uzo-oma ọnụ -uzo-ama ọnụ -ulegi ọnụ -nsi ọnụ
3	n'onwunwe n'onwere n'ọnụ -uzo-ama n'onodua n'onno
4	ọnụ ghi ọnụ -uzo-oma ọnụ -uzo-ama ọnụ -ulegi ọnụ -n
5	ọnụ lu ọnụ gi ọnụ ghi ọnụ -uzo-oma ọnụ -uzo-ama ọnụ
6	afo 1995 dika anumanu noo n'ọnụ onwu nke bawanyere pacenti
7	210. Otu ogwu antibiotic di ọnụ ala ka aga ekeesara nde mma
8	iri na abua;eji otu pearl me ọnụ uzo ama o bula nime ọnụ -uzo
9	abuo nke afo 1997 nchikota ọnụ ogugu umu anu mmiri ohia ka
10	buoma 31:17)) - ka egbugbere ọnụ ha da ogbi. Chineke gagbapu
11	buoma 31:17)) - ka egbugbere ọnụ ha da ogbi. Chineke gagbapu
12	agbaputawo ayi site n'obubu ọnụ nke iwu ahu ebe oghoro obub
13	gwo ahu isi. Out okwu akpu n'ọnụ ta nwere ngbakwasi ukwu ha
14	dozi ahu. Ndi ahu jikoro aka ọnụ were ruo ya bu oru bu ndi o
15	ike. ajuju na isi okwu ajuju ọnụ Itọ ọke udaolu (J 0) * Aba

16 na uzo aka ekpe na ha gbara ọnụ ogugu ra lute lue onwe ha.
17 ere amamihe (Jen 3:1). Obubu ọnụ ahu aburu ya puru iwepu iki
18 ere amamihe (Jen 3:1). Obubu ọnụ ahu aburu ya puru iwepu iki
19 abụo na anọ. Anọ (4) bu otu ọnụ agugu. O no n'etiti atọ n'i
20 ...ogbaranikiti okwu juru ọnụ ya... Ebube n'enwe agba dị
21 chara anucha nke uku ka okwu ọnụ Gi; oru Gi ahuwokwa ya n'ny
22 anya!) * Na akuku niile Na ọnụ ego di ala * Oru bu igba *
23 anya anyi wee biarue n'otu ọnụ uzo ejiri nkpuru okwu buru
24 agaghi enye Chukwu egbugbere ọnụ Ekpere kariri mkpuru okwu.
25 anyi, ka anyi jikoo aka ọnụ ka anyi bulie asusu na omen
26 ka na arutari oru iji mee ka ọnụ ogugu ndi na-enweta ulo oru
27 ise na asaa. Asaa (7) bu otu ọnụ agugu. O no na etiti ọnụ og
28 i na asatọ. Asatọ (8) bu otu ọnụ agugu. O no n'etiti ọnụ agu
29 ga – esi baa n'uzo ahu. Otu ọnụ uzo, otu mbanye, otu inwe o
30 nwe m bu Jehova no, rida rue ọnụ uzo ama Jerusalem‖ (Maika 1
31 nwe m bu Jehova no, rida rue ọnụ uzo ama Jerusalem" (Maika 1
32 nu bu nri nke mkpuruobi. Ibu ọnụ otu ugbo ma o bu abuo kwa i
33 ipines bu nke buru nsogbu di ọnụ ego di iri ijeri dollas. RT
34 a mgbe o bula obi ojoo ya na ọnụ ojoo ya na-acho jdiwa adi n
35 ya buru ndi nwanyi ihe oke ọnụ. O no n'etiti Togo na Naige
36 Chineke were obara techie n'ọnụ uzo ha inweta nzoputa site
37 ru na I na-ebu ọnụ n'ihi ibu ọnụ, isi na I ga-emeli ya enweg
38 u na I na-ebu ọnụ n'ihi ibu ọnụ , isi na I ga-emeli ya enweg
39 a eche nche, owee noro n'ime ọnụ mmiri ozuzo amalite wee ide
40 aria ola edo oma. Ihe di oke ọnụ ahua ka obu karia ruby, ihe
41 o nchekwa ha gbadatara obere ọnụ ego di naira ise, okwuru na
42 ekpere si naani n'egbugbere ọnụ. I ga-etinyeriri Obi gi n'i
43 Otu ekpere Wepu itu egwu ibu ọnụ kpee ekpere oge niile nke U
44 Otu ekpere Wepu itu egwu ibu ọnụ kpee ekpere oge niile nke U
45 umu umu gi bulu ndi eji eme ọnụ na okpulu anyanwu!!! Gi bu
45 ye ibe ihe meerenu. Mgbe igu ọnụ isi ruru, kwu na nari isii
46 ocha nke ime muo nke di oke ọnụ ahia, nke a na apughikwa ij
47 ocha nke ime muo nke di oke ọnụ ahia, nke a na apughikwa ij
48 Israel:"M gesuye kwa oku n'ọnụ uzo ama ya nile, ogerichapu
49 Israel:—M gesuye kwa oku n'ọnụ uzo ama ya nile, ogerichapu
50 i na ano ka nsogbu-a na echi ọnụ n'ala okachasi n'ime ezinul
51 ka Chineke wee meghara anyi ọnụ uzo ikwusa okwu ya, ka anyi
52 ka Chineke wee meghara anyi ọnụ uzo ikwusa okwu ya, ka anyi
53 South East ka nwe he bu na ọnụ ogugu mmadu.Ugbu a mba Sout
54 -—Ka ndi nemebi iwu…..idebe ọnụ ha du n'ala mo' (shoel (Abu
55 "Ka ndi nemebi iwu…..idebe ọnụ ha du n'ala mo' (shoel (Abu
56 lowa alowa. Ndigbo eji eme ọnụ ndewu nu. Ya ga zie iwe aya
57 ziri unu, ma obi site n'okwu ọnụ ayi, ma obu site n'akwukwo
58 mba nile jide, obuna ijide ọnụ ọnụ ala uwe nwoke bu onye J
59 mba nile jide, obuna ijide ọnụ ọnụ ala uwe nwoke bu onye J
60 ndi mba ukwu uwa kwesiri itinye ọnụ n'okwu di mkpa gbasara
61 ndi mba ukwu uwa kwesiri itinye ọnụ n'okwu di mkpa gbasara mmeg
62 tutere inyom n'umuntakiri bu ọnụ mmiri okwu okanmuta Akua ku
63 olu Na Afrika kere enyem aka ọnụ ego di nde us Dolla Otutu u
64 ihi na inyinya di nnukwu oke-ọnụ, otutu ndi mmadu anaghi eji
65 raist site na ndu gi na okwu ọnụ gi. Ogwuwgo nke Jisos gworo
66 anyi na-ekwu. I chikota okwu ọnụ, a ga m asi na ileba okwu a
67 anyi na-ekwu. I chikota okwu ọnụ, a ga m asi na ileba okwu a

68	odu aku na uba site na mbuda ọnụ ego ugwo onwa, oke ọnụ ahua
69	eji ha were na-aga. O di oke-ọnụ mgbe ahu. Maazi Dem_people
70	eji ha were na-aga. O di oke-ọnụ mgbe ahu. Re: Nrurita Uka (
71	guru:- —Na obugh nani n'okwu ọnụ ka ozioma ayi biaruru unu,
72	akpu nafo ya ta nemezu obubu ọnụ ahu aburu agwo mbu na Jen 3
73	akpu nafo ya ta nemezu obubu ọnụ ahu aburu agwo mbu na Jen 3
74	ghari: —obu onye nezisa okwu ọnụ ya n'uwa; ngwa ngwa ka okwu
75	ghari: "obu onye nezisa okwu ọnụ ya n'uwa; ngwa ngwa ka okwu
76	di nakpom asi; Rev. 21:21 Ọnụ -uzo-ama iri na abua ahu nke
77	nyere ndi haziri oru ahu oke ọnụ. Notu aka ahu James 1:14 ne
78	Andre Uweh: Ndi oma m ji eme ọnụ, unu mere ofuma na ihe unu
79	ec 05, 2010 Ndi oma m ji eme ọnụ, unu mere ofuma na ihe unu
80	Andre Uweh: Ndi oma m ji eme ọnụ, unu mere ofuma na ihe unu
81	m/237-ngaghari-iwe-kpu-oku-n'ọnụ -n'aga-n'ihu-ugbu-a-n'istanb
82	gha * Ngaghari iwe kpu oku n'ọnụ n'aga n'ihu ugbu a n'Istanb
83	gha * Ngaghari iwe kpu oku n'ọnụ n'aga n'ihu ugbu a n'Istanb
84	kpo nwa nna ahu jua ya ajuju ọnụ nke ndi di nkpa nime ogbako
85	kpo nwa nna ahu jua ya ajuju ọnụ nke ndi di nkpa nime ogbako
86	ya. E nwere ngozi si na ibu ọnụ aputa. O na-enyere giaka na
87	BUA: E nwere ngozi si na ibu ọnụ aputa. ahu, a tshimbila na
88	ere ozioma ndia, okwu akpu n'ọnụ bu na oria obula nke anemag
89	ere ozioma ndia, okwu akpu n'ọnụ bu na oria obula nke anemag
90	ioro maka nzuputa. Otutu ndi ọnụ na-eru n'okwu n'ala anyi bi

In the concordance lines above, *ọnụ* (mouth) taken together with other collocates, is rarely used to convey a literal meaning. For example, *ajụjụ ọnụ* (literal: question mouth) is found in lines 15, 84-85 and it means *interview*; and *eme ọnụ* (literal: doing mouth) in lines 45, 56, and 78-80 is the equivalent of *boastful* or *proud of*. The meaning of the MWE, *ajụjụ ọnụ*, is compositional because an interview involves asking questions. However, *ajụjụ* (question) and *ọnụ* (mouth) are lexically fixed. The feature of fixedness qualifies *ajụjụ ọnụ* (literal: question mouth; figurative: interview) as an MWE. In the second example, *to be proud of somebody* or *something* is to have an inordinate esteem or sometimes a reasonable feeling of superiority as to talents, beauty, wealth and rank, of an individual. It might sometimes be interpreted or viewed as a disdainful behaviour or treatment of insolence or arrogance of demeanour. However, *to be proud* can more rarely have the good connotation of a sense of delight or elation arising from some act, possession, or acquisition; *to boast* means to take pride or to exult in someone or in something. As in the case of pride, *boasting* can be used in both a good and a bad sense, the usage being determined by the context. Therefore, the meaning of the MWE, *eme ọnụ* (literal: doing mouth; figurative: boastful or proud of) is not associated with the meaning of the composing lexical items, *eme* (doing) and *ọnụ* (mouth).

Other extracted MWEs are *ibu ọnụ* (literal: to carry mouth) in lines 37-38, 86-87 means *fasting*, and *jikọrọ aka ọnụ* (literal: held hand mouth) which is the figurative equivalent of *combined* or *collborate* in line 14. *Ọnụ ọgụgụ* (literal: mouth reading) means *numbering system* in lines 9, 16, 26 and 53; and *oke ọnụ* (literal: male mouth) in lines 35, 46-47, 64, 69-70, and 77 is the equivalent of *expensive, luxurious, invaluable, inestimable, precious,* or *priceless*. *Oke ọnụ* (literal: boundary mouth) has an antonymous relation with *ọnụ ala* (literal: mouth land) which means *inexpensive, low-priced,* or *cheap*. *Ọnụ ala* (literal: mouth land) is found in line 7 and it is different from *ọnụ n'ala* (literal: mouth on land) in line 50 which means *to deal with*. *Ọnụ* (mouth) collocates with itself to form *ọnụ ọnụ* (literal: mouth mouth) in lines 58 and 59 to mean *fringes, borders, peripheries, or bounds* figuratively. In addition, *ọnụ* (mouth) collocates with *ego* (money) to form the MWE *ọnụ ego* (literal: mouth money) which means *price, fee, charge, amount* or *bill* found in lines 41 and 68. The

meaning of *ọnụ ego* (literal: mouth money), *price* or *amount* is idiomatic as the idea of *price* in trading system has nothing to do with *ọnụ* (mouth) neither is it always measured in terms of money.

Some other MWEs that are related to *ọnụ* are *ọnụ ụzọ* (literal: mouth road), lines 36, 51 and 52, which means a *way* or *entrance*. There is also *ọnụ ụzọ ama* (literal: mouth road compound) in lines 1-5, 30-31, and 76, which is the equivalent of *gateway*. *Ọnụ ụzọ* (literal: mouth road) and *ọnụ uzo ama* (literal: mouth road compound) differ in usage. Whereas the former is used for an entrance into a room or an apartment, the latter is used for an entrance into a yard, compound, or homestead. A gateway can be an assembly point, a wide space, which can not only serve as an information centre, perhaps for the arrival of travellers, but also for workmen and other visitors.

Ọnụ mmiri okwu (literal: mouth water speech), line 62, is the figurative equivalent of English *expression*, not in the sense of facial expression or countenance but in that of verbal expression. *Ọnụ mmiri okwu* has the same semantic value as *okwu ọnụ* (literal: speech mouth) in lines 65 and 66, which means *expression*. Also, *okwu a kpụ n'ọnụ* (literal: speech in the mouth) is found lines 88-89, means *talk of the day*. In addition, *iju ọnụ* (literal: to fill mouth) is found in line 20 as *juru ọnụ* (literal: filled mouth). In either tense form, *iju ọnụ* (literal: to fill mouth) and *juru ọnụ* (literal: filled mouth) is idiomatically used to express *shock, surprise, speechlessness*, and *dumbfounded*. Finally, *ikpụ ọkụ n'ọnụ* (literal: bearing fire in mouth), found in lines 81-83, means *intense, or extreme*. For example, its usage in the corpus: *Ngaghari iwe kpụ ọkụ n'ọnụ n'aga n'ihu ugbu a n'Istanbul* (literal: Protest that bears fire in its mouth is going on in Istanbul) figuratively means that the on-going protest in Istanbul is *severe* or *intense*. From the results presented above, it can be seen that collocates of *ọnụ* (mouth) concordance have helped in uncovering MWEs in Igbo. Table 1.1 presents the MWEs related to the word *ọnụ* (mouth) which have been discussed.

Table 1: Concordance-derived MWEs of *ọnụ*

S/N	MWE CANDIDATE	TRANSLATION	FIGURATIVE MEANING
1	*ọnụ ọgụgụ*	mouth counting	numbering system
2	*ọnụ uzo*	mouth way	entrance
4	*ọnụ ego*	mouth money	price
5	*ọnụ uzo ama*	mouth road compound	broadway
6	*otu ọnụ*	one mouth	once or one chance/opportunity
7	*oke ọnụ*	boundary mouth	exorbitant or expensive
8	*ibu ọnụ*	carry mouth	to fast
9	*ọnụ ọnụ*	mouth mouth	Fringes, edge, or brim
10	*ọnụ efù*	vain mouth	Free or without charge
11	*izara ọnụ*	answer mouth	defend
12	*ọnụ ala*	mouth ground/low	cheap, or affordable
13	*iri ọnụ*	ate up mouth	speechless or dumb-founded
14	*imenyu ọnụ*	to squeeze mouth	to deal severely
15	*imetu aka n' ọnụ*	to touch hand on mouth	to eat
16	*ime ọnụ*	to make mouth	to be proud of or brag with
17	*ịgụ ọnụ isi*	counting mouth head	census
18	*a kpu n'ọnụ*	keeping in mouth	talk of the day
19	*oku n ọnụ*	fire in mouth	forceful
20	*itu ọnụ*	to stretch or project mouth	to brag
21	*okwu ọnụ*	speech mouth	expression
22	*ajuju ọnụ*	question mouth	interview
23	*ọnụ ọjọọ*	bad mouth	ungracious

The table above has four columns. The first column is the serial number of the extracted MWE candidates. In the second column, the extracted Igbo MWEs are presented, while the literal English equivalent of the extracted MWEs are presented. The last column is the figurative or idiomatic equivalent of the MWEs.
We present extracted *anya* (eye) MWEs in the next section.

4.1.3. Extracted *anya* MWEs through concordance lines

The KWIC is *anya* (eye). Below is the presentation of the words that hang around *anya* (eye) in the corpus. We begin by presenting the concordance lines.

Figure 1.2: Concordance results of the word *anya* (eye)

1	o Ьula nwere ntozu oke nha anya icho ikpe ziri ezi n'ihu oh
2	ke o bula inweta nchedo nha anya n'usoro iwu. Madu nile tozu
3	ke o bula inweta nchedo nha anya n'usoro iwu. Madu nile tozu
4	ke o bula inweta nchedo nha anya n'usoro iwu. Madu nile tozu
5	a-arapi,zkwa ha n'otu ntabi anya mgbe o choro ikwusuata. dib
6	Gowon (Hausa-Benue) dochie anya ya. N'iwe na ọnuma ha wéré
7	amụma. Yabụ i ga-eji nkechi anya, eriri igwe nti na ihe eji
8	i m di na ya ugbu a. M were anya m b1, 1okesa ya, o zuola. Ka
9	ya ozi: "Biko, gbagoo. Lee anya ná mpaghara oké osimiri." Y
10	fta nna ya mesiara gia ya n'anya karia wee na-akpo ya; o wee
11	a ọsọ.' Ya mere, ndị ahu na-anya ụgbọ mmiri ahu jụrụ ya, sị:
12	a nke ya na Chineke na oile anya nk odi n'iru uwa achoputawo
13	a nke ya na Chineke na oile anya nk odi n'iru uwa achoputawo
14	wasi ya anya.Na otu ntabi anya, Eze etia nkpu, dika onye a
15	si nọrọ ya n'ebe dị ezigbo anya! Ọ kpụ okwu n'ọnụ chụrụ ha
16	Ya mere obugh ihe mgbagwoju anya na odigh nrutu aka obula di
17	Ya mere obugh ihe mgbagwoju anya na odigh nrutu aka obula di
18	unu. 14 Ya adịla mgbe m ga-anya isi, ọ gwụla ma ọ bụ ịnya
19	t nurv ya ga-azu wuru wuru; anya huru ga-agba obara O bughi
20	ala, ya kpudo. O nweghi ihe anya have gbaa mmee. The o pytara,
21	a ya dika ezi onye nnochita anya bu nani uzo di nkpiri nkpi
22	8:11). Ya mere site n'olile anya nka, ayi nechere "ngbaputa
23	rugh kwa ya, onoro n'ebe di anya n'ebe nmehie di (Ndi Hibru
24	ka ndu ya ma obu ihe nnochi anya onwe ya. N'ihi na obara nke
25	zoputa ya. Ot' nime ihe ido anya nke nchikota nke nnweko n'e
26	zoputa ya. Ot' nime ihe ido anya nke nchikota nke nnweko n'e
27	emetala ya mma kama ka turu anya ya. o kpagharia ogwlI n'eJu
28	ebuli onwe ya elu n'ebe dị anya.7 Ọ bụrụ na mụ ejee ije n
29	i anya ya, kama ọ na-anọchi anya kpukpe nile na-amịpụta mkpụ
30	nya n'ime ya.+ O wee na-atụ anya ka ọ mịa ezi mkpụrụ vaịn,+
31	i anya ya, kama ọ na-anọchi anya okpukpe nile na-amịpụta mkp
32	uto‖. Ya mere obu ihe doro anya na ot'mgbe, agamu Kraist; C
33	uto". Ya mere obu ihe doro anya na ot'mgbe, agamu Kraist; C
34	, ya mere anyị kwesịrị ịtụ anya ịghọta ya n'ụzọ kasị doo an
35	ya,+ jirikwa ndị gị dochie anya ndị ya.'"+ 43 Eze Izrel wee
36	ọ bụ n'ihi ya ka m na-ahuju anya ruo n'ókè nke ịnọ n'agbụ n'
37	ọ bụ n'ihi ya ka m na-ahuju anya ruo n'ókè nke ịnọ n'agbụ n'
38	eghị ya?+ Malite na mgbe dị anya gara aga, ọ bụ ihe m ga-eme
39	eghị ya?+ Malite na mgbe dị anya gara aga, ọ bụ ihe m ga-eme
40	kweghikwan ya h nwa o muru anya. cnncwa: o b gini butere ud
41	ọdụ na ya. N'oge na-adịghị anya, ị pụrụ ịzụlite àgwà bụ́: "
42	eji were oku dika ihe anahu anya maka ahuhu nmehie. Umuntaki

a; o wee gbatịa n’otu ntabi anya, wee malite inye Chineke o
liri anya wee nọrọ n’ebe dị anya hụ ebe ahu. 5 Ebreham wee s
hineke wee na-ele ụmụ Izrel anya, Chineke wee rịba ama. e we
ahu. 17 E wee mee ka o doo anya na ala Ifrọn nke dị na Mak
ahu. 17 E wee mee ka o doo anya na ala Ifrọn nke dị na Mak
ka ha we nye ayi ihe nlere anya zuru oke nke ogbugba ndu ah
ka ha we nye ayi ihe nlere anya zuru oke nke ogbugba ndu ah
ọ mụọ nwere mmekọrịta chiri anya ruo ọtụtụ ijeri afọ—ogologo
wa ịsụ ụzọ. N’oge na-adighị anya mgbe nke ahu gasịrị, ọ nakw
iwo ọnụ ụzọ ya ike. Ọ dighị anya, mmiri wee kpuchie ndagwuru
okwu bu ụzọ meworo ka o doo anya banyere nkwa dị iche iche B
i uto nelu igwe, kama oturu anya nbilite n’onwu “n’ubochi ik
i uto nelu igwe, kama oturu anya nbilite n’onwu “n’ubochi ik
i uto nelu igwe, kama oturu anya nbilite n‘onwu —n‘ubochi ik
ru ụkwụ unu n’oge na-adighị anya Ka obiọma Onyenwe anyị Jiz
ịakwute unu n’oge na-adighị anya ma ọ bụrụ uche Jehova,+ m g
iara zie unu ndi no nebe di anya ozioma nke udo‖. As 30:1) k
la ka ebe unu ga-aga dị oké anya. Riọsierenụ m arịrịọ ike.”+
Ya mere, ụmụnna m m hụrụ n’anya, guzosienụ ike,+ bụrụnụ ndị
Ya mere, ụmụnna m m hụrụ n’anya, guzosienụ ike,+ bụrụnụ ndị
eenụ ntị, ụmụnna m m hụrụ n’anya. Ọ bụ na Chineke ahurọghị
N’ihi ya, ụmụnna m m hụrụ n’anya, ndị agụụ ha na-agụsi m ike
aịst ume ka ha ghara inyoba anya n’echiche ndị na-emebi omum
wo ụlọikwuu m,+ n’otu ntabi anya, e bibikwara ákwà ụlọikwuu
ka, ogagh abu ihe ngbagwoju anya na oge ndu-modu nogbo mbu n
ka, ogagh abu ihe ngbagwoju anya na oge ndu-modu nogbo mbu n
ka, ogagh abu ihe ngbagwoju anya na oge ndu-modu nogbo mbu n
ọdụ ụfọdụ ndị anyị pụrụ ịtụ anya ịnụ ụtọ ha. 8 Mkpughe 21:4
do n'elu uwa. 2Ebe nleghara anya na nleli ikike mmadu nwegas
do n'elu uwa. 2Ebe nleghara anya na nleli ikike mmadu nwegas
udi ihe mere, beptisiara ha anya mmiri, nodu ala, o teghi ak
uche nke ufodu ndinyom wara anya ndi nasi na nwoke na nwanyi
uche nke ufodu ndinyom wara anya ndi nasi na nwoke na nwanyi
ka Solomon rapara n’ihu na anya… ndinyom we wezuga obi ya.
ka Solomon rapara n‘ihu na anya… ndinyom we wezuga obi ya.
te aka si abịa.” N’ụzọ doro anya, itinye ntụzi Bible ahu bụ
tu ụzọ o si eme nke ahu doo anya n’Abụ Ọma 32:8, bụ ebe any
ị n’ọnụ Shiol.+ 8 Otú ọ dị, anya m dị n’ebe ị nọ,+ Jehova bụ
mmadụ,” Setan mere ka o doo anya na ebubo ya agbasaghị nanị
aria. Kọnịliọs nọkwa na-atụ anya ha, ọ kpọkọwokwa ndị ikwu y
aria. Kọnịliọs nọkwa na-atụ anya ha, ọ kpọkọwokwa ndị ikwu y
ofe sara sara. O hiri aka n’anya ihu na o gbupjaghi ike okuk
ise RT @oma_jay: Amu jopuru anya "@DFinesse: Otu N’ihi nmehi
beokwu radio nke di ma ilee anya n'aka nri bulöögia. Yaabu p
beokwu radio nke di ma ilee anya n'aka nri bulöögia. Yaabu p
ọ dị pịịị,ma dị ọcha hichaa anya ihe ọ bụla ịka kamera prọgs
zo agecheta kwa otutu nhuju anya na ndu nke nagazigh agazi;
zo agecheta kwa otutu nhuju anya na ndu nke nagazigh agazi;
ị otú ahu, anyị kwesịrị ịtụ anya ịhụ ma ọ dighị ihe ọzọ, ụf
dị otú a, inwe nkekọ chiri anya n’ime ezinụlọ pụrụ ịbụ isi
zu ụgwọ ọrụ ha.” N’ụzọ doro anya, nke a bụ nkatọ a katọrọ nd
e onye nanị jee, Pọl hụsiri anya site n’ihe ọ kọwara dị ka “
gabu onye isi na ihe nlere anya ayi ma oburu na ayi anabata

97	gabu onye isi na ihe nlere anya ayi ma oburu na ayi anabata

In the concordance lines above, *anya* (eye), taken together with other collocates, is not often used to convey a literal meaning. For example, in line 50 *chiri anya* (literal: closed or blocked eye) is used figuratively to mean *cordial, warm, amiable, pleasant,* or *intimate*. *Chiri anya* (literal: closed or blocked eye) is often used in company of *mmekorita* (relationship) or *nkekọ* (bind) to form the MWEs *mmekorita chiri anya* (literal: relationship closed or blocked eye) which means *close, intimate,* or *cordial* relationship and *nkekọ chiri anya* (literal: binding closed eyes) figuratively means or *close, affectionate,* or *intimate* bond. There is also *doo anya* (literal: settle eye) in lines 47, 48, 54, 80, and 82, which means *clarity or lucidity*. *Doo anya* (literal: settle eye) is the simple present tense form of the infinitive *ido* (to settle) found as *anya* (eye) collocate in lines 25 and 26 to form the multi-word, *ido anya* (literal: to settle eye) which is used idiomatically to convey a thought *of clarity, precision, lucidity, or simplicity*. Its past tense form, *doro anya* (literal: settled eye) is found in lines 32, 33, 79 and 94. The variations in tense do not change the meaning of the expression. Also, *lelie anya* (literal: looking eye) in lines 14 – 19 figuratively mean *neglect or disregard*; and *nlere anya* (literal: looking eye) in lines 48, 49, 96, and 97 is the figurative equivalent of *example, pattern,* or *model*. In addition, there is *nnọchi anya* (literal: blocking eyes) in line 24 is part of a longer group *onye/ndi nnochi anya* (literal: one ot those blocking eye) which is the idiomatic equivalent of *representative(s)*. There is also *ntabi anya* (literal: bite off eye) which is also part of the group *otu ntabi anya* (literal: one chop off eye) found in lines 5, 14, 43, and 67. The MWE is the equivalent of *sudden, swift, or quick*. In addition, *olile anya* (literal: looking eye) is found in lines 11, 12, and 22 which is the figurative equivalent of *hope, expectation, or anticipation*. Other *anya*-MWEs include: *itụ anya* (literal: to throw eyes) in lines 34 and 71, which means *to expect*. *Wara anya* (literal: had broken or had cut eye) in lines 75 and 76 means *bravery* or *daring* while *mgbagwoju/ngbagwoju anya* (literal: twisted eye) in lines 16, 17 and 68-70 figuratively means *confusion*. *Nhuju anya* (literal: sight to the full) in lines 90 and 91 means *tribulation, misfortune, distress, or ordeal* while *anya mmiri* (literal: eye water) in line 74 means *tears.* Table 2 below presents the *anya*-MWEs extracted in this study.

Table 2: Corpus-based derived MWEs of *anya* (eye)

S/N	MWE CANDIDATE	TRANSLATION	FIGURATIVE MEANING
1	*ihụ n'anya*	see with/in eyes	loved
2	*adịghị anya*	not being of eye	not distant or faraway
3	*mma n anya*	good in eyes	mere face value
4	*eleghị anya*	not looking eyes	probably, perhapd, probably
4	*itụ anya*	measuring eyes	expectation or anticipation
5	*ụzọ doro anya*	a way that settles the eyes	clear way/manner
6	*anya n'ihu*	eye on face	dare or challenge
7	*ntabi anya*	blink of eye	Immediately/suddenly
8	*nlere anya*	to look at	example/pattern
9	*doro anya*	settled eyes	clear
10	*aba n anya*	aba and eyes	mesmerizing
11	*inọchi anya*	blocking the eyes	representing
12	*mgbagwoju anya*	eye twisting	confusing or puzzling
13	*atụghị anya*	not weighing eyes	not expecting
14	*anya na anya*	eye and eye	face-to-face
15	*ịju anya*	to fill eyes	surprising or wonderful
16	*aka n'anya*	hand in eyes	carefulness
17	*mmekọrịta chiri anya*	dealings that block the eyes	close or intimate relationship
18	*anya nwụrụ anwụ*	dead eyes	total blindness/not redeemable

19	*anya ala*	eyes on ground	careful
20	*anya nke uche*	eyes of the mind	imagine or visualize
21	*lelie anya*	looked the eye	belittle/ignore
22	*ọcha n'anya*	shining in the eyes	typhoid fever
23	*ileda anya*	looking downwards	despise, humiliate, or ignore
24	*ihuju anya*	to see eyes to the full	suffering
25	*ilegide anya*	looking at	care
26	*juru anya*	filled up eyes	surprised
27	*ilekwasi anya*	looking on to one	look up to one
28	*ikpachara anya*	clear eyes	carefully/intentionally
29	*ịgachi anya*	to walk block eye	to be regular
30	*tụrụ anya*	threw eye	expected
31	*anya ọkụ*	in hot eyes	intimidation
32	*na-anochi anya*	blocking eyes	represen
33	*na-atu anya*	measuring eyes	expecting
34	*na-ahuju anya*	seeing	in woes, misery, and anguish

Table 2 above, as in Table 1 presented earlier, has four columns. The first column is the serial number of the extracted MWE candidates. In the second column, the extracted Igbo MWEs are presented, while the literal equivalent of the extracted MWEs are presented. The last column is the figurative or idiomatic equivalent of the MWEs.

In this section, concordance lines have been used to extract MWEs that are related to *ọnụ* (mouth) and *anya* (eye). In other words, the analysis of concordance collocates of *ọnụ* (mouth) and *anya* (eye) have revealed a number of Igbo MWEs. In the following section, a discussion of the findings of this study is presented in relation to existing literature.

4.2. Discussion

MWEs that have *ọnụ (*mouth) display a variety of linguistic structures. For example, some of them are binominal (noun + noun) or what Sinclair (2005) refers to as bigram, an n-gram for n=2 and trigram for n=3 or a three-noun structure. *Ọnụ ụzọ ama* (literal: mouth road compound) is a trigram structure in which the key word, *ọnụ (*mouth), is post modified by two other nouns, namely: *uzo* (road) and *ama* (compound). *Ọnụ ụzọ* (literal: mouth road) is a fixed binominal expression in which *ọnụ (*mouth) is post modified by only *uzo* (mouth). Another *ọnụ* -related MWE trigram found in the corpus is *ọnụ mmiri okwu* (literal: mouth water speech). A semantic inquiry of the MWEs *ọnụ ụzọ ama* (literal: mouth way compound) and *ọnụ mmiri okwu* (literal: mouth water word) shows what might be a semi-idiosyncracy. The physical *ọnụ* (mouth) is not only the *roadway* or *entrance* into the human body, but also the organ of speech. In *ọnụ uzo ama* (literal: mouth road compound), there is a partial suggestion of entrance when *ọnụ* (mouth) collocates with *uzo* (road) and *ama* (compound). Also, because *ọnụ* is used for speech production, the overall meaning of *ọnụ mmiri okwu* (speech) is semi-idiosyncratic.

Additionally, *ọnụ* (mouth) is post modified by *mmiri* (water) and *okwu* (speech). Other *ọnụ*-related noun group extracted as MWEs in this study include: *ọnụ ala* (literal: mouth land), *okwu ọnụ* (literal: speech mouth), *ọnụ ego* (literal: mouth money), *ajụjụ ọnụ* (literal: question mouth), and *ọnụ ọnụ* (literal: mouth mouth). Whereas *ọnụ* (mouth) in *ọnụ ala* (literal: mouth land) and *ọnụ ego* (literal: mouth money) are post modified by *ala* (land) and *ego* (money) respectively, *ọnụ* (mouth) is pre-modified by *okwu* (speech) and *ajụjụ* (question) in *okwu ọnụ* (literal: speech mouth) and *ajụjụ ọnụ* (literal: question mouth) respectively. In *ọnụ ọnụ (literal: mouth mouth), ọnụ* (mouth) modifies itself.

Compound nouns are classified by Sag et al. (2002) as MWEs. However, those that can vary syntactically and that can obtain their meaning from constituting lexical items are not MWEs. The compound nouns that have been extracted from the current

study are fixed in syntactic structures. They are sometimes completely or partially idiomatic. For example, *ajụjụ ọnụ* and *ọnụ ego* which mean interview and price respectively are not only fixed expressions but are also figurative. In the first pair, *ajụjụ* (question) and *ọnụ* (mouth) are partially related to *interview*. However, interview is rather more of a *consultation* than *questioning*. And the face-to-face feature of an interview session is not deduced from *ajụjụ ọnụ.* In the second pair, the individual meaning of *ọnụ* (mouth) and *ego* (money) cannot be linked to the metaphoric equivalent of *ọnụ ego* (price). Whereas *price* is occasionally paid with money, this is not in all cases. *Price* can be paid with an experience, an action, or an event. For instance, success could be said to be the price of hardwork. From the foregoing, it could be seen that the meaning of *ọnụ* (mouth) is not reflected in the sum meaning of the MWEs that are related to it. Also, MWEs in which *ọnụ* is found, are fixed and cannot be changed into *ọnụ ajụjụ* or *ego ọnụ,* for example. Otherwise, the figurative or idiomatic meaning of the expression will be lost.

In some instances, *ọnụ* (mouth) either is post modified by a larger structure or as the complement in such structure like a phrase. For example, in *ọnụ n'ala* (literal: mouth on land), *ọnụ* (mouth) is post modified by *na* (a preposition) and *ala* (land, a noun) which combine to form the prepositional phrase, *n'ala*. In some other instnaces, *ọnụ* (mouth) is the complement of the preposition *na* in *kpu oku n'ọnụ* (literal: having fire in the mouth) and *akpu n'ọnụ* (literal: having in the mouth). In these instances, the prepositional phrases in which *ọnụ* is found form the object of *kpu* and *akpu* which are the present and progressive tense variants of the infinitive *ikpu*. *Ọnụ n'ala* is a metaphor for *experiencing hardship*. It is often used in company of the verb, *ichi* (to scrub). Therefore, *ọnụ n'ala* (literal: mouth on land) becomes part of the MWE: *ichi ọnụ n'ala* (literal: to scrub mouth on land) which means to deal with someone mercilessly. The meaning of the MWE, *ichi ọnụ n'ala* does not relate to the meanings of the individual words that make up the expression. So, the expression acquires semantic idiomaticity, which qualifies it as an MWE. Similarly, *kpu oku n'ọnụ* (literal: bearing fire in the mouth) is a figurative expression that is used to indicate that something is severe. *Kpu* (bearing) is the present continuous tense form of the verb *ikpu* (to bear). Literaly, the expression, *kpu oku n'ọnụ,* means a situation where one is carrying flame in one's mouth. On the contrary, *ikpu oku n'ọnụ* (literal: bearing fire in the mouth) simply means *severe*. The syntactic patterning of the constituting lexical items is fixed. As a result, any form of re-arrangement results in ill-formedness or deviance in Igbo lexis. In terms of morphology, variations can take place which may result in *kpu* (bearing) and *ikpu* (to bear). Yet, the MWE maintains its meaning. This shows the extent to which the sum meaning of the expression deviates not only from the meaning properties of the constituting items, but also from morphological and syntactic structures of the elements.

Also, *ọnụ* (mouth) is part of a verb phrase structure where it funtions as the object of the verb *ibu* (to carry) and *eme* (doing). *Ibu* (to carry) is in its infinitive form, whereas *obubu* (carriage) is a nominalized verb or gerund. *Eme* (doing) is the progressive tense form of the *ime* (to do) infinitive. In both instances, *ọnụ (*mouth) functions as the verb complement. The *ọnụ* (mouth) headword can also have an adjective to both pre-modify and post modify it. For example, there is *oke* (boundary), which means high when used in describing *ọnụ* as a premodifier, and *ojoo* (bad) used as a post modifier of *ọnụ.* From the foregoing discussion, it is evident that MWEs related to *ọnụ* occur in a variety of linguistic structures and that the *ọnụ* (mouth) headword can assume different case positions.

As in the case of *ọnụ* (mouth), *anya* MWEs are found in different linguistic structures and environments. For example, *olile anya* (*literal: looking eye*), *mgbagwoju anya* (*literal: twisted eye), nlere anya (*literal: looking eye), *ntabi anya* (literal: chop off eye), and *nhuju anya* (literal: sight to the full) are instances in which *anya* (eye) is premodified by nominal verbs or gerunds. In the case of *ntabi anya* (literal: chop off eye) the adjective *otu* (one) is included to form an expression that is not only fixed, but also figurative. The MWE *otu ntabi anya* (literal: one chop off eye) which means a blink of the eye. It is an expression that is synonymous to *instantly*. Similarly, *nlere anya* (literal: looking anya) is part of the expression: *ihe*

nlere anya (literal: something looking eye) which is a fixed nominal trigram in Igbo that is used to convey a thought of example, model, and pattern. As an MWE, *ihe nlere anya* is a fixed expression that acannot be dismembered or distorted in any way.

Also, as found in the data, *mgbagwoju anya* (literal: twisted eye) is, in all instances, found to be premodified by a noun *ihe* (something). So, there is always *ihe mgbagwoju anya* (literal: something twists eye) which means confusion. *Ihe mgbagwoju anya* (literal: something twists eye) is semi-fixed on the basis that *mbgagwoju anya* (confusing eyes) can be modified into the infinitive verb form to have *igbagwoju anya* to realize a fuzzy MWE. From the forgoing discussion, it can be said that apart from *olile anya*, the rest of *anya*-MWEs premodified by a gerund are trigrams.

Furthermore, a number of Light Verb Constructions (LVCs) are found in *anya*-related MWEs. For example, *ido anya* (literal: to settle eye), *na-anochi anya* (literal: blocking eye), *na-atu anya* (literal: throwing eye), *na-ahuju anya* (literal: seeing full eye), *itu anya* (literal: to throw eye), *huru n'anya* (literal: seen in eye), and *wara anya* (literal: had broken eye). Whereas *ido* (to settle) and *itu* (to throw) are verbs in their infinitive forms; *na-anochi* (blocking), *na-atu* (throwing), and *na-ahuju* (seeing full) are verb forms in the progressive tense; and *huru* (saw/seen) and *wara* (brok/broken) are the past tense forms of the respective verbs. The corpus also shows instances of *doro anya* (literal: settled eye) and *turu anya* (literal: thrown eye) which are respective past forms of *ido* (to settle) and *itu* (to throw) infinitives. Verbs in all instances of infinitive, progressive, and past tense forms take *anya* (anya) as their objects to verb phrases as MWEs.

Finally, in *mmekorita* (relationship) and *nkeko chiri anya* (literal: bond closed eye) the queried term (*anya*) also functions as the object of the verb in the past tense to form a verb phrase that is premodified by the nouns, *mmekorita* (relationship) and *nkeko* (bond). It is those combinations that produce MWEs such as *mmekorita chiri anya* (literal: relationship that closed eye) and *nkeko chriri anya* (literal: bond the closes eyes) which means *close and tight or intimate relationships respectively.* Also, whereas *anya* (eye) functions as the object of the preposition *na* (on) in a noun premodified prepositional phrase: *aka n'anya* (literal: hand on eye), a*nya* is post-modified by a prepositional phrase, *n'aka* (on hand). In summary, *anya* (eye) MWEs are not restricted to a specific linguistic environment. However, the corpus reveals that MWEs related to *anya* (eye) can be found in a number of linguistic environments.

4.3. Conclusion

This paper has demonstrated that concordance tool is an effective approach to MWE extraction in Igbo. Two parts of the human body were used as key words. Each line of the concordance contains an instance of the selected word, and the words that appear on either side of the word. These were examined through linguistic interpretation. Through such linguistic analysis corpus patterns were identified as MWEs.

Some *ọnụ* -MWEs that were extracted through the concordance programme include: *ọnụ ọgụgụ* (literal: mouth counting; figurative: numbering system), *ọnụ ụzọ* (literal: mouth road; figurative: entrance), *ọnụ ego* (literal: mouth money; figurative: price), *ọnụ ụzọ ama* (literal: mouth road compound; figurative: broadway), *otu ọnụ* (literal: one mouth; figurative: once or an opportunity), *oke ọnụ* (literal: male mouth; figurative: expensive, exhorbitant), *ibu ọnụ* (literal: to carry mouth; figurative: to fast), *ọnụ ọnụ* (literal: mouth mouth; figurative: edge, fringes/brim), *ọnụ efu* (literal: mouth vain; figurative: without charge free), *izara ọnụ* (literal: to answer mouth; figurative: to defend), *ọnụ ala* (literal: mouth land; figurative: affordable, cheap), *iri ọnụ* (literal: to eat mouth; figurative: surprise, speechless/dumbfounded), *imenyu ọnụ* (literal: to quench mouth; figurative: to deal with), *imetu aka n'ọnụ* (literal: to touch hand on mouth; figurative: to eat), *ime ọnụ* (literal: to do mouth; figurative: to be proud), *igu ọnụ isi* (literal: to count mouth head; figurative: census), *oku n'ọnụ* (literal: fire on mouth; figurative: intense, severe, extreme, or forceful), *itu ọnụ* (literal: to stretch mouth; figurative: to brag), *okwu ọnụ* (literal: talk mouth;

figurative: expression), *ajuju ọnụ* (literal: question mouth; figurative: interview), and *ọnụ ojoo* (literal: mouth bad; figurative: ungracious).

In addition, the following MWEs were extracted when *anya* (eye) was selected as the KWIC in the concordance search: *eleghị anya* (literal: not looking eye; figurative: perhaps, probably), *itụ anya* (literal to throw eye; figurative: anticipation, expectation), *ụzọ doro anya* (literal: way settle eye; figurative: lucid, simple, clarity), *anya n'ihu* (literal: eye on face; figurative: dare, taunt, challenge), *ntabi anya* (literal: chop off eye; figurative: quickly, swiftly, suddenly), *nlere anya* (literal: looking eye; figurative: example, model, pattern), *Aba n'anya* (literal: Aba in eye; figurative: mesmerize), *inọchi anya* (literal: to block eye; figurative: to represent), *mgbagwoju anya* (literal: twisting eye; figurative: puzzling, baffling, confusing), *atụghị anya* (literal: not throwing eye; figurative: not expecting, not anticipating), *anya n'anya* (literal: eye in eye; figurative: face-to-face), *iju anya* (literal: to fill eye; figurative: surprising, shocking), *aka n'anya* (literal: hand in eye; figurative: cautious, watchful, carefulness, sometimes, slap or smack), *mmekọrịta chiri anya* (literal: realionship block eye; figurative: intimate or affectionate relationship), *anya ala* (literal: eye land; figurative: carefulness), *anya nke uche* (literal: eye of mind; figurative: imagination), *lelie anya* (literal: looking eye; figurative: belittle, demean, disparage, ignore), *ọcha n anya* (literal: shining in eye; figurative: typhoid fever), *ileda anya* (literal: to look down eye; figurative: to dispise), *ihuju anya* (literal: to see full eye; figurative: to suffer, to grieve), *ilegide anya* (literal: to keep looking; figurative: to care for someone or sometimes to look up to someone for help), *iju anya* (literal: to fill eye; figurative: surprise, wonder, amaze), *ilekwasi anya* (literal: to look eye on top; figurative: to look up to someone), *aka ná ntị* (literal: hand on ear; figurative: to warn, to cautuion, to advise), *ikpachara anya* (literal: to scrape off eye; figurative: carefully or intentionally), *ịgachi anya* (literal: to go block eye; figurative: to be regular or consistent), *tụrụ anya* (literal: measured eye; figurative: expected), and *anya ọkụ* (literal: eye fire; figurative: intimidation or coersion). Consistent with the objective of this study, the conclusion is that, indeed, concordance tool is a productive approach in extracting Igbo MWEs.

References

Abraham, R. C. (1967). *The principles of Ibo*. Ibadan: Institute of African Studies, University of Ibadan, Occasional Publication No. 4.

Acholọnụ, C. (2011). *West-African origin of language: The case of Igbo–discussions on a New York Times article on the West African origin of modern languages*. Presented at the Conference on Indigenous Knowledge and the Challenges of the 21st Century, Institute of African Studies, UNN, Nsukka.

Adesoji-Farayibi, C. (2017). *Attitude towards Igbo Language in high schools in Nigeria. Insight into Nigerian indigenous language issues.* Lagos: Lambert Academic Publishing (LAP).

Aggarwal, N., Wunner, T., Arcan, M., Buitelaar, P., and O'Riain, S. (2012). *A Similarity Measure based on Semantic, Terminological and Linguistic Information,* In: In Reuse and Adaptation of Ontologies and Terminologies Workshop at 18th EKAW.

Alonso, C., and Araque, I. (2013). Corpus Pattern Analysis in determining specialised uses of verbal lexical units. *Terminàlia núm* (7) pp. 26-33.

Attia, M., Tounsi, L., Pecina, P., van Genabith, J., and Toral, A., (2010). *Automatic extraction of Arabic multiword expressions.* COLING 2010 Workshop on Multiword Expressions: from Theory to Applications. Beijing, China

Baldwin, T. (2005). The deep lexical acquisition of English verb-particles. *Computer, Speech, and Language, (Special Issue on Multiword Expressions)* 19.398–414.

Baldwin, T. (2006). *Compositionality and multiword expressions: Six of one, half a dozen of the other?* In proceedings of the Workshop on Multiword Expressions: expressions: A pain in the neck for NLP. In CICLing, (pp. 1–15).

Cabrio, E (2010). *Component-based textual entailment: A modular and linguistically-motivated framework for semantic inferences*. Ph.D. Thesis, Dept. of Information Engineering and Computer Science, University of Trento, Italy.

Church, K. and Hanks, (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics,* 16(1):22–29.

Cinkova, S., Holub, M., Rychly, P., Smejkalova, L., and Sindlerova, J. (2010). *Can corpus pattern analysis be used in NLP?* Springer-Verlag Berlin Heidelberg.

Collins, M. (2002). *Ranking algorithms for named-entity extraction.* Proceedings of ACL-2002

Dias, G., Lopes, G., and Guillore, S. (1999). *Normalising the IJS-ELAN Slovene-English parallel corpus for the extraction of multilingual terminology*. In: Monachesi, P. (ed.) Proceedings of the CLIN '99 Computational Linguistics in the Netherlands.

Emenanjo, E. (1976). Auxiliaries in Igbo syntax: A comparative study. *Studies in African Grammatical System* Vol 302 (2), 41-56.

Emenanjo, E. N. (2005). *ICT, globalization, and the future of human language*. In Ndimele (ed.) Globalization & the study of languages in Africa. Port Harcourt: Grand Orbit Communications & Emhai Press, 1-18.

Evert, S. (2004). The Statistics of Word Co-occurrences: Word Pairs and Collocations. Unpublished doctoral dissertation, Institut fur maschinelle Sprachverarbeitung, University of Stuttgart.

Fazly, A and Stevenson, S. (2007). *Distinguishing subtypes of multiword expressions using linguistically-motivated statistical measures.* Proceedings of the Prague Workshop on a Broader Perspective on Multiword Expressions, pp. 9–16. Association for Computational Linguistics

Flowerdew, J. (1995). The discourse of colonial withdrawal: A case study in the creation of mythic discourse. *Discourse and Society* 8, pp 453-477.

Hanks, P. (2002). *Mapping Meaning onto Use*. In: CORREARD, Marie Hélène (ed.). Lexicography and Natural Language Processing: a Festschrift in honour of B. T. S. Atkins. United-Kingdom: Euralex, Göteborg University, p. 156-198.

Hanks, P. (2013). *Lexical Analysis: Norms and Exploitations*. Massachusetts: The MIT Press.

Hornby, A. (2010). *Oxford Advanced Learner's Dictionary*. Oxford: Oxford University Press.

Housen, A. (2002). *A corpus-based study of the L2-acquisition of the English verb system.* In: Granger, S.; Hung, J.; Petch-Tyson, S. (Ed.). Computer Learner Corpora, Second Language Acquisition, and Foreign Language Teaching. Amsterdam: John Benjamins.

Hunston, S. (2002). *Corpora in applied linguistics.* Cambridge: Cambridge University Press.

Kennedy, G. (1998). Preferred ways of putting things with implications for language teaching. *Journal of Contemporary Issues in Education*, Vol, (43) (3) 335-373.

Kim, S. N. and Baldwin, T. (2010). How to pick out token instances of English verb-particle constructions. *Natural Language Engineering*, 10(2), pp. 97–113

Leech, G. (1991). *Corpora and theories of linguistic performance*. In J. Svartvik (Ed.). 105–122.

Lord, C. (1975). Igbo verb compounds and the Lexicon. *Studies in African Linguistics*.6 (1) p. 23-48

Maarouf, I. (2014). *Methodological aspects of corpus pattern analysis.* Research Institute in Information and Language Processing (RIILP), University of Wolverhampton, pp. 119-147.

Masini, F. (2005). Multi-word expressions between syntax and the lexicon: The case of Italian verb-particle constructions. *SKY Journal of Linguistics* 18: 145-173.

Mikulov, C. (2019). Crpus for lexicography and lexical computing. Corpora Digest, Vol 142, Issue 3, pp, 11-14.

Moiron, B. (2005). *Semantics-based multiword expression extraction*. MWE '05 Proceedings of the Workshop on a Broader Perspective on Multiword Expressions, pp. 25-32. Prague, Czech Republic — June 28 - 28, 2007. Association for Computational Linguistics.

Moiron, B. and Tiedemann, J. (2006). *Identifying idiomatic expressions using automatic word alignment.* In Proceedings of the EACL 2006 Workshop on Multi-word-expressions in a multilingual context, Trento, Italy.

Moon, R. (1998). *Fixed expressions and idioms in English. A corpus-based approach.* Oxford: Clarendon Press

Nge-Meier, P., and Bendor-Samuel, J. (1975). *A grammar of Izzi: An Igbo language variety*. Summer Institute of Linguistics publication. Norman: University of Okhlahoma Press, 1975.

Nkamigbo, C. (2012). Phonology in teacher education in Nigeria: The Igbo language example. *African Journal of Teacher Education.* Vol. No.1.

Nkamigbo, C. and Obiamalu, G. (2016). Phonological differences and the phonological Problems of Igbo Learners of Chinese. *Journal of Humanities and Social Science*, Vol. 21, (4), pp 52-61

Ofomata, C. (2007). *Utoasusu Igbo waka sinio sekondiri. Enugu*. Format Publishers (Nig).

Okudo, A., and Ifeagwasi, P. (2013). Sociolinguistic functions of Igbo language: Implications for education and mass communication. *Mediterranean Journal of Social Sciences*. Vol 4 No 11.

Onwuka, G. T. (2012). Factors affecting students' performance in Igbo language senior secondary school certificate examinations. *Institute of Education Journal*, Vol. 23. No.1. 86-97.

Pecina, P. (2006). *An extensive empirical study of collocation extraction methods.* In Proceedings of the ACL Student Research Workshop, pages 13–18, Ann Arbour, MI.

Popescu, O. (2014). *Learning corpus patterns using Finite State Automata.* FBK. Trento, Italy.

Sag, I., Baldwin, T., Bond, F., Copestake, A., and Flickinger, D. (2002). *Multi-word expressions: A Pain in the Neck for NLP*. LinGO Working Paper No. 2001-03. Stanford University, CA.

Schone, P., and Jurafsky, D. (2001). *Is knowledge-free induction of multiword unit dictionary headwords a solved problem?* In Proceedings of the 6th Conference on Empirical Methods in Natural Language Processing (EMNLP 2001), Hong Kong, China, pp. 100–108.

Schryver, G. (2002). Web for/as corpus: A perspective for the African Languages. *Nordic Journal of African Studies.* 11. 282-282.

Silva, J. (1999). *A local maxima method and a fair dispersion normalization for extracting multi-word units from corpora.* In Proceedings of the Sixth Meeting on Mathematics of Language (MOL6), pages 369–381, Orlando, FL, USA

Sinclair, J. (2003). *Reading concordances*. London, New York: Pearson Education.

Sinclair, J. (2004). *Looking Up: An account of the COBUILD Project in Lexical Computing*. London: Collins.

Sprenger, S. (2006). *Fixed expressions and the production of idioms*. Ph.D. dissertation. Nijmegen: Max Planck Instituut voor Psycholinguïstiek.

Taljard, E. & De Schryver, G-M. (2002). Semi-automatic term extraction for the African languages, with special reference to Northern Sotho. *Lexikos*, 12, 44–74.

Taljard, E. (2006). Corpus-based linguistic investigation for the South African Bantu languages: a Northern Sotho case study. *South African Journal of African Languages* 26(4). 14-27.

Uchechukwu, C. (2004). The cognitive basis of the Igbo verb cluster: The ontology of the –gbá cluster. In Valery Solovyev/Vladimir Polyakov (ed.), *Text Processing and Cognitive Technologies.* MISA, Moscow, 256-269.

Ugochukwu, F. (2008). Igbo dictionaries: A showcase for an open country. *Africana*, 2(1) pp. 1–32.

Venkatapathy, S. and A. Joshi (2005). *Using information about multi-word expressions for the word-alignment task.* In Proceedings of the COLING/ACL 2006 Workshop on Multiword Expressions: Identifying and Exploiting Underlying Properties, Sydney, Australia.

Vintar, S., and Fisier, D. (2008). *Harvesting multi-word expressions from parallel corpora.* In Proceedings of LREC, Marrakech, Morocco.

Ward, I. (1936). An Introduction to the Igbo Language. Cambridge: W. Heffer & Sons.

Xu, R., Q. Lu, and S. Li (2006). *The design and construction of a Chinese collocation bank.* In Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006), Genoa, Italy.