Problems with Formalistic Grammars in Analysing Colloquial Indonesian

Evan Keith

1 Introduction

The influence of modern urban colloquial Indonesian, or Bahasa Gaul, on standard Indonesian is a well-documented phenomenon (see Smith-Hefner, 2007, Sneddon, 2016). Despite this, extant descriptive grammars of Indonesian have tended to focus almost exclusively on the standardised formal written variety of the language. This has led to an underemphasis on describing the grammar and discourse patterns of Bahasa Gaul, even where these differ substantially from the standard language. This article will demonstrate this point by analysing a text in Bahasa Gaul according to the rules of grammars written by Dyen (1967), Kridalaksana et al. (1985), and Sneddon et al. (1996). The three grammars were deliberately selected so as to cover a diversity of time periods and authors. Each is separated by more than a decade from the others, and although each is written with a different audience in mind, each one purports to provide a 'descriptive' analysis of modern Indonesian grammar. The *Bahasa Gaul* sample text is taken from an online rant posted to Facebook in 2017, as quoted in a paper by Swandy (2017) on the unique discursive practices of *Bahasa Gaul* in online spaces.

2 Background

Bahasa Gaul, hereafter referred to simply as Gaul, is a blanket term used to refer to many informal varieties of Indonesian, spoken primarily in large urban centres such as Jakarta and Yogyakarta (Smith-Hefner, 2007). Swandy (2017) notes that the drivers behind the development of early Gaul in the Suharto era tended to be disadvantaged youth, in particular the urbanised poor, members of street gangs, and Indonesia's LGBT community. These early forms of Gaul were characterised by the development of new slang vocabulary through intentionally obtuse morphological patterns (in order to avoid being understood by non-members of the aforementioned groups), the simplification of complex and compound verbs, and heavy, unpredictable elision (Smith-Hefner, 2007; Kusuma and Mardijono, 2013. Gaul's origin from amongst sections of Indonesian society often considered 'undesirable' by mainstream society has historically contributed to a perception of Gaul as being unworthy of research or investigation. (Smith-Hefner, 2007) However, despite the best efforts at standardisation by the government of the Suharto era as discussed by Smith-Hefner (2007), the influence of Gaul on so-called 'standard' Indonesian is rapidly growing, even outside the communities where Gaul was originally spoken (Swandy, 2017; Smith-Hefner, 2007). Both the tendency towards elision of affixes

in early 'street' forms of Gaul and the influence of Chinese languages in more recent years have resulted in a shift towards Gaul being relatively isolating, in contrast to standard Indonesian, which is generally considered an agglutinative language (Kridalaksana et al., 1985; Tadmor, 2007). For these reasons, among others, the grammar of Gaul is different, sometimes radically so, from standard Indonesian.

The various forms of Gaul in common use today are undoubtedly different from the language described in early formalistic grammars of Indonesian. Pronoun lists given in Dyen (1967) and Sneddon et al. (1996) do not include, even in passing, the Hokkien-derived set of informal pronouns which Gaul uses as its standard repertoire. When Sneddon et al. (1996, p. 168) does acknowledge the existence of the informal pronoun set, he does so only in passing, providing only one example and admitting the incompleteness of his own review. However, it would be mistaken to conceive of Gaul as an abhorrent or divergent vernacular. Given the status of Indonesian as a second language to the vast majority of its speakers, slang and informal forms are especially vital in driving linguistic change, including grammatical change (Smith-Hefner, 2007). It therefore seems vital that a truly 'descriptive' grammar of modern Indonesian engage with those grammatical features which are characteristic of Gaul. Unfortunately, this has not been the case historically. To elucidate this point, this article will examine the failings of otherwise sturdy models of standard Indonesian grammar in analysing some sentences in Gaul.

3 Analysis

The following is an analysis of several sections of text from a rant posted on Facebook by an anonymous source, as quoted in Swandy (2017). The glosses for each excerpt are written according to the specifications of each grammar. Wherever the gloss does not align with the translation given, this is indicative of problems in the analysis suggested by the relevant grammar, as discussed in the text following the analysis. Note that this article uses the standard, 'perfected' orthography (*Ejaan Yang Disempurnakan*) as in common use after 1973 to render all Indonesian text. Where quoted texts written before 1973 have used other orthographies, and where informal online sources have used abbreviations and other 'text-speak' that does not reflect how a passage would be spoken if read aloud, this has been corrected to match the modern standard spelling.

3.1 Analysis Using Dyen (1967)

Figure 1: an excerpt of text from Swandy (2017), glossed using Dyen (1967)
Note: null signs represent morphemes absent in respective grammars.

Mem-per-tahan-kan	*sese-orang*	*itu*	*tidak*	*mudah.*
ACT-TR-withstand-TR	some.quantity-[+human]	those	NEG	easy.

It's not easy to stand some people.

Sering	*di-sakit-i*	*di-kecewa-in*	*di-bohong-i,*
Often	PASS-sick-frequently	PASS-dissappoint-Ø	PASS-lie-frequently

I often get hurt, let down, lied to,

hanya bisa sabar aja.
only can patient Ø
but I just have to be patient.

Banyak yang bilang bodoh.
Many which say stupid
There are many (people) who call me stupid.

Parsing these sentences using Dyen's rules as a guide, several strange things occur, all of which imply that Dyen's model is ill-suited to understanding Gaul. Perhaps the most easily remedied aspect of Dyen's analysis is the fact that he tends to analyse the meaning of multi-morphemic affixes as created by their underlying individual morphemes, rather than treating the meaning of these complex affixes as unique to each permutation. For instance, when analysing the first word in the text, *mempertahankan*, a transitive verb derived from the lexically ambiguous root *tahan*, Dyen's model would suggest that there are two constituent parts to the complex affix *memper- -kan*. These are: the circumfix *per- -kan*, itself consisting of the morphemes *per-* and *-kan* which each occur independently and in other complex affixes; and the prefix *me(N)-*, the N being a nasal consonant which shifts to the first place of articulation in the root. *Per- -kan* is called a 'transitiviser' by Dyen (p. 245), and *me(N)-* is referred to as an 'active' marker (p. 244). This analysis goes against the traditional understanding of *memper- -kan* as a single complex circumfix, an analysis more in line with elision patterns of Gaul. Given the tendency of Gaul to elide morphemes within these constructions without creating changes in meaning, it makes more sense to think of the complex affix *memper- - kan* as separate in meaning from its constituent morphemes; as a single circumfix whose elements can be elided at will to alter formality without changing its grammatical function. Given Dyen's willingness to analyse *per- -kan* as a circumfix despite the independent but related functions of both *per-* and *-kan* as monomorphemic affixes (pp. 244-245), it seems strange that he does not take it one step further and analyse *memper- -kan* as a circumfix.

The next sentence in the excerpt seems almost foolish to analyse with Dyen's grammar, since it is in the passive construction yet elides its object pronoun. Dyen does provide a model, although it is not particularly well-defined, of grammatical 'environments' (p. 57), in which an element or feature can be carried over from one sentence to the next and provide clues as to what has been elided. He uses this model to explain how, in colloquial use, a complex question can be answered with simply an aspect marker and a verb, leaving features like person and number to context since they have been specified in the 'environment' of the previous sentence. However, even in this model, this information must be specified in a prior sentence for it to be counted as part of the environment. Since the previous sentence contains no immediate indication of person or number, this would either need to be given by a null morpheme in the previous sentence, or the 'environment' model dropped. This explanation is indicative of a broader trend of explaining away a lack of explicit person, number, and TAM with reference to elision and context clues, rather than entertaining the notion that these features may not be encoded at all.

Reaching our final sentence in the excerpt, Dyen's syntax breaks down entirely. Parsing the final sentence in a way coherent with its actual meaning using the rules given by Dyen is impossible. To understand how Dyen's syntax fails to parse this sentence correctly, it is informative to look at a 'translation' of this sentence in more formal Indonesian. If we try to rewrite this sentence in formal language while altering a few words as possible, we get something resembling the following:

Ada	*banyak*	*yang*	*meng-ata-kan*	*saya*	*bodoh.*
Exist	many	which	ACT-word-BEN	1SG	stupid

There are many (people) who call me stupid.

Dyen's syntax handles this adapted sentence with ease. The addition of *saya* means that we can straightforwardly analyse *yang bilang saya bodoh* for what it is: a nominal expression. If this is the case, Dyen's syntax allows us to stack nominal expressions recursively, which enables us to parse *banyak yang bilang saya bodoh* as what Dyen (p. 70) terms a "yang-modified nominal", a more complex version of a nominal expression (note that Dyen was writing before the advent of X-bar theory). Add the existential quantifier *ada* to this nominal expression, and the sentence makes perfect sense. A literal understanding of this sentence by this analysis might sound more like 'there exist many (people) who say (that) I am stupid', the central verb being *ada* rather than *bilang*, but this still scans on a phrase-structure level. However, this analysis is not in line with elision patterns in colloquial use, as shown by the elided version in the original Gaul text. The fact that the subject pronoun is elided in the informal variant means that *bodoh* by itself cannot be analysed as a verb phrase (at least not without positing that person and number are implied by a null morpheme, which Dyen does not) This is also the result of the lexical ambiguity of simple (i.e. unaffixed) words in Gaul, as discussed in the following section. Because of this, the whole complex nominal expression that formed the basis of the analysis of the more formal sentence falls in a heap.

3.2 Analysis Using Kridalaksana et al. (1985)

Figure 2: an excerpt of text from Swandy (2017), glossed using Kridalaksana et al. (1985)

So	*kalau*	*nanti*	*gue*	*udah*	*benar benar*	*lelah,*
So	COND	later	1SG	PERF	true ADV	be.tired*

So if I end up getting really tired of all this,

pasti	*gue*	*bakal*	*ny-erah*	*sendiri*	*kok.*
certainly	1SG	plan.to	SIMUL-defend**	myself	EMP

I'm bound to fight back.

In some ways, Kridalaksana's model works better than Dyen's. For one, Kridalaksana's provides an account of simulfixation (see **), the process whereby a verb is made explicitly active by nasalising its initial consonant (p.20). While Dyen's grammar could easily be reworked to include this by positing that the 'activiser' prefix *meN-* is elided but its presence implied by

the shift to nasalisation, this is not done by Dyen himself, which represents a glaring omission. Furthermore, since this phenomenon is extremely common in colloquial usage, the reference to simulfixation means that fewer null morphemes are required to adapt our morphological model of verbs in standard Indonesian to make it correctly parse Gaul verbs. On top of this more flexible understanding of syntax and morphology that allows for a better parsing of Gaul phrases, Kridalaksana provides a fuller account of the morphological processes underlying word formation in Gaul. Phenomena such as the use of reduplication to imply entire clauses (p. 23), the use of abbreviation and back formation (pp. 23-24), and the set of loan pronouns used most commonly by Gaul (p. 35; p. 72) are all discussed at length while they are absent from Dyen.

There are still issues with Kridalaksana's model, however. Chief among these is how it handles adverbs. In Kridalaksana's model, adverbs can refer to either adjectives or verbs, and can come anywhere in the sentence, not just next to the word which they modify (p. 61). Additionally, aspect markers (e.g. *udah, terus*) tend to be analysed as adverbs (pp. 70-71), meaning that a great deal of lexemes which actually possess a unique grammatical function are simply labelled 'adverbs', and the syntactic rules governing their placement not specified in adequate detail. This is compounded by the ambiguous ways in which Gaul identifies its lexical categories morphologically (or rather, how it often does not). For instance, in the text above, *lelah* could be parsed either as an adjective or a verb – each makes equal sense, and since the word is unaffixed, as is extremely common in Gaul, we are left with no morphological clues as to its lexical category. The best we can do is infer its lexical category through context. In the gloss above, *lelah* is analysed as a verb ('be tired', see *), and *benar-benar* as an adverb which modifies it, although since Kridalaksana allows adverbs to modify adjectives (p. 61) this is only one of two permissible analyses.

3.3 Analysis Using Sneddon et al. (1996)

Figure 3: an excerpt of text from Swandy (2017), glossed using Sneddon et al. (1996)

Terus	*banding-in*	*ke-salah-an*	*gue*	*sama*	*lo*
HAB	compare-Ø	NMLZ-wrong-NMLZ	Ø	with	Ø

And if you compare my mistakes to yours,

itu	*apa*	*se-besar*	*ke-salah-an*	*lo*	*punya...*
PROX.DEM	what	EQU-big	NMLZ-wrong-NMLZ	Ø	have

What have I ever done that's as bad as the mistakes you've made?

Sneddon's grammar is the most extensive and also the most prescriptive of the three surveyed. It is also, however, the only of the grammars not to prescribe generative rules for phrase structure. It is therefore difficult to evaluate the strength of the syntactic model it advocates for parsing entire sentences, but its analysis of the functions of individual words and morphemes can still be reviewed. Much like Dyen, Sneddon's analysis depends on the highly agglutinating morphology of standard Indonesian, and therefore fails to cope with the more isolating inclinations of Gaul. An example of the danger of relying on this complex morphology can be seen in the final noun phrase of the glossed excerpt, *kesalahan lo punya.* Although context clues

indicate that this is actually a noun phrase using the passive construction, the scant affixation gives us no way of deducing this by looking at the phrase in isolation. For this reason, there is no sensible way of interpreting this phrase with Sneddon's grammar. For one, Sneddon states that all transitive verbs must precede the objects to which they refer if they appear unaffixed (p. 277). Given that the root *punya* appears here completely unaffixed, with no morphological clues to either its passitivity or its transitivity, it would appear by Sneddon's analysis that the sentence is missing an object after *punya*. This is clearly erroneous, as the object is *kesalahan*. To make this noun phrase fit within Sneddon's model, we would either have to either posit that the passiviser is simply being elided, or that transitive verbs within noun phrases are not bound to the head. Neither of these concessions is particularly helpful, as both create additional ambiguity when identifying the structure of noun phrases.

Another, more minor flaw in Sneddon's analysis is his selective vocabulary. Take for instance the word *terus*. It is mentioned a number of times (pp. 218-219 inter alia), both as a standalone aspect marker and as the root of complex words, but always carrying the meaning of 'habituality'. While this meaning does exist in Gaul, its much more common function, and indeed its true function in the above text, is as a complementiser, carrying a secondary implication of causality. The word lagi, which has a very similar meaning, is subject to the same narrowing of meaning by Sneddon (pp. 211-212 inter alia). As another example, although my gloss records the informal second person singular pronoun lo for ease of reading, this pronoun is not present in Sneddon. This is indicative of a broader trend in Sneddon not to provide vocabulary which is incredibly common in Gaul, including words which have a specific grammatical function and would typically appear in a grammar on lists of pronouns, aspect markers, and such. Although Sneddon's grammar is extensive and well-equipped for analysis of written and formal language, it lacks the flexibility in both its syntax and vocabulary required to analyse Gaul. It is important to note that this is not so much a criticism of Sneddon's grammar when used to analyse standard Indonesian; it merely suggests that a model capable of handling formal standard Indonesian which does not make note of differences in register may be inadequate when dealing with Gaul. Sneddon himself has stated that contemporary Indonesian is "essentially diglossic", and that the grammar of Gaul should be viewed as separate from that of the standard language (Sneddon, 2003, pp. 17–18).

4 Conclusion

As stated in the introduction, the selection of grammars from across several decades was deliberate. The Indonesian language is subject to rapid change driven by a number of sociolinguistic factors. As such, grammars from as late as the 1960s suffer from a variety of inaccuracies which are the result of asynchronous language shift rather than poor research. However, as this article has demonstrated, the rules outlined in grammars of formal Indonesian struggle to cope with the grammar and discourse patterns of Gaul, irrespective of their date of publication. Furthermore, the tendency of supposedly 'descriptive' grammars of Indonesian to draw primarily from written sources, a behaviour which all three surveyed grammars admit to, has problematic implications for these grammars' ability to describe everyday spoken Indonesian. Whether or not the difference between standard Indonesian and Gaul truly does represent a

diglossia, as stated by Sneddon, is beyond the scope of this article, but what is clear is that the two are meaningfully distinct. Furthermore, since linguistic research into the grammar of Gaul is scant and typically conducted through the paradigm of written Indonesian, we are left with no alternative but to refer to ill-fitting formalistic grammars when analysing Gaul. Since, as discussed above, the prevalence and influence of Gaul is growing rapidly, this situation is undesirable both for linguists and for everyday speakers of Gaul alike.

References

Dyen, I. (1967). *A descriptive indonesian grammar*. Yale University. https://books.google.com.au/books?id=-q9pPAAACAAJ

Kridalaksana, H., Montolalu, L. R., Utorodewo, F., Elias, S., Sutami, H., Siswanto, M. H., Bala, S. E., & Djelalu, S. (1985). *Tata bahasa deskriptif bahasa indonesia: Sintaksis*. Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan. https://books.google.com.au/books?id=AWS5AAAAIAAJ

Kusuma, V. S., & Mardijono, J. J. (2013). The word formation processes of indonesian slang words used in cosmo girl! indonesian teenage magazine. *Kata Kita*, *1*(1), 217–222.

Smith-Hefner, N. J. (2007). Youth language, "gaul" sociability, and the new indonesian middle class. *Journal of Linguistic Anthropology*, *17*(2), 184–203. http://www.jstor.org/stable/43104150

Sneddon, J. (2003). *The indonesian language: Its history and role in modern society*. UNSW Press. https://books.google.com.au/books?id=A9UjLYD9jVEC

Sneddon, J. (2016). *Colloquial jakartan indonesian*. Canberra.

Sneddon, J., Adelaar, A., Djenar, D., & Ewing, M. (1996). *Indonesian reference grammar*. Allen & Unwin. https://books.google.com.au/books?id=dRhV%5C_KgfNo8C

Swandy, E. (2017). Bahasa gaul remaja dalam media sosial facebook. *Jurnal Bastra*, *1*(4).

Tadmor, U. (2007). Grammatical borrowing in indonesian. In Y. Matras (Ed.). J. Sakel (Ed.), *Grammatical borrowing in cross-linguistic perspective* (pp. 301–328). Mouton de Gruyter.