



Language Specific Peculiarities Document for IGBO as Spoken in NIGERIA

1. Special handling of dialects

1.1 Overview

Igbo is a Niger-Congo language (Volta-Niger) spoken in south-eastern Nigeria by approximately 25 million people. There are at least 30 different dialects of Igbo; descriptions vary widely as to the number depending on the definition applied. To cite Nwaozuzu (2008, p iv), “...almost every community living as few as three kilometres apart has its few linguistic peculiarities. If these tiny peculiarities are isolated and considered to be able to assign linguistic independence to each of these communities, we shall therefore be boasting of not less than one thousand languages in what we now know as the Igbo language.”

Despite the complexity of the task, Nwaozuzu (2008) classifies the Igbo language into the following groups of dialects based on phonological, morphological and lexical features: West Niger, East Niger, East Central, Cross River, North-Eastern, South-Eastern, South-Western and Northern. We have used this classification to help determine the choice of dialects for collection (see section 1.2 below).

1.2 Written and spoken Igbo – standard and dialects

1.2.1 Igbo dialect collection

Taking into consideration the logistical considerations in the field, the degree of dialectal variation in the spoken language (outlined below), and mutual intelligibility considerations, three dialects have been collected for this project: Owerri, Onitsha and Ngwa. The table below shows the dialects collected for the speech database.

Dialects	Dialect Groups	Cities (States)	Region
Owerri	East Central	Owerri (Imo State)	Imo State and parts of Abia State (Umuahia North/South)
Onitsha	East Niger	Onitsha (Anambra State)	Anambra State and parts of Abia State (Umu-Nneochi)
Ngwa	South-Eastern	Ngwa (Abia State)	Abia State (Ukwa East/West, Isiala Ngwa North/South, Osisioma Ngwa, Obioma Ngwa, Aba South/North)

1.2.2 Standard Igbo and spoken dialect variation

Standard written Igbo is a variety developed by the Society for Promoting Igbo Language and Culture (SPILC hereafter) on the basis of Owerri/Umuahia (East Central Group) and other dialects in Abia state and Imo state (Ikekeonwu, 1991). It is not spoken by any speech community, but is used in academic works, text books, education and standardized examinations. The Owerri and Umuahia dialects supply the bulk of the lexical items and

grammatical standards in standard written Igbo (Oraka, 1983). However, standard Igbo orthography does not reflect pronunciation peculiarities specific to the Owerri or Umuahia dialects, e.g., the nasalized and aspirated consonants found in these two dialects are not represented (Ugorji, 2005; Ohiri-Aniche, 2007).

There is no agreed-upon standard spoken variety of Igbo, in the way that the Oyo dialect is considered standard Yoruba, and the Kano dialect is considered standard Hausa (Arokoyo, 2012; Newman et al., 2001). Williamson (1972) and Nwachukwu (1983) suggest that the Onitsha dialect¹ is the closest Igbo comes to a spoken standard.

Onitsha (East Niger dialect group) is often used as a lingua franca among speakers of different dialects because of its relatively limited and simple phonological features (Williamson, 1972). However Onitsha diverges from standard written Igbo in a number of respects. For example, it has distinctive lexemes and morpho-phonological features (cf. Clark, 1990).

Ngwa (South-Eastern dialect group) is a dialect that is mutually intelligible with Owerri, Umuahia and Onitsha (Clarla, 2013; Appen Igbo linguist consultants, personal communication), but which differs from both Owerri and Onitsha dialects in the set of phonemic distinctions it makes (see Igbo Phone Chart in Section 7).

Although Owerri, Onitsha and Ngwa are mutually intelligible, they have a significant number of different phonological and morphological features and lexemes (Nwaozuzu, 2008). Owerri is known for more complex phonology such as palatalization, nasalization, and aspiration whereas these phonological features are absent in Onitsha, which has the simplest phoneme inventory. Ngwa lacks palatalization and exhibits less nasalization and aspiration than Owerri. Differences in the pronunciation of standard Igbo lexical items in these regional dialects can be seen in the following table, together with some dialect-specific lexical items.

Standard Written Igbo	Owerri	Ngwa	Onitsha
ocha 'cleanliness'	ocha	uchha	ocha
ndu 'life'	ndhu	ndhu	ndu
kpu 'drag'	ḳp̣u	kpu	kpu
atu 'buffalo; example; chewing stick'	ãtu	athu	atu
eŋe 'clothes'	eŋe	nweyi /eŋe	aŋe/ave
ihu 'face'	ih̃wu	ihwu/ĩhu	iru
unyaahu 'yesterday'	echi gara aga	echi gara aga	nnyaafu
oru 'work'	õru	õru	olu

¹ Onitsha refers to varieties of the Onitsha dialect such as Otu-Onitsha and Enu-Onitsha. Enu-Onitsha is a dialect spoken by indigenous speakers of Onitsha in a sub-area of the town of Onitsha called "Onitsha Inland Town", while Otu-Onitsha is the levelled variety of Onitsha spoken in the "Waterside" (or market) area of the town of Onitsha.



Standard Written Igbo	Owerri	Ngwa	Onitsha
ivu 'load'	ivu	ivu	ibu
gaala 'has gone'	gaana	gaala	jego

Dialectal differences such as those described above can be found reflected in writing, especially in literary writing (Ugorji, 2005) even though they deviate from the standard orthography. Appen used a combination of automated and manual checks to find and correct non-standard spellings. In cases where there is no standard spelling for a dialect-specific lexical item, Appen determined a single spelling to be used in consultation with its language experts.

2. Deviation from native-speaker principle

Only native speakers of Igbo were collected for the speech database.

3. Special handling of spelling

Since 1854, the Igbo language has been written in three different orthographies (Onukawa, 2011). Modern spelling is ostensibly standardized in the Onwu script, but, given the dialectal variation that exists, standardization of spelling is still an issue. In addition, Onukawa (p.3) comments that:

... most Igbo do not show any interest in their language. As far as they are concerned, Igbo should be a language of “anything goes”. It does not matter how the language is spoken and written. Again, most of the Igbo are neither aware of the existence of the standard Igbo dialect (Igbo Izugbe) nor an acceptable orthography.

The Society for the Promotion of Igbo Language and Culture (SPILC) had in the past made recommendations on spellings especially on word boundaries, morphemes/words that should be written together and ones that should be written separately (Emenanjo et al., 1990). Our standardization approach is based on these recommendations and will use the Onwu script.

Code-switching is found in native Igbo utterances in the database; speakers primarily used English and Nigerian Pidgin words and grammar.

Igbo can be considered a morphologically agglutinative language where grammatical functions are expressed by affixation, and therefore the morphology-syntax boundary is difficult to draw. To standardize some of the most frequent function words and morphemes, the spelling of pronominal and verbal prefixes, and enclitics has been defined as follows:

- <e>/<a> is treated as a pronoun and thus written separated from the verb when there is no other adjacent pronoun
- <e>/<a> is treated as a subject marker and thus written attached to the verb when there is a pronoun following the verb

- <i> is treated as an infinitive, and it is written attached to the verb when preceded by a pronoun. If there is no pronoun preceding <i>, then <i> itself is presumed to be the pronoun and thus written separately from the verb
- <kwa> and its allomorphs are written attached to the verb only in cases where there is a subject marker at the beginning of the verb, and certain inflectional affixes at the end of the verb (following <kwa>)
- <na> as a verbal auxiliary (as opposed to a conjunction) is written with a hyphen to the following verb (which always begins with <e>, <a>, or <i>).

4. Description of character set used for orthographic transcription

The Onwu alphabet uses mostly Latin characters, with 8 dotted characters (four lower case and four upper case). The Unicode code range is thus: Standard Latin 0042-007A, plus punctuation, plus the dotted characters 1ECA-1ECD, 1E44-1E45, 1EE4-1EE5. The dotted characters (i/ı, o/ö, u/ü, ñ/ Ñ) are contrastive as they correspond to distinctive phonemes in the phone inventory. However, due to a general lack of phonological and orthographic awareness, the dotted characters are frequently incorrectly used or substituted by the corresponding undotted characters (Ojukwu et al., 2011). For this reason, the diacritic characters shown in the table below are not represented in the transcription.

Unicode codepoint	Name of Letter	Onwu	IPA	ABH XSAMPA
1ECB	Latin Small Letter I With Dot Below	ı	ɪ̣	ı
1ECA	Latin Capital Letter I With Dot Below	İ	ɪ̣	ı
1E45	N With Dot Above, Latin Small Letter	ñ	ɲ	N
1E44	N With Dot Above, Latin Capital Letter	Ñ	ɲ	N
1ECD	O With Dot Below, Latin Small Letter	ö	ɔ̣, ɒ̣	O
1ECC	O With Dot Below, Latin Capital Letter	Ö	ɔ̣, ɒ̣	O
1EE5	U With Dot Below, Latin Small Letter	ü	ʊ̣	U
1EE4	U With Dot Below, Latin Capital Letter	Ü	ʊ̣	U

Similarly tones can be represented orthographically by, for example, acute, grave and macron for high, low and downstepped tone respectively. However, since tone-marking is not a usual practice nor represented in the Onwu script, tone markers do not appear in the transcription.

The following is the list of characters that appear in the orthographic transcription:

Unicode	Igbo upper case	Unicode	Igbo lower case
U+0041	A	U+0061	a
U+0042	B	U+0062	b
U+0043	C	U+0063	c
U+0044	D	U+0064	d
U+0045	E	U+0065	e
U+0046	F	U+0066	f
U+0047	G	U+0067	g
U+0048	H	U+0068	h
U+0049	I	U+0069	i
U+004a	J	U+006a	j
U+004b	K	U+006b	k
U+004c	L	U+006c	l
U+004d	M	U+006d	m
U+004e	N	U+006e	n
U+004f	O	U+006f	o
U+0050	P	U+0070	p
U+0051	Q	U+0071	q
U+0052	R	U+0072	r
U+0053	S	U+0073	s
U+0054	T	U+0074	t
U+0055	U	U+0075	u
U+0056	V	U+0076	v
U+0057	W	U+0077	w
U+0058	X	U+0078	x
U+0059	Y	U+0079	y
U+005a	Z	U+007a	z

5. Description of Romanization scheme

None.

6. Description of method for word boundary detection

Words are delimited by whitespace.

7. Table containing all phonemes in the stipulated notation

The phonemic transcription of the words in this database uses X-SAMPA symbols, which can be found at <http://www.phon.ucl.ac.uk/home/sampa/x-sampa.htm>. The total number of phonemes is 78. There are 69 consonants, (of which 3 are also syllabic) and 9 vowels (all monophthongs). There are 3 tones, which are not marked in the lexicon (see Notes on Tone below). There are no foreign phonemes in this Phone Chart, although foreign words are not nativized in the orthography, they are phonologically adapted, with the exception of syllable structures. The chart below is based on various sources such as Ikekeonwu (1985; 1991),



Nwaozuzu (2008), Mbah & Mbah (2000), Nwosu (2013) and Ugorji (2009), and includes dialectal variations.

IGBO PHONE CHART

TYPICAL IGBO CORRES- PONDENCE	UNICODE	IPA	Appen SAMPA			COMMENTS
			Owerri	Onitsha	Ngwa	
CONSONANTS						
b	0062	b	b	b	b	
		b ^h	bh		bh	aspirated
bi	0062 0069	b ^j	bj			palatalized
ch	0063 0068	tʃ	tS	tS	tS	
		tʃ ^h	tSh		tSh	aspirated
		tʃ̃	tSn			nasalized
d	0064	d	d	d	d	
		d ^h	dh		dh	aspirated
di	0064 0069	d ^j	dj			palatalized
f	0066	f	f	f	f	
fi	0066 0069	f ^j	fj			palatalized
g	0067	g	g	g	g	
		g ^h	gh		gh	aspirated
		g̃	gn			nasalized
gi	0067 0069	g ^j	gj			palatalized
gb	0067 0062	g̃b b	gb	gb (Otu- Onitsha) b_< (Enu- Onitsha)	gb	

TYPICAL IGBO CORRESPONDENCE	UNICODE	IPA	Appen SAMPA			COMMENTS
			Owerri	Onitsha	Ngwa	
gh	0067 0068	ɣ	G	G (Otu-Onitsha) /j/ before front vowel, /w/ before back vowel (Enu-Onitsha)	G	
gw	0067 0077	g ^w	gw	gw	gw	labialized
		g ^{wh}	gwh		gwh	labialized aspirated
h	0068	h̥	H	H	H	
		h̃	Hn		Hn	nasalized
hi	0068 0069	h ^j	Hj			palatalized
j	006A	dʒ	dZ	dZ	dZ	
		dʒ ^h	dZh		dZh	aspirated
		dʒ̃	dZn			nasalized
k	006B	k	k	k	k	
		k ^h	kh		kh	aspirated
ki	006B 0069	k ^j	kj			palatalized
kp	006B 0070	\overline{kp} p̥	kp	kp (Otu-Onitsha); p_ _{<} (Enu-Onitsha)	kp	
kw	006B 0077	k ^w x ^w	kw	kw (Otu-Onitsha), xw (Enu-Onitsha)	kw	labialized
		k ^{wh}	kwh		kwh	labialized aspirated
l	006C	l	l	l	l	
li	006C 0069	l ^j	lj			palatalized
m	006D	m	m	m	m	
		m̩	m=	m=	m=	syllabic nasal

TYPICAL IGBO CORRES- PONDENCE	UNICODE	IPA	Appen SAMPA			COMMENTS
			Owerri	Onitsha	Ngwa	
mi	006D 0069	m ^j	mj		mj	palatalized
n	006E	n	n	n	n	
		ɲ̥	n=	n=	n=	syllabic nasal
		ɲ ²	N	N (Otu-Onitsha) Nw (Enu-Onitsha)	N	
		ɲ ³	N=	N=	N=	syllabic nasal
nw	006E 0077	ɲ ^w	Nw	Nw	Nw	
ny	006E 0079	ɲ	J	J	J	
p	0070	p	p	p	p	
		p ^h	ph		ph	aspirated
pi	0070 0069	p ^j	pj			palatalized
r	0072	ɹ r	r	4 - intervocalically r - elsewhere	r	
		ɹ̃	rn		rn	nasalized
ri	0072 0069	r ^j	rj	rj	rj	palatalized

² These correspond to the dotted character, i.e., ɲ̥.

TYPICAL IGBO CORRESPONDENCE	UNICODE	IPA	Appen SAMPA			COMMENTS
			Owerri	Onitsha	Ngwa	
s	0073	s	s	s	s	
		š	sn		sn	nasalized
si	0073 0069	s ^j		sj		palatalized
sh	0073 0068	ʃ	S	s	S	
t	0074	t	t	t	t	
		d	d_< - when co-occurring with /l, O, U, a/			
		ʔ	ʔ			
		t ^h	th		th	aspirated
ti	0074 0069	t ^j	tj			palatalized
v	0076	v	v	f (Otu-Onitsha) v (Enu-Onitsha)	v	
		ṽ	vn		vn	nasalized
vi	0076 0069	v ^j	vj	vj	vj	palatalized
w	0077	w	w	w	w	
y	0079	j	j	j	j	
z	007A	z	z	z	z	
		ž	zn		zn	nasalized
zi	007A 0069	z ^j	zj	zj	zj	palatalized

TYPICAL IGBO CORRESPONDENCE	UNICODE	IPA	Appen SAMPA			COMMENTS
			Owerri	Onitsha	Ngwa	
VOWELS						
a	0061	a	a	a	a	
e	0065	e	e	/e/ when co-occurring with /i, u, o/	e	
		ɛ		/E/ when co-occurring with /l, O, U, a/		
i	0069	i	i	i	i	
		ɪ ³	ɪ	ɪ	ɪ	high front vowel with retracted tongue root
o	006F	o	o	o	o	
		ɔ̣ ⁴	ɔ	ɔ	ɔ	(mid-)low back vowel with retracted tongue root
u	0075	u	u	u	u	
		ʊ ⁵	ʊ	ʊ	ʊ	high back vowel with retracted tongue root

OTHER SYMBOLS	
.	syllable break
#	word boundary

Notes on Tone

Although tone marks can be used, it is not a common practice nor are they represented in the normal orthography. For this reason, tone marks are not used in the orthographic

³ This corresponds to the dotted character, i.e., ȳ.

⁴ This corresponds to the dotted character, i.e., ɔ̣.

⁵ This corresponds to the dotted character, i.e., ʊ̣.

transcription or in the lexicon. There is also significant inter-transcriber inconsistency in lexicon tone marking due to contextual variation in the realization of underlying tones, variation that is driven in part by grammatical phenomena.

7.1 List of rare phonemes

The aspirated, nasalized, palatalized, and labialized consonants are rare in the database and these include:

Phonemes			
aspirated	nasalized	palatalized	labialized aspirated
bh	Hn	fj	kwh
dZh	dZn	gj	gwh
dh	gn	kj	
gh	rn	tj	
dZh	tSn	vj	
kh	vn	mj	
ph	zn	Hj	
tSh	sn		

7.2 List of foreign phones

Most loan words are phonologically conditioned into the Igbo (Ekpe, 2011; Nkamigbo, 2009, 2010). However, some of the palatalized consonants, primarily /kj/ and /vj/, occur in in the database to represent loan and foreign words.

8. Other language items

8.1 Table of Digits

Numeral	Igbo
0	adigi efu
1	out
2	abuo
3	ato
4	ano
5	ise
6	isii
7	asaa
8	asato
9	iteghete itoolu

8.2 Other Numbers

Numeral	Igbo
10	iri
100	nari
1,000	puku
10,000	puku iri
100,000	puku nari
1 million	nde
10 million	nde iri
1 billion	ijeri

8.3 Spelling Alphabet

Igbo Alphabet	SAMPA
A	a
B	b l
C	tS l
D	d l
E	e
F	f l
G	G l
H	H l
I ⁶	i l
J	dZ l
K	k l
L	l l
M	m l
N ⁶	n l N l
O ⁶	o O
P	p l
Q	kw l
R	r l
S	s l
T	t l

⁶ Representing pronunciations of alphabets with and without dot (i.e., I, Ñ, Q, U).

Igbo Alphabet	SAMPA
U ⁶	u U
V	v l
W	w l
X	e . k l . s l
Y	j l
Z	z l

9. References

- Adegbija, E. E. 2004. *Multilingualism: A Nigerian Case Study*. Africa World Press, Trenton, NJ.
- Arokoyo, B. E. 2012. A comparative phonology of the Olùkùmi, Igala, Owe and Yoruba languages. In *Proceedings of Towards Proto-Niger Congo: Comparison and Reconstruction International Congress*, Paris, September 2012, pp. 18-21.
- Clara, U. U. 2013. Igbo Language and Identity. *Journal of Literature and Art Studies*. 9 (3). pp. 576-588.
- Clark, M. 1990. The Tonal System of Igbo. *Publications in African Languages and Linguistics*. Foris, Dordrecht.
- Ekpe, M. B. 2011. Nigerian English in the determinant triangle. *Studies in Literature and Language*, 2 (2), pp. 93-99.
- Emenanjo, E. N, Umeh, I. A. O. and Ugoji, J.U 1990. *Ọkaasụsụ Igbo, Igbo Metalanguage, Incorporating Recommendations of the Standardization Committee of SPILC, vol 1*. University Press Limited, Obadam.
- Ikekeonwu, C. I. 1985. Aspects of Igbo dialectology: A comparative phonological study of Onitsha and central Igbo dialects. *Journal of West African Languages*, 15 (2), pp. 95-109.
- _____. 1991. Igbo. *Journal of the International Phonetic Association*, 21, pp. 99-101.
- Mbah, B. M., and Mbah, E. E. 2000. *Topics in Phonetics and Phonology: Contribution from Igbo*. AP Express Publishers, Nsukka.
- Newman, R. M., and Newman, P. 2001. The Hausa lexicographic tradition. *African Journals OnLine*, Lexikos, 11, pp. 263-286.
- Nkamigbo, L. C. 2009. Global perspectives on multilingualism: Concepts, issues, problems and prospects pertaining to Igbo. *International Journal of Research in Arts and Social Sciences*, 1, pp. 133-144.
- _____. 2010. Phonology in teacher education in Nigeria: The Igbo language example. *African Journal of Teacher Education*, 1 (1), pp. 48-63.
- Nwachukwu, P. A. 1983. *Towards an Igbo Literary Standard*. Kegan Paul International, London.
- Nwaozuzu, G.I. 2008. *Dialects of the Igbo language*. University of Nigeria Press Ltd, Nsukka.



- Nwosu, V. 2013. *Okwu Ngwa Dictionary*. Home and Abroad USA Publishing, Raleigh.
- Ohiri-Aniche, C. 2007. Stemming the tide of centrifugal forces in Igbo orthography. *Dialectical Anthropology*, 31 (4), pp. 423-436.
- Ojukwu, O. N. and Ndimele, R.I. 2011. Igbo Orthography and Orthography Principles: A Critique. *Journal of Institute for Nigerian Languages*, 2 (2). Online version: <http://inlan.edu.ng/igothographojundi.htm>
- Onukawa, M. C. 2011. Writing in the Igbo language: standards and trends. *Journal of Institute for Nigerian Languages*, 2 (2). Online version: <http://www.inlan.edu.ng/wrinigboanu.html>
- _____. August 7th, 2014. Writing in Igbo Language. <http://nollyculture.blogspot.com.au/2014/08/writing-in-igbo-language-standards-and.html>, retrieved December, 2014.
- Oraka, L. N. 1983. *The Foundations of Igbo Studies*. Onitsha: University Publishing Co, Onitasha.
- Ugorji, C. U. C. 2005. Dialect communities and language vitality: the case of Igbo. *Journal of West African Languages*, 17 (1).
- _____. 2009. A sociophonological model for orthographies in the new world democracy. *Modern Journal of Applied Linguistics*, 1 (6), pp. 360-381.
- Williamson, K. 1972. *Dictionary of Ònìchà Igbo*. Ethiope Press, Bernin.