

A MULTIMODAL VISUAL ANALYSIS

Temilade Afolasade Olatunji and Festus Moses Onipede

Federal Polytechnic, Ilaro, Department of General Studies

ABSTRACT: *This paper is concerned with the multimodal visual analysis of selected second language (l2) English textbooks written by Nigerian authors for the teaching of English phonemes in secondary schools. The paper compared the multimodal visual resources used by the various authors for the teaching of English phonemes, to know if they make use of similar semiotic mode; how the audio-visual mode was used effectively, and the implication of English textbooks without audio-visual mode. The data for this paper were extracted from the selected textbooks and analysed using Halliday's (1976) three metafunctions of language, which was developed in Kress and Van Leeuwen's (1996; 2006) Grammar of Visual Design (GVD). Findings showed that the authors of the selected textbooks used different multimodal resources such as images, verbal texts and instructions to discuss the selected phonemes. We found out that while NCE and MBE used images to support the verbal texts in order for learners to understand the effective ways of learning English phonemes, NOSEC and EFEL did not make use of images, they only gave verbal texts and created rooms for teachers to guide the learners. In sum, none of the selected textbooks used audio-visual resources in form of CD-ROMs to aid learners' understanding.*

KEYWORDS: multimodality, phonemes, textbooks, audio-visual, metafunction.

INTRODUCTION

The major purpose of communication is for meaning making. If a sender has not received feedback, he/she has not disseminated information. It is obvious that textbook serves as a teaching tool widely used to pass across information between the authors, teachers and students. The teaching of English phonemes in a second language environment poses some challenges to both textbook designers, teachers and students. It is expected that the teaching of English phonemes should be taught in such a way that students learn these phonemes easily just as the first language students acquired them. If first language students acquired the phonemes of English easily, it is therefore left for the textbook authors to create learning environment that will allow the l2 students acquire the correct pronunciation. One of the best methods of teaching English phonemes in l2 environment is through multimodality. Multimodality deals with the interpretation of semiotic resources such as images, signs, symbols, etc to investigate how multimodality texts make meaning.

The mastery of English phonemes can only be taught effectively if visual, audio and textual aspects of the lessons are adequately discussed. A learner can interpret multimodal resources and thereby create meaning in different ways, by forming an overall understanding of the subject. Multimodal features of English textbooks are very important where all different resources of textbook (writing,

images, symbols, speech and moving images in digital textbooks) are put in place for meaning potential (kress,2011). Phonemes can be taught effectively if the readers are able to make use of contextual clues to create meaning with the help of different modes, (kress, 2010; Selander & kress, 2010).

The teaching of English phonemes in l2 environment can be effectively taught through teacher's instructions, images, and audio- visual elements. If these elements feature in l2 English textbooks, the students/learners will easily understand the topic, but if otherwise, the topic would remain difficult to understand. Authors or teachers of English textbooks play an important role in instructing learners to make sense of and explore the visual and verbal resources in the texts (image-text relation or the 'co-articulation of image-verbiage') (unsworth , 2006: 1165).

With the increasing use of multimedia technology in education and the growing importance of multimodal l2 (second language) teaching and learning, it becomes necessary to explore the multimodality of the present Nigerian l2 English textbooks as a way to helping students learn English phonemes in an easiest way.

This study is guided by the following research questions: Do the textbook authors/designers make use of similar semiotic mode? Do the authors/designers effectively use audio-visual mode? What are the implications for English textbook without audio-visual mode?

LITERATURE REVIEW

Textbook and Semiotic Modes

Textbooks make use of semiotic modes to communicate meaning. Textbooks are written materials which comprise instruction materials for the teaching of a particular topic. Textbook features icons, drawings, graphics, photos, colours, video recording that can be accessed on the web. (Alvarez, 2016).

A mode is a set of socially and culturally shaped resources for making meaning. According to kress (2009), mode can be best represented with what and how semiotic entries are connected. A textbook producer together with teachers may use audio-visual as part of his communication modes which will make meaning to the readers. Second language English textbooks make use of semiotic mode to communicate meaning, which can be explored through all modes of communications- image, writing, typography and layout.

It is the duty of textbook designers /authors together with teachers to combine different signs to aid the textual component of a given topic, most especially the phonemes of English. It is expected that a textbook discussing English phonemes, should be made to include both written and audio-visual resources for effective teaching and learning.

Again, textbook authors and teachers should ask themselves the following questions before they teach English phonemes. One, what is my preferred social relation with my imagined audience and how can I best realise it? Two, how is the subject content best shaped and realised to represent my theoretical conception of the subject while maximizing the learners' engagement? If

appropriate answers are given to the above questions, the authors/teachers will know the best way to teach the topic.

Relationship between Text and Image

Words alone cannot effectively communicate meaning in the teaching of English phonemes. Hence, ‘the writing of authors sits alongside the images provided by visual artists and the layout of the Graphic Designers’ (Bezemer & Kress, 2010:10). It shows that words together with images help learners to grasp the content of the lesson.

How images are arranged in textbooks determine the way learners will quickly learn the topic. Based on Marsh and White's (2003) work, they identify 49 functions from the relationship between text and image. The functions of these relationships can be merged into one. In view of Hodnett (1988), the function of image should be categorised into three –top decorate, to inform and to interpret. Textbook authors and teachers should merge these three functions above in the teaching and learning of English phonemes.

Moreover, meaning potential can be created in the language of images via colour, tone, angle, framing and among other categories. The image represents an aspect of the world (abstract or concrete), plays a part in some interaction with the viewers and combine visual elements into coherent whole (Salbego et al, 2015).

Students or learners can also understand English phonemes through the interpretation of images in the light of the three meta functions which can help them interpret the verbal text in L2 textbooks (Royce, 2007). If phonemes are taught with verbal language, there must be room for appropriate images that will accompany them. In view of Kress (2001:67), ‘verbal language is being displaced as a communication mode by images, in many sites of public communication: whether in school textbooks, in newspapers, in reports produced in institutions of all kinds, in the electronic media, and in information and communication technologies in general. Images are ceased to be there as more illustration, that is, an embellishment of the central, the linguistics nor socio linguistics in any longer sufficient as the theoretical enterprise to account fully and plausibly for central aspects of representation and communication.’

Past and Present Studies

There is lots of researches on representation in learning resources from different perspectives, e.g, Nanyang Salbego, 2015; Udenwa VC and Nriokanta, 2008; Szpyra-Kozłowska, 2015; Bakliti Niana and Rev. Esther, 2012; Ajayi, 2012; Giaschi, 2000; Tan, 2012; Li, 2011; Bezemer and Kress, 2010; Liu and Diarming, Qu, 2014; Nelson, 2006 and Jewitt, 2008.

Giaschi (2000) pointed out that sometimes textbook producers relayed message through the image across textbooks and such might not be carried by the linguistic text. In Nelson's (2006) work, he argued that multimodal practices allow English learners the opportunities to negotiate and communicate meanings in different media than English language. Udenwa and Nriokota (2008) evaluate English language textbooks used in the junior secondary schools in Nigeria, and they found out that the textbooks are difficult for the intended reader and imbalance in the integration

of English language and Literature-in-English. Jewitt (2008) suggested that teachers may use multimodal texts as “the basis for critical engagement, redesign, or the explicit teaching of how modes construct meaning in specific genres.” Jewitt’s view is that teachers can help learners to interpret text and the learners will relate this text to the given images. Again, English teachers need to link textual features of textbooks to social contexts and politics that inform them. For example, teachers need to teach students how textbooks’ structures and visual resources interact and integrate to convey biases and prejudices and how such features can be interpreted within particular socio-cultural contexts and through specific social practices. In Li’s (2011) work, ‘How teachers deploy multimodal textbooks to enhance English language learning’, he compared the listening and speaking course book and the reading and writing course book of *New Horizon College English* and argued that the different tasks and cognitive characteristics of different classes determined mode selection. In addition, Tan (2012) examined “Advanced with English for Senior high school students”, and he claimed that the lovely layout and image facilitated literacy learning and stimulated students interest in participation in learning practice, and the text and image co-existed to make meaning, while the image repeated the text content more intuitively. Liu and Dianning (2014) explored the multimodality of EFL textbooks for Chinese college students, and they found out that EFL textbook shows the importance of properly arranging the different modes in a double page spread. The textbooks compilers also consider the learners’ English proficiency and appropriately employ the variety and number of multimodal resources to achieve optimal intersemiotic complementarity. In order to effectively teach English pronunciation in EFL classroom, Szpyra-Kozłowska (2015) stated that pronunciation can only be effectively taught and at the same time attractive-learner-friendly and teacher-friendly, by developing a holistic multimodal approach to EFL phonodidactics which cater for students’ different learning styles and combination of elements of auditory, articulatory, cognitive and multisensory training. Multisensory training will help students with different learning modalities –auditory, visual, tactile and kinesthetic (Celce-murcia et al, 1996).

To sum up, these previous studies are helpful to this present study, but none of them has been able to examine the multimodal visual analysis of selected Nigerian L2 English language textbooks. There are presently different 12 English textbooks written by Nigerians, and it is necessary for us to carry out a comparative study of these textbooks to know how each of this textbook makes use of multimodal resources to explain English phonemes. The present study therefore aims to fill the gap that has not been filled by the previous studies.

Theoretical Framework

According to Halliday (1976), language is used to perform three functions- ideational, interpersonal and textual. Ideational meaning concerns human experience of the world. This function relates to how words are used to express actions, objects, places, events, people, things, qualities, and ideas. The Interpersonal meaning deals with interaction between a speaker and listener. This speaker-listener interaction is what Van Lier (2004) referred to as “roles” and identities of participants, socio cultural issues, relationships, power and control” The textual meaning concerns the relation that holds among the texts in order to create a coherent messages that is relevant to the context of the situation. This coherent messages can be realized through

given and new information. Halliday's three metafunctions of language were later developed by Kress and Van Leeuwen in 1996, 2006, to the analysis of multimodal texts. Through visual grammar, language performs three functions- representational, interactive and compositional. Representational social semiotics concerns how different people in visual image were represented and relate to each other; interactive social semiotics considers the kind of participant-viewer social relation; and the compositional social semiotics concerns different compositional configurations that realise different meanings.

The theory of social semiotics has been extended by some scholars to explain how different semiotic modes serve as resources for critical interpretations of texts and build on individuals' multifarious subjectivities (Ajayi, 2012, cited in Cope and Kalantzis, 2009, Kern, 2006, and Lemke, 2015). It is therefore expected of both textbook authors and teachers to teach students how textbooks structures and audio-visual resources interact to derive meaning from the learning of English phonemes.

METHOD

For the present study, images from four textbooks were selected and analysed in the context of multimodal visual social semiotic (Kress and Van Leeuwen, 1996; 2006), based on our experience and familiarity with them. The selected books- *New Oxford Secondary English Course* (NOSEC) for junior secondary school one (Banjo et al, 2014), *New Concept English* (NCE) for senior secondary school one (Ademola-Adeoye et al, 2014), *Exam Focus English Language* (EFEL) for senior secondary schools (Banjo et al, 2018) and *Macmillan Brilliant English* (MBE) for junior secondary school one (Osisanwo et al, 2012). The selected textbooks are well known English textbook series in Nigeria and used in a number of junior and senior secondary schools. These textbooks present students with specified proficiency level, and concerned activities and texts for beginners since this is the phase in which learners may need more support and images might help to reinforce their understanding of the content.

As stated in the theoretical framework, the three metafunctions developed by Kress and Van Leeuwen (2006) in their *Grammar of Visual Design* (GVD) have grounded our qualitative analysis in this paper and our study tries to show benefits from interpreting images coupled with texts or language activities. Along with the image analysis, examples from the textual activities are provided. The analysed images were named as images A1, A2, B1, B2, C1 and C2. In junior English textbooks, we examined phonemes /a:/, /p/ and /b/. For the senior category, phonemes /ɜ:/, /t/ and /d/ were examined. We refer to textbooks makers as authors/teachers, and the viewers as students/learners of English phonemes observing the image analysed.

Data Presentation

The Vowel /ɑ:/

A. /ɑ:/ is a long vowel. To pronounce it, the front of the tongue remains low on the floor of the mouth; the back is pushed backward. The mouth is fully open. The lips are not quite spread; they form an oval shape.

B. The vowel /ɑ:/ occurs in the following sentences as shown. Listen carefully as your teacher reads.

- I saw Martin at the park.
- If you work hard, you will pass.
- Danladi's aunt is a pharmacist.
- Remain calm; don't panic.
- A new car for one million naira? That's a good bargain.
- He laughs best; he laughs best.

C. Listen again as your teacher reads. Repeat after him.

D. Notice that in writing, the vowel /ɑ:/ may take the letter a as in: pass, father; or ar as in: car, park, bargain; or al as in: calm, palm, psalm. In a few cases, it uses au as in: laugh, aunt (auntie); or o as in: clerk, sergeant.

E. Here are more sentences in which the vowel occurs. Read them aloud.

- Everyone knows that the car runs faster than the cart.
- The captain ran fast past the goalkeeper.
- Our branch chairman is a police sergeant.
- Armed robbers often operate in darkness.
- Mark scored a higher mark than John.
- The night guard raised an alarm.

Image A1: NOSEC Book 1 (junior)

Consonants /p/ and /b/

A. You remember what a consonant is, do you? It is a sound produced when the air from the lungs is totally or partially obstructed somewhere in the mouth. In some consonants, when activity is going on in the voice box, we hear the sound loud. The sound produced is described as **voiced**. In some others, no activity goes on in the voice box. The air appears to start moving in the mouth only. In this case the sound is heard faintly, like a whisper. It is called a **voiceless consonant**.

B. Consonant /p/

The air for this sound is in the mouth. There is no activity in the voice box. The lips are pressed hard together. The air in the mouth pops out suddenly with force. Place the back of your hand in front of your mouth and force the air out as you say /p/. The air hits the back of your hand with a very faint sound. It is as if you throw the air out. /p/ is a **voiceless sound**.

C. Say the following words: pat, pass, pit, help, map, wrap, tap, perhaps, replace, hope, pen.

D. Read the following sentences after your teacher:

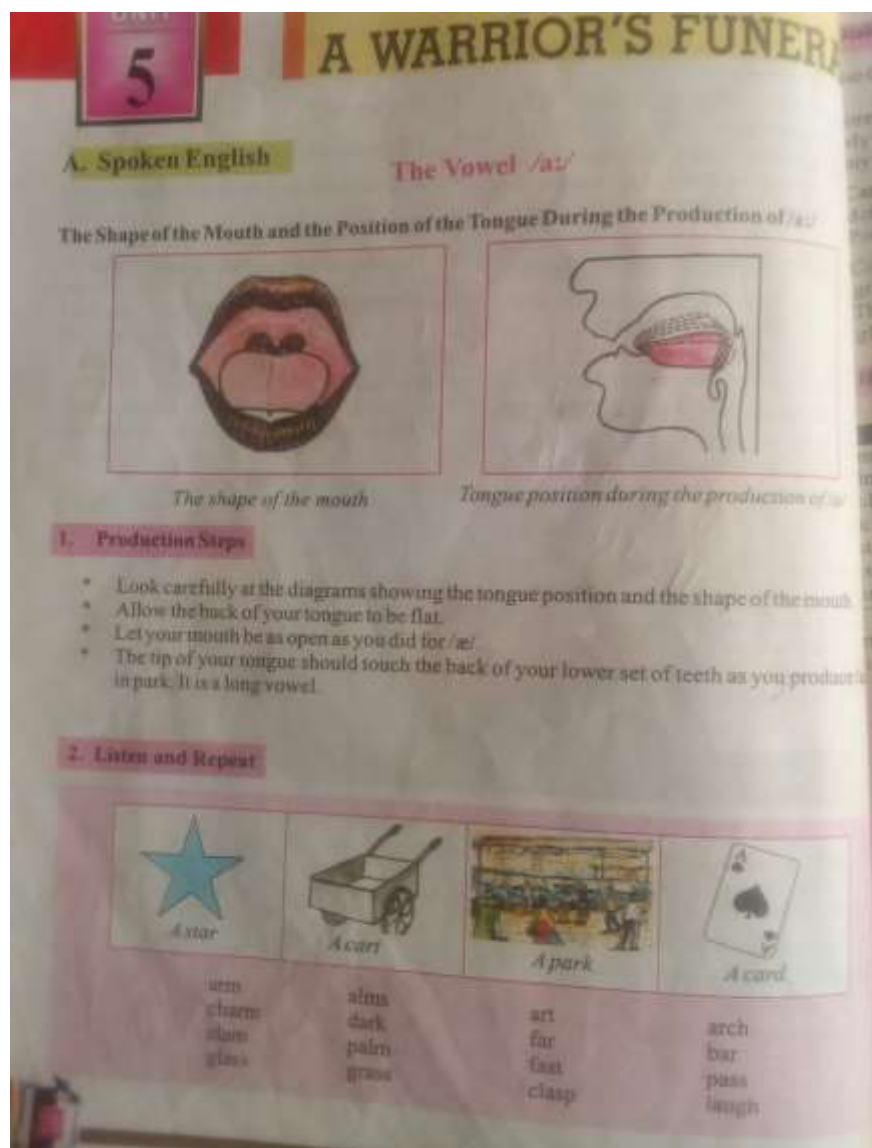
- Please lend me your pen.
- We gave a present to the school principal.
- If you press hard, the seeds will pop out.
- Our spaniel has three pups.
- The pupils are proud of their school.

E. Consonant /b/

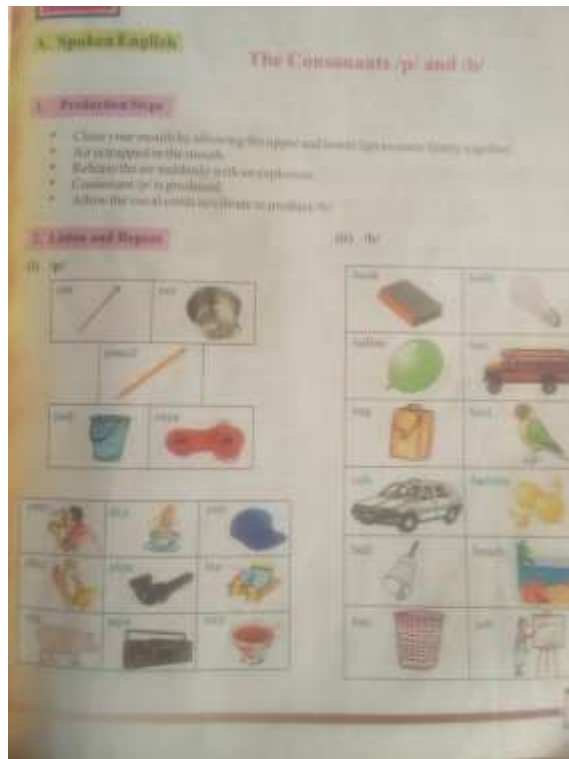
For this sound, the air from the lungs passes through the voice box. The elements in the voice box are active, flapping against each other. The lips are pressed together. The air is pushed out suddenly but gently. Because there is activity in the voice box, you hear the sound clearly as you say /b/. It is a **voiced consonant**.

F. Say the following words: bat, bus, bass, bit, robe, rub, bathe, abode, bribe, probable, bulls.

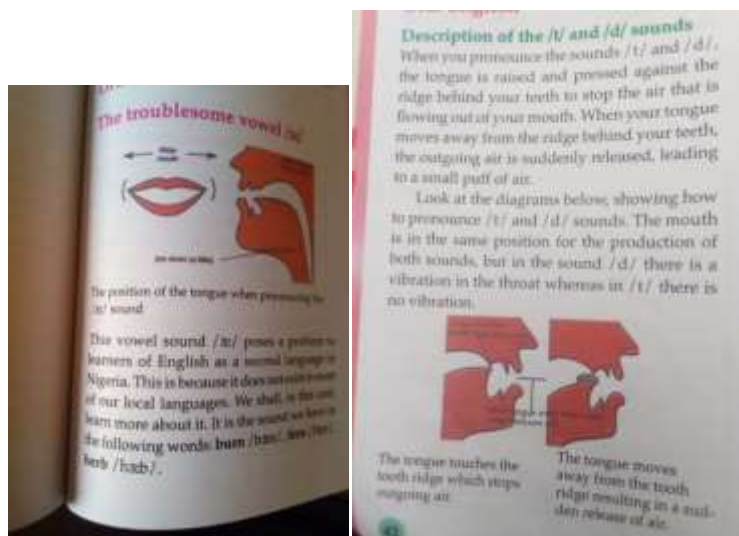
Image A2: NOSEC Book 1(junior)



Images B1: MBE 1 (junior)



Images B2: MBE Book 1 (senior)



Images C1 & C2: NCE Book 1-3 (senior)

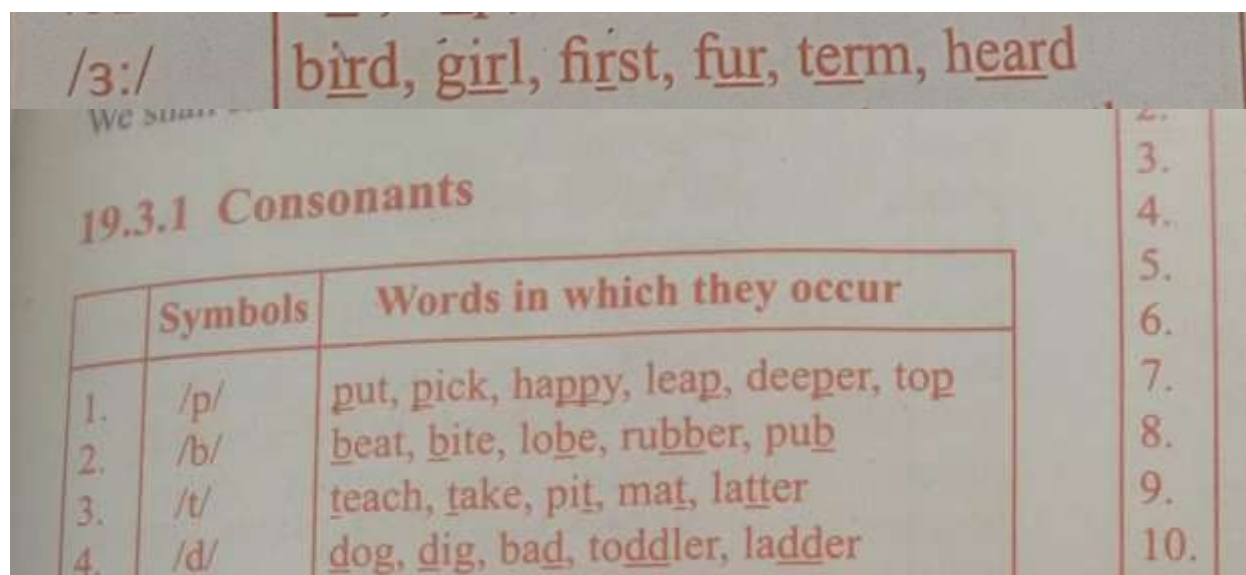


Image D1 & D2: EFEL Book 1-3 (senior)

Table 1: Multimodal resources in each textbook

Textbook's Name	Multimodal Element(s)	Topic
NOSEC Book 1	verbal texts, instructions	phonemes /a: , p, b/
MBE Book 1	verbal texts, images, instructions	phonemes /a; , p, b/
NCE Book 1 (senior)	verbal texts, images, instructions	phonemes /ɜ: , t, d/
EFEL Book 1-3 (senior)	verbal texts	phonemes /ɜ: , t, d/

ANALYSIS AND DISCUSSION

For the purpose of learning, textbook authors must have ideas to share with both teachers and learners (ideational/representational), the message they shared between them and their students, that is, the interaction between the textbook authors and their learners (interpersonal/interactive), and the linkage or coherence in the content of their discussion (textual/compositional), must be appropriately structured to provide students/learners with an effective learning environment. Table 1 summarised the multimodal resources found in each textbook.

From image A1, the textbook's authors make use of verbal element, and a teacher acts as an instructor to students learning of phoneme /a:/. The description of /a:/ starts with the position of the tongue- low, spread, and exercises were given to produce the /a:/ speech sound of n the given words. The description of /a:/ does not make use of image. Image A2 comprises verbal element and there is room for teachers to guide learners to master the correct speech sounds. Phonemes /p/ and /b/ are described based on the state of the glottis- voiced or voiceless sounds. Learners are given instructions to follow based on their teacher's order. The textbook gave class exercises for learners to practice. The meaning potential is shared between the authors and learners the moment they follow the instructions.

Images B1 and B2 present how Phonemes /p/ and /b/ are described in MBE book 1. The multimodal elements used in the description of the Phonemes are images and verbal elements which serve as instruction to learners. There is a photo showing the shape of the mouth (spread) and the position of the tongue during the production of /a:/. The learners are instructed to allow the back of their tongue to be flat; their mouth should be open, and the tongue should touch the back of lower set of the teeth. The given exercises highlighted some words with pictures to aid students' pronunciation. Likewise, consonants /p/ and /b/ are described with production steps, which instructed the students to close their mouth by allowing the upper and lower lip to come firmly together; the air should be trapped in the mouth; the air should be released with explosion and there should be vibration of vocal cords for the production of /b/. The textbook authors further gave a list of English words with pictures for easy pronunciation. Here, the teacher produced each word and the learners imitated him. Among these words are pin, pot, pail, rope, book, bulb, balloon, bus etc. For /p/ and /b/ respectively.

Images C1 and C2 present Phonemes /3:/, /t/ and /d/ in NCE for senior secondary school book 1. The textbook starts the description of phonemes /3:/, /t/ and /d/ with two pictures showing the position of the tongue. According to NCE textbook's authors, vowel /3:/ poses a problem to the l2 learners in Nigeria. This problem arises because the three major languages in Nigeria do not have the /3:/ vowel sound. The challenge is evident in the following words: burn /b3:n/, fern /f3:n/, herb /h3:b/. Again, students are also given the production steps for consonants /t/ and /d/. For the production of /t/ and /d/, the tongue is raised and pressed against the ridge behind the teeth to stop the air that is flowing out of the mouth the moment the tongue moves away from the ridge behind the teeth, the outgoing air is suddenly released, leading to explosion. The verbal description is accompanied with two pictures- one showing when the tongue moves away from the tooth ridge which results in a sudden release of air. The visual components of images B1, B2, C1 and C2 disseminate visual meanings which are realised via images, colours, framing and saliency. The compositional/textual components of these images are appropriate because the moment a learner sees the photos, he/she will interpret the pictures along with the verbal texts. For examples, the pictures of cart and car will aid learner's understanding of vowel /a:/

The last images, that is, D1 and D2 of FEEL, itemised only phonemes /3:/, /t/ and /d/ with words in which they occur. The given examples feature /3:/ in bird, girl, first, fur, term, heard; /t/ teach, take, pit, mat, and latter, and phoneme /d/ dog, dig, bad, toddler and ladder. The verb texts in images D1 and D2 can only make learning easy if the teacher guiding the learners is competent to do so. In a situation where a learner finds himself/herself in a learning process without the presence of a teacher, the learning of these phonemes will be impossible.

Showing the relationships between verbal texts and images, MBE and NCE discussed phonemes /a:/, p, b, 3:, t, d/ with appropriate images that allowed readers to understand the correct pronunciation. The colours used in designing the speech organs and the words in which these Phonemes are pronounced, assisted the readers to involve in the learning process. The way images accompanied the given words corroborates Yassine (2014) claimed that "the inclusion of more images does not only transform the appearance of the teaching content but also contributes in reshaping the relations between teaching and learning by providing more resources for meaning

making." Again, the colours of the images- red, yellow, blue, green, black, white and orange, lured learners to engage themselves in learning.

Furthermore, NOSEC and EFEL did not include images in the teaching of the same phonemes presented by MBE and NCE. This showed that students had challenges comprehending the pronunciation tips. We all know that readers will quickly understand a topic accompanied with visual resources easily without any challenge. If the teaching of English Phonemes features Multimodal pedagogy, learners would have the opportunity in the second language environment to expand their literacy practices, which will enable them to integrate different representational models- languages, images, colours and non-textual features to mediate interpretation of texts (Ajayi, 2011, p. 62).

FINDINGS AND CONCLUSION

This paper has examined the Multimodal visual analysis of the teaching of English Phonemes in 12 English textbooks written by Nigerians. The findings from the selected textbooks revealed that the authors of the textbooks used different methods for teaching English Phonemes. While MBE and NCE used verbal texts and images in the teaching of English Phonemes, NOSEC and EFEL made use of verbal texts; NOSEC only accompanied the verbal texts with production instructions. Here, the teacher is left to guide the learners to follow the instructions written in the textbooks. Finally, none of the textbooks used audio resources. Since textbooks are written to guide both teachers and students, there must be improvement on the part of the textbook authors to include audio resources in the teaching of English phonemes. The audio resources can be in form of CD-ROMs. Once audio-visual resources are included in the teaching of English Phonemes, learners/students will find it easy to learn the correct pronunciation.

Recommendations

Textbooks written in 12 environment must create rooms for learners to learn the English Phonemes without facing any challenge. It should also be stated that not all teachers of English language are able to interpret the Multimodal aspect of the textbooks, but author provision of audio-visual CD-ROMs will best assist the teachers teaching English Phonemes. It is therefore, according to O'Halloran & Lim (2011) the "responsibility of educators to understand the multimodal ways in which knowledge is presented and, beyond that, to teach students to access, appraise and appreciate the Multimodal texts which they will inevitably encounter."

References

- Ajayi, L. (2012). How teachers deploy multimodal textbooks to influence English language learners. *TESOL Journal*, Vol. 6:16-35.
- Ajayi, L. (2011). Video "reading and multimodality: A study of ESL/literacy pupils' interpretation of "Cinderella" from their socio-historical perspective. *Urban Review: Issues and Ideas in Public Education*, 44(1), 60-69.
- Alvarez, J. (2016). Meaning making and communication in the Multimodal age: ideas for language teachers. *Colomb. Applied Linguistics*. J., 18(1), 98-115.
- Ademola-Adeoye, F., Adam, Q., Eto, J., Eyisi, J., Adekunle, A., & Adepoju, B. (2014). *New concept English for junior and senior secondary schools*. Lagos: Lead Africa PLC.

- Banjo, A., Adeniran, A., Akano, A., & Onaga, U. (2014). *New Oxford secondary English course for junior secondary schools*. Ibadan: University Press PLC.
- Banjo, A., Ayodele, S., Ndahi, K., & Akano, A. (2018). *Exam focus English language for senior secondary schools*. Ibadan: University Press PLC.
- Bakhtiniana and Tev Esther. (2012). Multimodal genres in textbooks: are students being schooled for visual literacy? *Discurso*. vol. 7(2). Sao Paulo (Ed.)
- Bezemer, J & Kress, G. (2010). Visualizing English: A social semiotic history of a school subject.
- Celce-Murcia, M., Brinton Dim., & Goodwin, J. (1996). Teaching pronunciation: A reference for teachers of English to speakers of other languages. Cambridge: CUP.
- Giaschi, P. (2000). Gender positioning in education: a critical image analysis of ESL textbooks. *TESL. Canada Journal* 18(1), 32-46.
- Halliday, M. A. K. (1976). System and function in language. London: Oxford University Press.
- Jewitt, C. (2008b). The visual in learning and creativity: A review of literature . England: Creative Partnerships.
- Kern, R. (2006). Perspectives on technology in learning and teaching languages. *TESOL Quarterly*, 40(1), 183-210.
- Kress, G. (2009). Multimodality. A social semiotic approach to contemporary communication. London: Routledge.
- Kress, G & Van Leeuwen. (2006). Reading images: *The grammar of visual design*. 2nd ed. London: Routledge.
- Kress, G & Van Leeuwen, T. (2001). *Multimodal Discourse: The modes and media of contemporary communication*. London: Arnold.
- Kress, G & Van Leeuwen. (1996). *Reading images: The grammar of visual design*. London: Routledge.
- Li, S. (2011). Multimodal discourse analysis of New Horizon College English coursebook. MA Thesis, Southwest University, China.
- Marsh, E. E & White, M. D. (2003). A taxonomy of relationships between images and text. *Journal of Documentation*, 59(6), 647-672.
- Nelson, M. (2006). Mode meaning and synaesthesia in multimedia l2 writing. *Language Learning and Technology*, 10(2), 56-76.
- O'Halloran, K.L., & Kim Fei, V. (2011). Dimension of multimodal literacy . *Videnomlasning*. Number 10, September 2011, 14-21. National Videocenter for Learning: Denmark.
- Osisanwo, W., Alabi, V., Owoeye, B., & Onuoha, A. (2012). *MacMillan Brilliant English for junior secondary schools*.
- Royce, T. D. (2007). Multimodal communication competence in second language contexts. In T. D. Royce; W. L. Bowcher(eds.), *New Directions in the Analysis of Multimodal Discourse*, Mahesh, Lawrence Erlbaum, pp. 361-403.
- Salbego, N., Heberie, V. M., & Gabriela, M. (2015). A visual analysis of English textbooks: Multimodal scaffolded learning. *Calidos Caprio*. vol. 13(1). pp. 5-13.
- Szpyra-Kozłowska, J. (2015). English pronunciation pedagogy in EFL instruction: A research

- based approach. Bristol, Buffalo, Toronto: Multilingual matters.
- Tan, Y. (2012). A multimodal discourse of the senior high school English textbooks: Advance with English. MA Thesis, Hunom University of science and Technology, China.
- Udenwa, V. C & Ikonta, NIR. (2008). Evaluation of English language textbooks used in the junior secondary schools. In International Journal of Educational Research 4(1), pp. 167-171.
- Unsworth, L. (2001). Teaching multiliteracies across the curriculum: Changing contexts of text and image in classroom practice. Buckingham: Open University.
- Van Lier, L. (2004). *The ecology and semiotics of language learning perspective: Sociocultural perspective*. Boston: Kluwr Academy Publishers.
- Unsworth, L. (2006). Image/text relations and intersemiosis: towards Multimodal text description for multiliteracies education. In International Systemic Functional Congress, 33, Sao Paulo, 2006 Proceedings. Sao Paulo, PUC-SP, 1: 1165-1205.
- Xiqin Liu & Dianning QU.(2014). Exploring the multimodality of RFL textbooks for Chinese college students: A Comparative study. In RELC Journal.
- Yassine, S. (2014). Multimodal design of EFL textbooks: A social semiotic multimodal approach. In Anglisticum Journal (IJLLIS) 3(12).