



An Analysis of “English for Nusantara” Textbook Through Multimodal Approach

¹Rahma Wati Arani, ²Kurnia Budiyaniti

¹UIN Sultan Syarif Kasim Riau, ²UIN Sultan Syarif Kasim Riau
[¹rahmawatiarani78@gmail.com](mailto:rahmawatiarani78@gmail.com), [²kurnia.budiyanitiuin@gmail.com](mailto:kurnia.budiyanitiuin@gmail.com)

Corresponding Author

Abstract

The purpose of this research is to analyze how visual images and verbal texts combine to expose visual meaning in the “English for Nusantara” textbook for Grade 7. The researcher used a qualitative research design with content analysis, focusing on Chapters 1 through 5. This study employed Kress and van Leeuwen’s multimodal discourse theory—representational, interactive, and compositional modes. Findings reveal that the combination of images and texts in the textbook effectively scaffolds student comprehension. Images help clarify the task, enhance engagement, and support language acquisition, making the textbook a valuable resource for EFL students. This study highlights the textbook’s alignment with the Merdeka Curriculum and its use of multimodal resources to promote literacy.

Keywords:

English Textbook; Visual Meaning; English for Nusantara

This is an open access article under CC by 4.0 license



INTRODUCTION

Textbooks are used as reference materials for standards-based learning in a specific field of study, written by subject matter experts. For example, in learning context, textbook is used as the main materials in learning process. Textbook contains verbal text and visual images which use more than one semiotic modes at the same time, which is known as multimodal. In fact, “it is difficult these days to find a single text which uses solely verbal English” (Goodman, 1996). Then, the use of images helps learners have a thorough understanding of topics and contents of the textbook. Even so, students still require the guidance of their teacher in order to read these multimodal texts and apply the linguistic features and competencies that are emphasized in textbooks. Since the textbook offers numerous visual aids for the materials, it is depictions and features a large cast of characters.

By believing that, to teach English Foreign Language (EFL) learners improving their English skills in the class, teachers can foster learners’ language learning processes by directing them to read images, instead of



only written texts in textbook. It contributes to create imagination in learners' mind because visual elements are considered as any visual material such as drawings, and photos included in a textbook which facilitates learning (Tomlinson, 2003) and the meaning of images (and of other semiotic codes) is always related to, and in a sense, depended on verbal text (Barthes in Kress & van Leeuwen). Therefore, there is a relation between the text and the visual or image.

English textbooks can be used as the main material to facilitate language learning in classroom and act as guidance for learning process. English for Nusantara class 7 (in revised edition 2022) is a new English textbook prepared by the ministry of education and culture of Indonesia. This textbook's material or content was created with the curriculum 2013 syllabus in mind. The materials are primarily intended for senior high schools in their initial years of study and are only available in English. This textbook covers social skills and fundamental competencies. Listening, reading, speaking, and writing are among the language skills that are supported in each chapter and by the goals. In addition, this textbook is integrated with multimodal resources (e.g., image, color, layout, typography, and font) to communicate messages in short, this textbook uses images to convey certain ideas or messages, so researcher interested in learning more about the multimodal nature of visual text. For this reason, educators must comprehend how pictures are woven into the textual material of this book. This point is important because English language learners who do not have the skills to interpret images in textbooks may not understand how social, political, and ideological messages are conveyed to them and how they are positioned as both english language learners and as social subjects (Giaschi, 2000).

The book's focus on using English language variations relevant to nusantara's communication needs adds value, promising practicality in language mastery. Research on the extent to which "English for Nusantara" aligns with local curriculum standards and its influence on literacy development will contribute to assessing its overall contribution to the educational landscape in the Nusantara region. Multimodal is a term used to refer to the way people communicate using different modes simultaneously (Kress & van Leeuwen, 1996), defined as the use of multiple semiotic modes in the design of products or simultaneous semiotic events, and the specific ways these modes are combined to strengthen, complement, or arrange in a particular order (Kress and van Leeuwen, 2001).

Multimodal can also be considered a technical term aimed at indicating that, in the process of meaning-making, humans utilize



various verbal and visual semiotics (Iedema, 2003). Meanwhile, Chen (2010) defines multimodal as all verbal and visual semiotic sources that can be used to realize texts and levels of dialogic engagement in a textbook. Strategies or approaches used to gather, examine, and evaluate data consisting of many modes of communication, such as text, images, sound, and movement, are referred to as multimodal research methods.

English for Nusantara was chosen as the study's focus for a number of important reasons. First of all, this textbook was created especially to aid in the application of the Merdeka Curriculum, which gives educational institutions the freedom to create lesson plans based on the requirements and traits of their students. The effectiveness of the textbooks used in this curriculum must be assessed as part of the government's efforts to raise educational standards.

In order to improve students' comprehension and involvement in learning English, English for Nusantara incorporates multimodal components, including spoken language and visual images. Examining how these components interact can reveal information about how meaning is created and communicated to students as well as how successful this multimodal approach is in the Indonesian educational setting. There is currently no precise information on the adoption of English for Nusantara in the sources that are currently available, including the number of schools that use this textbook. But according to the data that is currently available, about 140,000 schools have signed up to use the Merdeka Curriculum. It follows that this textbook may be used extensively in educational institutions that have embraced the curriculum.

The importance of multimodal into textbook because the goal of the , multimodal approach is to help students become readers and producers of multimodal approach by highlighting the role that different sources play in giving texts meaning and the ways that particular decisions contribute to the desired communicative outcomes. in a textbook which facilitates learning (Tomlinson, 2003) and the meaning of images (and of other semiotic codes) is always related to, and in a sense, depended on verbal text (Barthes in Kress & van Leeuwen).

The text and the visual, or image, are related. It is believed that teachers can support students' language learning processes by having them read visual materials rather than just written texts in order to help English foreign language (EFL) learners improve their English proficiency in the classroom inside a textbook. It contributes to create imagination in learners' mind because visual elements are considered as any visual material such as drawings, and photos include.



METHOD

In this study is a content analysis, that is method is qualitative. Qualitative research is an approach used to understand and explore human experiences in social and cultural contexts. According to Creswell (2014) this research focuses more on gaining an in-depth understanding of issues rather than measurement or statistics. According to Merriam (2009) Qualitative research is an approach used to understand the meaning and experiences of individuals in specific situations. This research relies on interviews, observations, and documents as the main sources of data. According to Sugiono (2017) Qualitative research is a type of research that seeks to understand the phenomena occurring with research subjects holistically and in depth. This research aims to explore the meaning of an event, occurrence, or social phenomenon in the context of real life, without emphasizing numerical measurements. The data used in qualitative research is descriptive and collected through various techniques such as interviews, observations, and document studies.

Content analysis research is essential for learning important things about different facets of communication and information sharing. With the use of this technique, researchers can spot patterns and trends in the materials they have examined, providing a thorough understanding of how specific subjects change over time. A quantitative evaluation of the predominance of particular messages or viewpoints is made possible by content analysis, which quantifies the frequency and distribution of particular words or themes. Furthermore, it is an effective instrument for identifying prejudices or specific points of view in the text, encouraging critical assessments of how information is presented.

In this study, researcher chose text that followed by picture taken from an English textbook for Nusantara Class VII (in revised edition 2022) prepared by the Ministry of Education and Culture of Indonesia for junior High School as the object of the study to be analyzed. Text-pictures from Chapter 1 until 5 were taken as the main source of the data and researcher took text-pictures to be analyzed.

The documentation method in qualitative research is used to collect, analyze, and interpret data obtained through various documents or written sources. This method aims to gain an in-depth understanding of the phenomenon being studied, by paying attention to the context, relationships, and meanings contained in the documents. Researcher analyze textbook English for Nusantara seventh grades in junior high school.

The researcher collected data using triangulation in the documentation of five chapters of the "English for Nusantara" textbook. Triangulation is a methodological approach used to enhance the credibility



and validity of research findings by incorporating multiple perspectives, methods, or data sources. When documenting the "English for Nusantara" textbook through multimodal analysis, triangulation involves integrating diverse data sources, theoretical frameworks, and methods to provide a comprehensive understanding of the multimodal elements within the textbook.

The purpose of this study's data analysis is to assess how the multimodal components in the English for Nusantara book aid in students' comprehension of the English language. Kress and van Leeuwen's (2006) Multimodal Discourse Analysis (MDA) approach, which incorporates representational, interactive, and compositional modes, was used to conduct the analysis. Information was gathered from the textbook's Chapters 1 through 5 and examined in light of how text and images work together to enhance students' learning.

This study uses a indicators to assess the English for Nusantara textbook's efficacy, including : how well the content complies with the Merdeka Curriculum, how well students engage with the material through multimodality, and how easily students connect verbal and visual information to understand English language concepts. Feedback from teachers and students about how the textbook aids in their learning process can also be used to evaluate effectiveness in addition to content analysis. As a result, this study looks at both how the textbook uses multimodal elements and how it helps students learn and advance their English language proficiency.

RESULTS AND DISCUSSION

The results include the realisation of visual meanings and visual scaffolding in the English textbook for Nusantara seventh graders in junior high school. Studying images from Chapters 1 through 5, the researcher focusses on multimodal text, which consists of both verbal and visual material. Some of the varied visuals serve a communicative purpose. Chapters 1 through 5 of the book contain 49 photos that need to be examined.

1. Relation of Visual and Verbal Text to Expose Visual Meanings

Based on Halliday's social semiotics, Kress & van Leeuwen (1996; 2006) establish a technique of social semiotic analysis of visual communication and produce a descriptive framework of multimodality, allocating representational, interpersonal and



compositional interpretations of pictures.

2. Results of analysis 49 images :

a. Representational mode

Language structures of all kinds, as well as the visual processes that go along with them, are associated with particular roles played by the participants. Thus, the analysis of the inter semiotic relationship between visual and verbal modes in representational terms requires the identification of the represented participants, whether animate or inanimate, the processes or the activity described, the attributes or the qualities of the participants and, finally, the circumstances in which the action is being developed (Unsworth and Wheeler, 2002; Unsworth, 2008).

b. Interactive mode

As for the interpersonal features of multimodal texts, the relationships between the visual, the producer and the viewer are to be considered, together with the analysis of the inter semiotic mood in both verbal and visual modes (Unsworth and Wheeler, 2002). In the verbal or visual modes, the instructional parts are associated with commands the readers making offers or requiring some kind of action of them. Thus, the verbal and visual modes support each other to make the speech functions clear. Among the visual techniques used to analyze interpersonal meaning is the absence or presence of facial expressions towards the viewer, gestures which make commands, and offers of information or offers of goods and services to the viewer (Royce, 2007)

C. Compositional mode

In addition, the compositional features analysis in the multimodal text is related to the layout of the page in order to determine the extent to which the visual and verbal elements achieve the coherence to the unit. Another the inter semiosis in compositional terms also involves the relation to the main principles of composition, that is, the distribution of the information value, visual salience (size, colour, focus, ...) and visual farming (Kress & van Leeuwen, 1996, 2006; Royce. 2007; Unsworth, 2008). The placement of elements to the left (given information) or to the right (new information) of the double spreads, the relative size of the figures in the image and the use



of framing are all relevant factors of the compositional meaning (Unsworth and Wheeler, 2002, p. 70; Kress & van Leeuwen, 1996, 2006).

The representational mode, which shows various human and non-human participants, actions, and events, lends further credence to the images' informational and functional qualities. They depict accurate depictions of the world that are consistent with the students' everyday experiences. Non-realistic visual arts are currently pervasive in educational materials (Chen, 2010; Martin, 1985), and fictional beings with senses and sensibilities, like talking animals and personified objects, have become more common in books, visual aids, and other educational materials. The widespread use of fantasy in educational settings is said to run the risk of promoting a "childist" ideology (Martin, 1985; p. 55), which undervalues students' capacity to learn the language from texts that deal seriously with a variety of real contexts, even though playful cartoons have the advantage of encouraging students' motivation and interest in learning the foreign language.

Therefore, it appears that the best option for junior high school is a realistic style in the textbook being studied. Commonplace people, things, and locations are depicted in this book in a realistic manner. In defining the context or situation, however, the majority of the represented participants who engage in conversation, play, reading, or other activities without any background or with minimal contextualization could hardly relate to the students. This discussion is further supported by the findings of the interactive mode analysis. Students may find it much easier to use the images to interpret the textual portion of the activity if the shot types (long, medium, or close) suggest some social distance. Additionally, the use of perspective from an oblique angle suggests a sense of detachment and is reinforced (Kress & van Leeuwen, 2006). In this regard, the selection and portrayal of participants is commendable; by strengthening a feeling of collective identity and solidarity, these representations encourage students to identify with and participate in the images (Van Leeuwen & Jewitt, 2001). Also, it is made public at the representational It means nothing, such as favouring one specific profession, social class, or ethnic group over another. In other words, the participants' appearance reflects their respect for Indonesian religious customs. The images' localized representation of the practices is a great asset.

The first and final factor has to do with the image modality. The



modality is based on historically and culturally established criteria of what is real, as Kress & van Leeuwen (2006) pointed out. The technology used in EFL textbooks appears to be expanding quickly these days due to the rise of globalization. In addition to the represented pictures' naturalistic modality, this fact might help the students comprehend the scenes being depicted. A spectrum of color modality for naturalistic coding orientation is distinguished by Kress & van Leeuwen (2006). It extends from "full color saturation to the absence of color; that is black and white." Because the data are different, the results differ from those of earlier studies. In summary, the visual aids in the textbook can be seen as helping students meet the learning goals during the teaching and learning process. Particularly with the 2013 Curriculum, students are expected to be more engaged in independent exploration. The findings also suggest that the images support scaffold learning since they are a component of the overall meaning, as the verbal text and visual images work together to reveal meaning.

CONCLUSION

The conclusion is that all of the multimodal indicators representational, interactive, and compositional as defined by Cress and Van Leeuwen have been satisfied. The pictures in the book accurately depict things, circumstances, or ideas that complement the English language learning content, demonstrating the representational mode. Through the use of perspective, character expressions, and visual components that actively involve students in the comprehension process, the images' interaction with the readers reflects the interactive mode. In the meantime, the compositional mode, can effectively connect information to the book's coherent layout, color scheme, and text-image relationships.

Taking into account the findings and importance of this study, Researcher propose that future research, students, and English teachers will all benefit from the findings.

First and foremost, Researcher advise English teachers to get a better grasp of multimodality since it will be crucial for deciphering the pictures in the "English for Nusantara" textbook. Teachers who want to improve their teaching methods should focus on how multimodal elements are used in visual images.

Second, if students are familiar with multimodality, they will be better able to interact with the pictures. Even though multimodality may not be explicitly taught to students, teachers can help them understand the visuals by providing guidance. This will help them better understand the spoken texts that go along with the images.



Finally, it would be advantageous to interview students in order to increase the study's credibility for subsequent research. Deeper understanding of their comprehension and experiences with the multimodal components in the textbook may result from this.

REFERENCES

- Baddeley, A. D., & Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics*, 21(8), 627–635.
<https://doi.org/10.1080/00140137808931764>
- Chen, Y. (2010). The semiotic construal of attitudinal curriculum goals: Evidence from EFL textbooks in China. *Linguistics and Education*, 21(1), 60–74. <https://doi.org/10.1016/j.linged.2009.12.001>
- Christie, F. (2005). *Language education in the primary years*. UNSW Press.
- Creswell, J. W. (1994). *Research design: Qualitative and quantitative approach*. Sage Publications.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage Publications.
- Cunningsworth, A. (1995). *Choosing your coursebook*. Heinemann.
- Danes, F. (1974). Functional sentence perspective and the organization of the text. In F. Danes (Ed.), *Papers on functional sentence perspective* (pp. 106–128). Academic Press.
- Giaschi, P. (2000). Gender positioning in education: A critical image analysis of ESL texts. *TESL Canada Journal*, 18(1), 32–46.
<https://doi.org/10.18806/tesl.v18i1.902>
- Gibbons, P. (2015). *Scaffolding language, scaffolding learning: Teaching English language learners in the mainstream classroom* (2nd ed.). Greenwood Publishing Group.
- Halliday, M. A. K. (1978). *Language as social semiotic*. Edward Arnold.
- Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). Edward Arnold.
- Iedema, R. (2003). Multimodality, resemiotization: Extending the analysis of discourse as multi-semiotic practice. *Visual Communication*, 2(1), 29–57. <https://doi.org/10.1177/1470357203002001751>



- Kress, G., & van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. Routledge.
- Kress, G., & van Leeuwen, T. (2001). *Reading and writing with images: A review of four texts*. In *Reading images: The grammar of visual design*.
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.
- Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Arnold.
- Kress, G. (2009). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Kress, G., & van Leeuwen, T. (1992). Structures of visual representation. *Semiotica*, 89(4), 497–513. <https://doi.org/10.1515/semi.1992.89.1-2.497>
- Maffia, A. M. C., Dias, L. S. M., Brauna, R. C. A., & Cruz, R. (2003, August). Analysis of teachers' performance in the selection program of science textbook. Paper presented at the *Meeting of the European Science Education Research Association*.
- Marefat, F., & Marzban, S. (2014). Multimodal analysis of gender representation in ELT textbooks: Reader's perceptions. In *International Conference on Current Trends in ELT*. Allameh Tabataba'i University, Tehran, Iran.
- Merriam, S. B., & Tisdell, E. J. (2009). Dealing with validity, reliability, and ethics. In *Qualitative research: A guide to design and implementation* (pp. 209–235). Jossey-Bass.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Sage Publications.
- Moya, J., & Jesús, M. (2008). Compositional, interpersonal and representational meanings in a children's narrative: A multimodal discourse analysis. *University of Castilla-La Mancha, Spain*.
- Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific Region. *TESOL Quarterly*, 37(4), 589–613. <https://doi.org/10.2307/3588214>
- Salbego, N., Hebele, V., & Silva, M. (2015). A visual analysis of English textbooks: Multimodal scaffolded learning. *Universidade Federal de Santa Catarina*.
- Sugiono. (2017). *Metodologi penelitian kuantitatif, kualitatif, dan R&D*.



Alfabeta.

- Sweet, M. (2009). Writing learning objectives. University of Texas at Austin.
<http://www.utexas.edu/academic/diia/assessment/iar/students/plan/objectives/bloom-emphasize.php>
- Tarigan, H. G., & Tarigan, D. (1986). *Telaah buku teks bahasa Indonesia*. Angkasa.
- Tahririan, M., & Sadri, E. (2013). Analysis of images in Iranian high school EFL course books. *Sheikhbahae University*.
- Thibault, P. J. (1991). *Social semiotics as praxis: Text, social meaning making and Nabokov's Ada*. University of Minnesota Press.
- Tomlinson, B. (Ed.). (2023). *Developing materials for language teaching*. Bloomsbury Publishing.
- Van Leeuwen, T., & Jewitt, C. (Eds.). (2001). *The handbook of visual analysis*. Sage.
- Yasin, M., et al. (2012). A visual analysis of a Malaysian English school textbook: Gender matters. *Universiti Kebangsaan Malaysia*.