KnE Social Sciences
IICB
1st Indonesian International Conference on Bilingualism
Volume 2023

Research Article

A CORPUS-BASED STUDY ON THE DEVELOPMENT OF THE JAVANESE LANGUAGE AMONG MULTILINGUAL CHILDREN IN EAST JAVA

Evynurul Laily Zen, Aulia Apriana

Universitas Negeri Malang, Malang, Indonesia

Abstract.

This paper seeks to examine multilingual children's Javanese across different grades of primary schools with a focus on proficiency, vocabulary, and grammatical development. The data were partially taken from CBLING (Corpus of Bilingual Learners' Languages), a corpus developed by Universitas Negeri Malang. Using AntConc as a corpus development tool, we measured Javanese proficiency using the Mean Length of Utterances (MLU) method. Our findings on the development of Javanese proficiency and vocabulary growth indicated insignificant improvement across grades. However, the average vocabulary density grows significantly from the first two lower grades to the three higher ones. In terms of grammar, hybrid words showing morphological blendings, such as *mastekno* or *isikno* were observed in the participants' writings. Taken together, our findings are essential for teaching and learning Javanese as a regional home language and generating an important linguistic feature of 21st-century multilingual children's Javanese.

Keywords: *corpus study, Javanese language, multilingual children, East Java*

Corresponding Author: Evynurul Laily Zen; email: evynurul.laily.fs@um.ac.id

Published 27 April 2023

Publishing services provided by Knowledge E

Selection and Peer-review under the responsibility of the IICB Conference Committee.

1. INTRODUCTION

Javanese being the language with the largest number of native speakers in Indonesia has received scholarly attention with major explorations being centred on its description. In the area of phonology, Dudas (1946), Yallop (1982), Fagan (1988), Hayward and Muljono (1991), and Hayward (1993) have conducted studies exploring the rich and fascinating sound patterns of Javanese. Meanwhile, Horne (1961), Dudas (1946), and Uhlenbeck (1978) have focused their works on the area of morphology. Furthermore, Sumukti (1971), Suharno (1982), and Robson (2002) have investigated the syntactic patterns in this language. Then, in the area of socio-pragmatics, Kartomihardjo (1971), Djajengwasito (1975), and Wolff and Poedjosoedarmo (1982) investigated the practical use of the Javanese language.

These early theoretical groundings on the Javanese language have focused on idealized monolingual speakers where it was not only supposed to be but was also still possible to find Javanese monolingual speakers. However, Indonesia's linguistic situation has undergone a significant shift. The rise of pan-national identity after 1945 independence – almost two decades after the first national movement initiated by the Indonesian Youth Congress – proclaimed the use of Indonesian as the language of unity in 1928. The official status of Indonesian has consequently brought a huge impact on the existence of the regional home language.

As far as the regional language shift is concerned, studies have reported the shift of domain of use of Indonesian from public to private, from schools and offices to homes and neighbourhoods. Specifically, the 1990 national census reported an increase in native speakers of Indonesian and a decrease in native speakers of the regional home language with globalization and global mobility taking this language-change phenomenon to the next level. Therefore, further scientific works that scrutinize the acquisition of regional home language among young children become necessary.

Our current paper takes Javanese as the point of departure and intends to see how this language is acquired among primary-school learners in East Java as the home province of the Javanese language. In particular, we look at vocabulary development because vocabulary or word has been seen to constitute the building blocks for language production. We also seek to explore the nature of grammatical development among our participants. Here, the acquisition of grammatical knowledge in a multilingual setting is highly possibly extended to the domain of cross-linguistic transfer. Therefore, we assume that our participants' background languages – Indonesian and/or English – will influence the production of the target language, which is Javanese. More importantly, we extend our analysis to compare the learners' Javanese production across grades.

2. METHOD

This corpus-based study makes use of some parts of CBLING (Corpus of Bilingual Learners' Languages) – a corpus developed by Universitas Negeri Malang (UM) containing a collection of both spoken and written production of primary school children's languages in East Java. At the current stage, CBLING has stored approximately 576 recorded spoken productions and 2.530 written productions in Javanese, English, and Indonesian. With a specific mention of the written datasets, CBLING has compiled 220.090 tokens in total and continues to collect more data in the future.

For our current analysis, however, we focused on the written production of Javanese among first to fifth graders. The written production we use here was based on the production of a personal narrative on the topic of 'my favourite toy(s)'. In locating the intended language features, we utilized a corpus development tool of AntConc. We also measured the Javanese proficiency of these children using the Mean Length of Utterances by Word (MLUw) method.

3. FINDINGS AND DISCUSSION

In this study, we look at Javanese proficiency, vocabulary, and grammatical development across grades in which the key findings are elaborated below.

3.1. Javanese Proficiency of Multilingual Primary-School Learners

As aforementioned, in measuring the Javanese proficiency of children under investigation we specifically used the MLU by Word (MLUw) method, which is a measure of linguistic productivity and widely used in language acquisition studies (Allen & Dench, 2015; Dethorne, Johnson, & Loeb, 2005; Parker & Brorson, 2005; Voniati, 2016). The MLUw is operated by dividing the total number of words by the total number of complete and intelligible utterances in each of the participants' language production. The results in Table 1 explain that Javanese proficiency across different grades varied insignificantly, where gaps in scores between grades tend to be small. For example, the gap in the average score of Grade 1 and Grade 2 learners is 1.14 points, meaning that Grade 2 learners are only 1.14 higher than their Grade 1 peers. The point of improvement across two grades – from Grade 1 to Grade 3 – does not also show a significant increase, which is 1.53 points. As it is seen that the increase in average scores from Grade 1 to Grade 5 is 3.66 points, we can assume that the learners' Javanese proficiencies do not improve significantly, despite long hours of learning in schools.

Table 1: Results of MLUw Scores.

Grades	Average of Word Types	Average of Tokens	Average of Utterances	Average of MLUw Scores
1	21	33	5	6.69
2	23	38	5	7.83
3	36	60	7	8.22
4	39	65	8	8.87
5	35	61	6	10.35

However, when we look at the aspect of productivity – word types, tokens, and size of utterances – we see that Grade 5 learners produced more varied types of words denoting that the higher the grade the more varied the words one could produce, even though it does not seem to increase significantly. Similarly, concerning word tokens or the total number of words in one single essay, we observe that Grade 5 learners were able to produce a greater number of tokens compared to Grade 1 learners despite the same length of utterances between these two grades.

3.2. Vocabulary Development of Javanese among Multilingual Primary-School Learners

Concerning vocabulary development, we analyse three aspects: diversity, density, and sophistication (Perfetti & Hart, 2001). In the first aspect, we look at a broad diversity of vocabulary in one complete sentence in a piece of writing. In the corpus tool, the aspect of diversity is indicated by word types. Table 2 presented the results of our analysis where the gap of an average number of word types produced by the lowest and the highest grade was as varied as 14 points.

Table 2: Vocabulary Diversity across Grades.

Grades	Diversity (Word Types)	Density (Token Size)
1	21	33
2	23	38
3	36	60
4	39	65
5	35	61

Looking more closely at the datasets, we found that the three most frequently used content words in all essays are *aku* 'I', *dolanan* 'toys', and *seneng* 'happy' – pronoun, noun, and adjective respectively – while for the function words, the three most frequently used words are the conjunction *lan* 'and', the demonstrative *iku* 'that', and the adjective clausal marker *sing* 'which'. Regarding the production of nouns especially those related to the main topic of the writing task, our findings interestingly indicate a proportional divide between traditional and modern toys or games. For the traditional ones, the frequency of word occurrence as reported in the corpus analysis is *bal-balan* (279), *dakon* (138), *enklek* (92), *layangan* (56), *golekan* (39), *gobak sodor* (36), *tekongan* (33), *egrang* (23), *delikan* (20), *congklak* (12), *singitan* (10), *bentengan* (9), *jumpritan* (6), and *gangsing* (5), whereas the order of frequency for the modern toys or games is *lego*

(124), *boneka* (117), *game* (58), *slime* (50), *hp* (46), *squishy* (29), *uno* (27), *barbie* (24), *tablet* (9), *monopoli* (9), and *hot wheels* (6).

For the second aspect of vocabulary development, we analyse the level of density. In this aspect, we measure the quantity of content vocabulary by looking at the result of token analysis in the corpus tool (see Table 2). Table 2 informs us that the average of word tokens grows significantly from the first two lower grades (1 and 2) to the three higher grades (3 to 5). In Table 3, we also provide a summary of the highest and lowest token size in each of the grades investigated, with the results demonstrating a slow but steady improvement in vocabulary development.

TABLE 3: Vocabulary Density across Grades.

Grades	Highest Token Size	Lowest Token Size
1	70	4
2	129	10
3	136	22
4	116	27
5	121	25

In the last aspect, we explore the nature of vocabulary sophistication among these young Javanese learners. Vocabulary sophistication refers to the ability of learners to incorporate infrequent and more complex vocabulary in their writing. In our analysis, we determine sophisticated words by removing (1) non-dictionary words, (2) most common words, and (3) Indonesian words. This way, we find rare words which we analyze qualitatively. In this context, we identify those rare words produced by Grade 1 learners including *watu* 'stone', *wakeh* 'many/much', *uwenak* 'so delicious', and *tapekno* 'however'. In Grade 2 writings, we found *wesoh* 'wash hand', *wenehi* 'to give', *wedine* 'worried', *tokno* 'expose', *tibake* 'surprisingly', in Grade 3 *tumbasno* 'buy me', *tetulung* 'to help each other, and *ngentekake* 'to finish', in Grade 4 learners produced *telung* 'three', *suek* 'torn', *sepuro* 'sorry', *nyegat* 'get in the way', *nukokno* 'buy for', and in Grade 5 *sakmolene* 'after returning', *pungkasane* 'at the end', *nyritakno* 'to tell a story', *nggarai* 'cause', and *ngetokno* 'to bring out.

In the course of vocabulary development, prior works have indicated the interplay between vocabulary knowledge and reading skills as well as teaching instructions (Quinn, Wagner, Petscher, & Lopez, 2014; Silverman et al., 2013). Furthermore, vocabulary growth is sensitive to input and experience (Hart & Risley, 1992). Here, we see that our corpus-based findings can inform respective language teachers about the stages of vocabulary development of their students so teachers can provide rich language input as well as relevant vocabulary instructions.

3.3. Grammatical Development of Javanese among Multilingual Primary-School Learners

For this section, we observe the production of temporal markers and morphological mixing that appear to be unique. In the first point of observation, the most appeared temporal markers in our participants' production are perfective and past marker '*wes/uwes/uwis*' (Example 1 – 2) and future marker '*arep*' (Example 3), with acceptable word order.

Table 4

1	*Koncoku*	*wes*	*iso*	*sepatu rodaan*
	Friend-my(POSS)	have(PER)	able	roller skate
	My friend has been able to play roller skate			

Table 5

2	*aku*	*wes*	*njenengi*	*golekanku*
	I	have(PER)	name	doll-my(POSS)
	I have given my doll a name			

Table 6

3	*koncoku*	*arep*	*nyilih*	*bonekaku*
	friend-my(POSS)	will	borrow	doll-my(POSS)
	My friend will borrow my doll			

The second grammatical item we found interesting is morphological mixing in which our participants tended to mix the free morpheme of one language with the bound morpheme of the other language. Presented in Table 4, we can see that Indonesian's free forms typically function as the base words to where Javanese affixes are attached to.

It is worth noting, however, the Indonesian words being mixed are commonly non-standard or colloquial. For example, participants have modified the standard verb of Indonesian *melanjutkan* into its colloquial form of *ngelanjutkan*, then the affix *-kan* is substituted by the equivalent affix no to create a strong nuance of 'Javanese'. Similar morphological processes also occur for *nggantikno* 'to substitute', *mastekno* 'to ensure', and *isikno* 'to fill in'.

TABLE 7: Morphological Mixing.

Words	Morphemes	Mixed Elements
nggantikno 'menggantikan'/'to substitute'	[ganti] + [-no]	Nonstandard Indonesian *'ngganti'* and Javanese suffix [-no]
mastekno 'memastikan'/'to ensure'	[pasti] + [-no]	Nonstandard Indonesian *'masti'* and Javanese suffix [-no]
ngelanjutno 'melanjutkan'/'to continue'	[lanjut] + [-no]	Nonstandard Indonesian *'ngelanjut'* and Javanese suffix [-no]
katane 'katanya'/'it is said'	[kata] + [-ne]	Indonesian *'kata'* and Javanese suffix [-ne]
semuane 'semuanya'/'all/everything'	[semua] + [-ne]	Nonstandard Indonesian *'semua'* and Javanese suffix [-ne]
isikno 'isikan'/'to fill in'	[isi] + [-no]	Indonesian *'isi'* and Javanese suffix [-no]

4. CONCLUSION AND RECOMMENDATION

To conclude, our analyses demonstrated that Javanese proficiency does not seem to improve significantly across grades. Furthermore, while the vocabulary diversity as indicated by the average number of word types does not show a significant difference, the vocabulary density shown by the average size of word tokens is found to differ significantly between the lower-grade learners (Grade 1 and 2) and the higher-grade learners (Grade 3 – 5). Regarding grammatical development, hybrid words in the form of morphological mixing have been featured in our participants' production. Here, the hybrid words become interesting evidence of cross-linguistic transfer.

Taken together, our findings are essential for heritage language pedagogy and practices, particularly in the context of Javanese. Also, our findings can be used to generate an important linguistic feature of Javanese among 21st-century multilingual children. In this way, our study may therefore be used as an empirical basis to predict the outlook of Javanese in Indonesia.

References

[1] Allen SE, Dench C. Calculating mean length of utterance for Eastern Canadian Inuktitut. First Lang. 2015;35(4–5):377–406.

[2] Dethorne LS, Johnson BW, Loeb JW. A closer look at MLU: what does it really measure? Clin Linguist Phon. 2005 Dec;19(8):635–48.

[3] Djajengwasito S. Javanese speech style: A multiple discriminant analysis of social constraints (Doctoral dissertation, Cornell University, Ithaca, New York, United States).

[4] Dudas KM. The phonology and morphology of modern Javanese (Doctoral dissertation, University of Illinois, Urbana-Champaign, Illinois,United States).

[5] Fagan JL. Javanese intervocalic stop phonemes: The light/heavy distinction. In: McGinn R, editor. Studies in Austronesian Linguistics. Athens: Ohio University Center for International Studies; 1988.

[6] Hayward K, Muljono KM. The dental/alveolar contrast in Javanese. Bull Sch Orient Afr Stud. 1991;54(1):126–44.

[7] Hayward K. /p/ vs. /b/ in Javanese: Some preliminary data. SOAS Working Papers in Linguistics and Phonetics. 1993;3:1–99.

[8] Hart B, Risley TR. American parenting of language-learning children: persisting differences in family-child interactions observed in natural home environments. Dev Psychol. 1992;28(6):1096–105.

[9] Horne EC. Beginning Javanese. New Haven, London: Yale University Press; 1961.

[10] Kartomihardjo S. Ethnography of communication codes in East Java. Canberra: The Australian National University; 1971.

[11] Parker MD, Brorson K. A comparative study between mean length of utterance in morphemes (MLUm) and mean length of utterance in words (MLUw). First Lang. 2005;25(3):365–76.

[12] Perfetti CA, Hart L. The lexical basis of comprehension skill. In: Gorfein DS, editor. Decade of behavior. On the consequences of meaning selection: Perspectives on resolving lexical ambiguity. Washington (D. C.): American Psychological Association; 2001. pp. 67–86.

[13] Quinn JM, Wagner RK, Petscher Y, Lopez D. Developmental relations between vocabulary knowledge and reading comprehension: a latent change score modeling study. Child Dev. 2015;86(1):159–75.

[14] Robson S. Javanese grammar for students. Melbourne: Monash University Press; 2002.

[15] Silverman RD, Proctor CP, Harring JR, Doyle B, Mitchell MA, Meyer AG. Teachers' instruction and students' vocabulary and comprehension: an exploratory study with English monolingual and Spanish-English bilingual students in grades 3-5. Read Res Q. 2013;49(1):31–60.

[16] Suharno I. A descriptive study of Javanese. Canberra: The Australian National University; 1982.

[17] Sumukti RH. *Javanese morphology and morphophonemics* (Doctoral thesis, Cornell University, Ithaca, New York).

[18] Uhlenbeck EM. Studies in Javanese morphology. The Hague: Martinus Nijhoff; 1978.

[19] Voniati L. Mean length of utterance in Cypriot Greek-speaking children. J Greek Linguist. 2016;16(1):117–40.

[20] Wolff JU, Poedjosoedarmo S. Communicative codes in Central Java. New York: Cornell University; 1982.

[21] Yallop C. The phonology of Javanese vowels. A. Halim, L. Carrington, S.A. Wurms (Eds.). Third International Conference in Austronesian Linguistics. 1982;4:299–319.