LANGUAGE DOCUMENTATION PRACTICES IN INDONESIA: A CRITICAL THEORY PERSPECTIVE

Satwiko Budiono[1]*; Selly Rizki Yanita[2]

[1-2]Research Center for Language and Literature Preservation, National Research and Innovation Agency
Jakarta, Indonesia 12710
[1]satwiko.budiono@brin.go.id; [2]selly.rizki.yanita@brin.go.id

ABSTRACT

The research examined the potential of language documentation practices in Indonesia. Using principles from critical theory, it argued that language documentation was different from language description, which was classified as monodisciplinary linguistics. Rather, language documentation had a broad scope of potential uses beyond descriptive linguistics. This was because language documentation could be linked to other linguistics fields such as sociolinguistics, oral tradition, oral history, and others. The core of language documentation was the natural language recordings that can be used in various fields, making language documentation part of interdisciplinary linguistics. Besides, multiple national and international institutions recognized its value. Meanwhile, language documentation in Indonesia was still closely linked to traditional descriptive or theoretical linguistics and was considered a data collection process rather than an independent field in linguistics. Furthermore, natural language recordings were still considered unscientific by the Indonesian academic community, especially in linguistics fields. This was because natural language recording could not be a final project for linguistics students or an output of linguistics research in Indonesia. Research output was still limited to products such as dictionaries, teaching materials, grammar, language maps, and so on. The research used a descriptive qualitative approach. Data collection included close reading books, journals, and other relevant academic materials. Therefore, the research argues that natural language recordings should be considered a valid scientific activity as a final project for linguistics students or the output of linguistic research in Indonesia. Thus, the number of natural language recordings in Indonesia will increase in national and international repositories. In fact, this can contribute to linguistic research because many linguistics scholars can utilize these language documentation results.

Keywords: critical theory, language documentation, descriptive linguistics, natural language recording, Indonesia academia

INTRODUCTION

Language documentation in Indonesia is still closely linked to language descriptions in the field of descriptive linguistics or theoretical linguistics. This is due to prevailing views about language documentation among language and linguistics researchers in Indonesia. Currently, most Indonesian language documentation focuses on phonology, morphology, and syntax (Moro, 2019; Sawaki, 2019; Klamer & Saad, 2020; Balukh, 2021). Many scholars treat language documentation as merely a method of collecting data for the descriptive linguistics field.

This view is also closely tied to the emergence of the digital era in which traditional descriptive linguistics has been employed to facilitate data processing. However, language documentation cannot be equated with descriptive linguistics.

The predominant view in Indonesia is that language documentation is part of descriptive linguistics. This is because language documentation is still not considered an independent field alongside descriptive or theoretical linguistics. Prior to the digital age, the Indonesian study of linguistics systems in descriptive and theoretical linguistics, such as phonology, morphology, and syntax, was

heavily monodisciplinary. Past scholars of descriptive linguistics often use questionnaires when collecting data (Handayani, 2019; Prihartono, 2019; Erniati, 2020; Garing et al., 2021). However, this data collection method is unnatural because the questionnaire is situational and prone to bias. Descriptive linguistics is also not considered of interest to Indonesian students or scholars and is often avoided compared to other approaches due to the aforementioned monodisciplinary nature of the field.

Things have changed since the emergence of the digital age. Language documentation is now considered an update of descriptive and theoretical linguistics. Preparations of language systems are now frequently carried out using an interdisciplinary language documentation approach (Amith, 2020). Language documentation has begun using natural data because language recording is more comprehensive (Seyfeddinipur & Rau, 2020). This is characterized by incorporating many types of language use, including oral tradition, conversation, poetry recitals, mantra chanting, and so on (Danerek, 2018). Nevertheless, digital-age language documentation in Indonesia is still considered unscientific by the Indonesian academic community, especially in linguistics fields. This is because natural language recording cannot be a final project for linguistics students or an output of linguistics research in Indonesia. Research output is still limited to products such as dictionaries, teaching materials, grammar, language maps, and so on. Besides that, current language documentation practice relies heavily on audiovisual language recordings, which are still not widely accepted as valid language documentation results by Indonesian scholars. This view is based on the aforementioned enduring traditional paradigms in the Indonesian linguistics field.

Meanwhile, there is still no Indonesian research on critical theory in relation to the field of linguistics. Fields that have incorporated research on critical theory include counseling (Indri, 2023), computer and information science (Yaniasih, 2020), sociology (Sholahudin, 2020), and cultural anthropology (Irwan et al., 2022). On the other hand, linguistic research outside Indonesia has used this critical perspective theory approach. It is starting from Dash (2024) who discussed the transition of the concept of traditional dialectology into language documentation; Schulze (2015), who linked it to sociolinguistics; Varela and Lee (2018), who discussed the concept of language documentation that can be used in performance arts, to Koster (2020) who related to second language acquisition. This condition makes this research something new in terms of data analysis because there has been no research on the critical theory that discusses language documentation practice in Indonesia, which is often associated with descriptive linguistics.

The substantive critical task of science is to reflectively analyze the social context that is embedded not only institutionally but also methodologically in scientific practice, which in turn allows people to

determine the accuracy of scientifically produced information (Ryen, 2020). Criticism here refers to continuous correction, improvement, and refinement (Zamiri & Esmaeili, 2024). Darling-Hammond et al. (2020) note that criticism will never be complete because all developments depend on special historical conditions.

Considering the significant lack of critical theory implementation in Indonesian linguistic research, this research examines language documentation in Indonesia from a critical theory perspective. Using critical theory, it examines and critiques prevailing Indonesian views of language documentation. Additionally, it explores the differences between language documentation, descriptive linguistics, and other fields and demonstrates that language documentation has long been misinterpreted and misunderstood in Indonesia.

METHODS

The research takes a descriptive qualitative approach. According to Creswell and Creswell (2018), this approach has different data collection, analysis, interpretation, and report writing methods from a quantitative approach. Qualitative methods can include the documentation of narratives, phenomena, ethnography, and case studies. In this research, descriptive qualitative methods suit the circumstances of language documentation research in Indonesia. Data collection includes close reading books, journals, and other relevant academic materials.

The collected works are then interpreted as per critical theory to evaluate whether they contain adequate representations of the current situation in Indonesia. This is done in accordance with the opinion of Oliva, Olivia, and Novara (2020) that critical theory can reveal the effectiveness of developing theory into practice in a given context. Sholahudin (2020) notes that a critical theory perspective can contribute to the development of awareness of social realities and emancipatory practices. This aligns with this research, in which Indonesians view language documentation and outline its unique qualities as a field. Thus, it seeks the liberation of Indonesian linguistics research from the shackles of the old, established paradigms that have formed prevailing academic thought and work patterns.

This research uses four principles as the foundation of its analysis (Freyenhagen, 2017; O'Mahony, 2023). First, critical theory does not just understand or interpret the world but to change it. Second, research must also question and challenge the assumptions that exist in the context of that research, reviewing older, established paradigms as part of its work. Third, research must be action-oriented with the goal of facilitating positive change or intervention. Finally, research does not just carry out criticism but must build or bring changes and explicitly recognize and promote emancipatory values. This principle

emphasizes that the research position must be clear and prove its usefulness. These four principles serve as the basis for the discussion in this research.

RESULTS AND DISCUSSIONS

As mentioned in the previous section, this discussion refers to the principles of critical theory as guidelines for analysis. Using these principles develops a critical approach for evaluating language documentation. This section first discusses the unique process of language documentation, followed by current practices in language documentation data collection. Finally, it explores the wide potential of language documentation use.

As explained in the introduction, As explained in the introduction, language documentation is still considered unequal to other linguistic fields in Indonesia, such as descriptive linguistics, sociolinguistics, dialectology, and others. However, language documentation is not part of descriptive linguistics; the two fields have markedly different approaches to data collection, analysis, and utilization of the results. However, it cannot be denied that language documentation produces more grammar than any other output, so language documentation is often identical to descriptive linguistics in Indonesia. Meanwhile, Fitzgerald (2021) suggests that the drive to language documentation is causing a shift in contemporary linguistics. This shows a new interest in the technology-based descriptive linguistics movement. Descriptive linguistics, which is considered old, is updated with language documentation that utilizes technology. In fact, language documentation can produce other outputs besides grammar based on descriptive linguistics.

The basic task of descriptive linguistics is the preparation of a language system (Voelkel & Kretzschmar, 2021), including the preparation of phonological systems, morphology, syntax, and orthography. Data collection often uses questionnaires as elicitation material in the form of words and sentences formulated according to the research target language. Analysis in descriptive linguistics is differentiated based on its level, and each level has its own analysis structure. For example, phonological analysis can be differentiated based on vowels, consonants, phoneme descriptions, sound distribution, suprasegmentals, morphophonemics, and so on (Rhodes & Campbell, 2018). Phoneme descriptions can be further divided into descriptions of vowels, consonants, and others.

In fact, the compilation of a language system included in the descriptive linguistics approach is categorized as a form of language conservation in Indonesia. Although the terms are different, the implementation of language conservation still refers to the descriptive linguistics approach. An interesting thing is the definition of documentation in the Regulation of the Minister of Education and Culture Number 42 of 2018 concerning national language and literature policies. The regulation explains that one of the efforts to preserve languages in Indonesia is documentation. More specifically, the definition of documentation in the regulation refers to efforts to compile language systems, recording, and developing conservation of endangered languages. Thus, it can be seen that language documentation, descriptive linguistics, and conservation are interpreted the same even though they are actually very different in implementation.

The process of language documentation is also distinct from descriptive linguistics and just as complex. In accordance with Austin and Himmelmann,
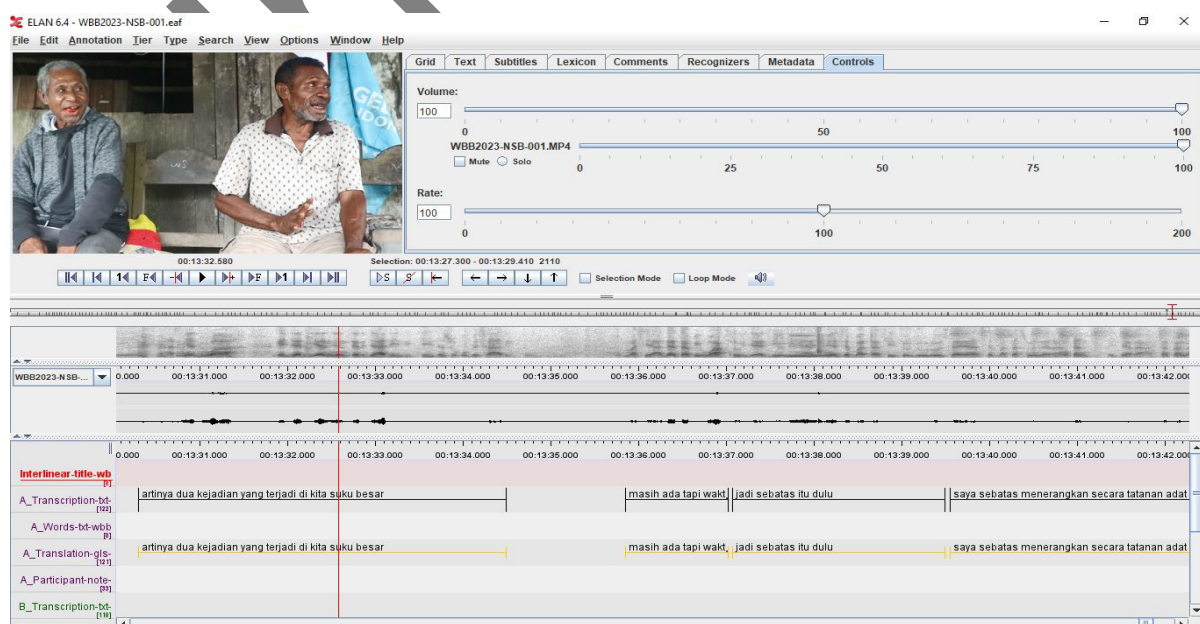

Figure 1 Data Transcription Process

the language documentation process includes (1) recording, (2) capture or digitalization, (3) data analysis or transcription, (4) archiving, and (5) mobilization (Rahima, 2024). This process makes it easier for researchers to obtain natural data from informants directly without any bias or intervention from factors outside the language itself. The digitalization process transfers analogue recordings to the digital domain. All recorded material is saved in digital form according to the type of language used. This is intended to make it easier for researchers to see the context of language use in a target language.

The data transcription process includes not only transcription but also translation, annotation, and metadata notation (Figure 1). Transcription involves writing down spoken language from recordings into phonetic and orthographic forms. After that, annotations and metadata notation assist in accurate data management. Archiving is the process of creating archive objects, determining access rights, and using the recording results. Usually, these results are entered into a national repository or international repository so that they can be preserved and accessed. Language records entered from national and international repositories require quite long and strict verification times, similar to scientific articles submitted to scientific journals. Finally, mobilization is the publication or distribution of materials.

Figure 2 Data Collection with Questionnaire

Figure 3 Data Collection with Recording

In language documentation, much of the current data collection practice refers to Himmelmann and Austin (Akpan, Urua, & Ekpenyong, 2018). Here, data collection does not use elicitation questionnaires as in descriptive linguistics (Figure 2). Instead, it takes the form of audio or video recordings of language uses like poetry recitals, songs, mantras, folk tales, instructions, lectures, conventions, traditional ceremonies, traditional games, and so on (Figure 3). Data recording techniques are adapted to field conditions and the type of data that needs to be recorded. For example, it is recording monologue folklore rather than recording activities involving multiple people, such as traditional ceremonies, theatrical performances, or conversations.

Based on the processes elucidated in the previous sections, language documentation is neither simple nor easy. Moreover, the need for language recordings to be able to accommodate all domains of language use is great. There are no definite and clear standards regarding the duration of recordings taken during language documentation. The more recordings obtained, the more language data can be retrieved. Even before entering the mobilization process, language documentation analysis is complicated. This is what makes language documentation not just a data collection technique.

Language documentation is considered part of interdisciplinary linguistics, as stated in the Regulations of the Head of the Indonesian Institute of Science Number 1 concerning guidelines for the selection of researchers' fields of expertise (2016). This is because language documentation has an enormous scope, including art, oral traditions, traditional games, local knowledge, and material artifacts like traditional houses. This broad scope makes it too complex to be just a data collection technique. Moreover, language documentation analysis also includes phonetic transcriptions in the fields of phonology, orthographic writing, and translation. Several aspects of language documentation analysis have become part of interdisciplinary analysis.

Language documentation can, therefore, be considered a scientific linguistic field independent of descriptive linguistics and other linguistics fields. As can be seen, it includes the same level of complexity in its techniques, data analysis, and utilization of the results.

The difference between the usage of The difference between the usage of language documentation and descriptive linguistics is highly misunderstood. Language description is only one of many possible uses of language documentation. The utilization of language documentation can also take other forms, such as in compiling dictionaries in the field of lexicography. Language documentation can also be used in education, such as in making teaching materials. The many uses of language documentation stem from its position as an interdisciplinary field of linguistics that overlaps well with other fields. Chelliah (2021) adds that it can be useful for providing materials for language learning, literacy,

and traditional knowledge. It can also provide the scientific community with a unique perspective on climate change, traditional ecological knowledge, musical traditions, biodiversity, and much more.

The use of language documentation starts with archiving language records. Once archived, they can be into public data banks, which can then be utilized in various fields. One example of a national repository is the Language Documentation of Indonesia (LADIN) under the National Research and Innovation Agency (BRIN Indonesia). Meanwhile, examples of international repositories include the Pacific and Regional Archive for Digital Source in Endangered Cultures (PARADISEC), which belongs to the Australian National University (Figure 4), and Scholar Space run by the University of Hawaii at Manoa (Figure 5), and the Endangered Language Archive (ELAR), courtesy of the Arcadia Foundations Endangered Language Documentation Program (Figure 6). The large number of international repositories indicates that language documentation is a field of linguistics that is of international interest and receives extensive funding for target language recording.

One source of international funding is the Endangered Language Documentation Program (ELDP), held annually. Recipients of ELDP funding receive language documentation training first in Berlin. After that, they carry out data collection of their target languages. The results have been entered into international repositories such as ELAR, PARADISEC, and so on. Even though ELDP funding recipients are not required or obligated to publish scientific articles, many do so.

In contrast, Indonesia's linguistic community still focuses on publications in scientific articles. As a result, language recordings are not archived and are only available to the researchers who made them rather than the wider national and international research community. However, storing and archiving language documentation by national and international repositories provides an example of good practice for future development. The former paradigm that has previously been established can be replaced with a new mechanism that emphasizes a wide range of benefits.

Language recordings could even provide an alternative to an undergraduate thesis, especially for students majoring in linguistics. This is in accordance with the independent learning program *Merdeka Belajar* within the framework of an independent campus called *Kampus Merdeka* (Directorate General of Higher Education, 2020). In this program, a thesis is not a requirement for student graduation. It can be replaced with a project or other relevant assignment according to the development and skills in their field. These conditions mean that students with a specialization in linguistics could create language recordings using a language documentation approach as a project instead of a thesis.

Figure 4 PARADISEC Repository

Figure 5 Scholar Space Repository

Figure 6 ELAR Repository

CONCLUSIONS

Based on these discussions, it is evident that language documentation is different from language description. Descriptive linguistics is only one result of using language documentation. The core of language documentation is a natural language recording that can be used in various fields, bringing language documentation into the realm of interdisciplinary linguistics. Language documentation has a broad scope, covering oral traditions, arts, local knowledge, traditional games, and material culture. It can thus be considered a unique linguistic field unto itself.

Language recording as a process and product of language documentation is still not considered a scientific activity in Indonesia. However, the process of analyzing language documentation in language recordings is complex and well-respected. As can be seen, many international funding organizations require

language recording as their research output rather than the publication of articles. If Indonesian scholars are free from the long-established views disregarding language recordings, language documentation practices developed in the digital age could be freed from the shadow of descriptive linguistics.

In fact, language recording can enhance the emancipatory value of language documentation in general, allowing natural language recording to be considered a scientific activity in Indonesia as a final project for linguistics students or the output of linguistic research. Thus, the number of natural language recordings in Indonesia will increase in national and international repositories. Hopefully, this can contribute to linguistic research because many linguistics scholars can utilize this language documentation results in many ways and other approaches or disciplines.

REFERENCES

Akpan, A., Urua, E.-A., & Ekpenyong, M. (2018). Bridging the gap between language documentation and description for effective linguistic research. *Journal of the Linguistics Association of Nigeria, 3*, 9-20.

Amith, J. (2020). Endangered language documentation: The challenges of interdisciplinary research in ethnobiology. *Language Documentation & Conservation Special Publication: Interdisciplinary Approaches to Language Documentation, 21*, 72-112.

Balukh, J. (2021). A grammar of Dhao: An endangered Austronesian language in Eastern Indonesia. *Wacana, 22*(2), 522-533. https://doi.org/10.17510/wacana.v22i2.1005.

Chelliah, S. L. (2021). *Why language documentation matters*. Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-030-66190-8

Creswell, J. W., & Creswell, J. D. (2018). *Resign design: Qualitative, quantitative and mixed methods research. In Introducing English Language* (Fifth Edit). Singapore: Sage Publications.

Danerek, S. (2018). Documentation of Palu'e; Storytelling and folklore. *Wacana: Journal of the Humanities of Indonesia, 18*(3), 718-745. https://doi.org/10.17510/wacana.v18i3.634.

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science, 24*(2), 97-140. https://doi.org/10.1080/10888691.2018.1537791.

Dash, N. S. (2024) Documentation and digitization of endangered indigenous languages: Methods and strategies. *Proceedings of the International Mother Language Day 2024*. Colombo, Sri Lanka. pp. 5-24.

Directorate General of Higher Education (Direktorat Jenderal Pendidikan Tinggi). (2020). *Buku panduan merdeka belajar Kampus Merdeka.* Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. https://dikti.kemdikbud.go.id/wp-content/uploads/2020/05/Buku-Panduan-Merdeka-Belajar-Kampus-Merdeka-2020-1.pdf.

Erniati, N. (2020). Sistem fonologi bahasa Melayu dialek Ambon. *BEBASAN Jurnal Ilmiah Kebahasaan Dan Kesastraan, 6*(2), 113-124. https://doi.org/10.26499/bebasan.v6i2.116.

Fitzgerald, C. M. (2021). A framework for language revitalization and documentation. *Language, 97*(1), e1–e11. https://doi.org/10.1353/lan.2021.0006.

Freyenhagen, F. (2017). Critical theory's philosophy. In *The Cambridge Companion to Philosophical Methodology* (pp. 356–378). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/9781316344118.019.

Garing, J., Agus, N., Arisnawati, N., & Mappau, R. (2021). Sistem fonologi bahasa limola/Phonology system in limola language. *Aksara, 33*(1), 153-168. https://doi.org/10.29255/aksara.v33i1.671.153-168.

Handayani, R. (2019). Penyusunan sistem morfologi sebagai upaya pendokumentasian bahasa: Prefiksasi bahasa budong-budong. *Sawerigading: Jurnal Bahasa dan Sastra, 25*(2), 99-106. https://doi.org/10.26499/sawer.v25i2.606.

Indri, M., D. B. (2023). Filsafat ilmu bimbingan dan konseling di Indonesia: Telaah kritis teori dan praktis. *Chalim Journal of Teaching and Learning (CJoTL), 1*(2), 209-217. https://doi.org/10.31538/cjotl.v1i2.371.

Irwan, I., Siska, F., Zusmelia, Z., & Meldawati, M. (2022). Analisis perubahan peran dan fungsi keluarga pada masyarakat Minangkabau dalam teori feminisme dan teori kritis. *Satwika: Kajian Ilmu Budaya dan Perubahan Sosial, 6*(1), 191-205. https://doi.org/10.22219/satwika.v6i1.19383.

Klamer, M., & Saad, G. (2020). Reduplication in Abui: A case of pattern extension. *Morphology, 30*(4), 311-346. https://doi.org/10.1007/s11525-020-09369-z.

Koster, J. (2012). Theories of language from a critical perspective. In *The Cambridge Handbook of Second Language Acquisition* (pp. 9-25). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139051729.003.

Moro, F. R. (2019). Loss of morphology in Alorese (Austronesian): Simplification in adult language contact. *Journal of Language Contact, 12*(2), 378-403. https://doi.org/10.1163/19552629-01202005.

O'Mahony, P. (2023). Introduction to special issue: The critical theory of society. *European Journal of Social Theory, 26*(2), 121-135. https://doi.org/10.1177/13684310231154081.

Oliva, A., Olivia, A., & Novara, I. (2020). *Marx and contemporary critical theory: The philosophy of real abstraction.* London, UK: Palgrave Macmillan.

Prihartono, W. (2019). Sistem fonologi dalam bahasa nedebang. *Jurnal Lingko: Jurnal Kebahasaan dan Kesastraan, 1*(1), 32-54. http://dx.doi.org/10.26499/jl.v1i1.27.

Rahima, A. (2024). Language documentation as a revitalization effort regional languages are starting to extinct. *International Journal of Multidisciplinary Sciences and Arts, 3*(3), 72-76. https://doi.

org/10.47709/ijmdsa.v3i3.4123.

Republic of Indonesia. (2016). *Regulations of the Head of the Indonesian Institute of Science Number 1 concerning guidelines for the selection of researchers' fields of expertise.* Jakarta, Indonesia: State Secretary.

Republic of Indonesia. (2018). *Regulation of the Minister of Education and Culture Number 42 of 2018 concerning national language and literature policies.* Jakarta, Indonesia: State Secretary.

Rhodes, R., & Campbell, L. (2018). The goals of language documentation. In K. L. Rehg & L. Campbell (Eds.), The *Oxford Handbook of Endangered Languages* (pp. 106-122). Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190610029.013.7.

Ryen, E. (2020). Klafki's critical-constructive didaktik and the epistemology of critical thinking. *Journal of Curriculum Studies, 52*(2), 214-229. https://doi.org/10.1080/00220272.2019.1657959.

Sawaki, Y. W. (2019). Meneropong tipologi bahasa-bahasa di Papua: Suatu tinjauan singkat. *Linguistik Indonesia, 36*(2), 129-143. https://doi.org/10.26499/li.v36i2.79

Schulze, I. (2015). Factorizing the sociocultural dimension of language documentation. *Iran and the Caucasus, 19*(4), 311-333. https://doi.org/10.1163/1573384X-20150403.

Seyfeddinipur, M., & Rau, F. (2020). Keeping it real: Video data in language documentation and language archiving. *Language Documentation & Conservation, 14*, 503-519.

Sholahudin, U. (2020). Membedah teori kritis mazhab Frankfurt: Sejarah, asumsi, dan kontribusinya terhadap perkembangan teori ilmu sosial. *Journal of Urban Sociology, 3*(2), 71-89. https://doi.org/10.30742/jus.v3i2.1246.

Varela, M. E., & Lee, N. H. (2018). Language documentation: A reference point for theatre and performance archives? *International Journal of Performance Arts and Digital Media, 14*(1), 17-33. https://doi.org/10.1080/14794713.2018.1453242.

Voelkel, S., & Kretzschmar, F. (2021). Language documentation and descriptive linguistics. In *Introducing Linguistic Research* (pp. 79-105). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/9781316884485.006.

Yaniasih, Y. (2020). Teori kritis terhadap analisis sitasi untuk kajian kuantitatif sains dan evaluasi kinerja riset. *Berkala Ilmu Perpustakaan Dan Informasi, 16*(1), 127-141. https://doi.org/10.22146/bip.v16i1.72.

Zamiri, M., & Esmaeili, A. (2024). Methods and technologies for supporting knowledge sharing within learning communities: A systematic literature review. *Administrative Sciences, 14*(1), 1-34. https://doi.org/10.3390/admsci14010017.