# A LEXICAL-FUNCTIONAL GRAMMAR REPRESENTATION OF INDONESIAN VERBAL SENTENCES

### **Fuad Cholisi**

#### Abstract

This paper presents a Lexical-Functional Grammar (LFG) description of Indonesian structures with a verbal predicate. The similarity of Indonesian and English in this type of construction has enabled the application of the original patterns of LFG for the English structures on its Indonesian counterparts. However, some adjustment has to be made in the description of the constituent structure (c-structure). The Indonesian constituent structure here is unique in that whilst it is organized endocentrically, it uses lexocentric means of function identification. Another different feature description that distinguishes the Indonesian LFG representation from the English one is the absence of the features for tense and agreement in the feature structure (f-structure) due to the fact that Indonesian structures are not subject to tense and number agreement. The number feature, however, appears in the cstructure merely to show the status of the subject in terms of singularity or plurality. In addition to the distinctive descriptions above, some constraints and thematic arguments based on the Lexical Mapping Theory have to be applied in the phrase structure rules to accommodate some peculiar characteristics of Indonesian verbal structures, such as those dealing with the position of adjuncts and oblique objects.

**Keywords:** Lexical-Functional Grammar (LFG), constituent structures (c-structure), feature structure (f-structure), phrase structure rule, lexical entry, sub-catagorization, predicate-argument structure.

Formal system of linguistic representation or grammar formalism, which is supposed to bridge the gap between linguistics and computer technology for language processing, has so far received very little attention among scientists and academicians in Indonesia. There is only very little interest in research in this subject area. One of such grammar formalisms that has gained a worldwide attention is Lexical Functional Grammar (henceforth LFG), which was introduced by Kaplan and Bresnan (1982) (see also Cholisi 2010). This formalism has developed so rapidly in recent years that the formalism representation has become extensively varied when compared to its original forms.

When discussing the grammar representation for a particular language, the standard assumption is that all syntactic operations should be expressed and represented in terms of constituent structure, categories and linear order relations, in

jsh Jurnal Sosial Humaniora, Vol 6 No.1, Juni 2013

which the c-structure can be organized in either endocentric mode or lexocentric mode (cf. Bresnan 2001, Dalrymple 2001, Falk 2001). In this research, the LFG descriptions, particularly the c-structures and their rules, are organized in the conventional way, i.e. endocentrically but with lexocentric means of function identification as those found in the early development of this formalism. As Bresnan (2001) suggests, this approach should be applied to languages with a NP < XP subject-predicate construction, where XP may be a predicate phrase of any range of categories VP, NP, AP or PP. Indonesian obviously belongs to this group of languages.

### **Canonical Indonesian Verbal Sentences**

Indonesian verbal sentences have exactly the same standard order as that of English sentences. The commonest structures of Indonesian verbal sentences are presented below:

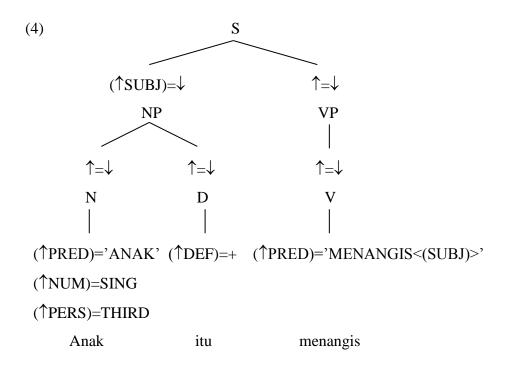
- Anak itu menangis child the cry 'The child cried.
- (2) Rani membeli sebuah buku.Rani buy a book'Rani bought a book.'

Whilst (1) is a typical structure of sentences with an intransitive-verb predicate, (2) represents that of sentences with a transitive-verb predicate. Both sentences are glossed into their past-tense English equivalents for reasons of convenience. Although sentences with the verbs expressing actions like (1) and (2) are likely to show the past time occurrence rather than the present, tense is not specified in Indonesian and so it is omitted from the f-structures of the LFG descriptions.

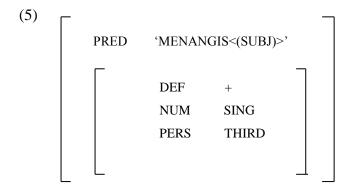
To account for (1), the set of standard phrase structure rules commonly used in LFG for English is appropriate, with the order of the daughters of NP switched round:

(3) a. 
$$S \rightarrow NP VP$$
  
 $(\uparrow SUBJ) = \downarrow \uparrow = \downarrow$   
b.  $NP \rightarrow N D$   
 $\uparrow = \downarrow \uparrow = \downarrow$   
c.  $VP \rightarrow V$   
 $\uparrow = \downarrow (\uparrow OBJ ) = \downarrow$ 

The first level of LFG representation of (1) can then be described in terms of its complete annotated c-structure as illustrated in (4).



There is one point to notice before we deal with the description of the fstructure. Indonesian verbs differ from English verbs in that they are not subject to tense and number. Consequently, the features for tense and number are not required in the description of the f-structure, i.e. there are no AGR features on verbs. From the c-structure (4), the f-structure for (1) can be built as seen in (5).



To see how the semantic arguments are mapped onto the f-structure, the lexical entry for *menangis* is given in (6).

(6) menangis: V, ( $\uparrow$  PRED) = 'MENANGIS <(SUBJ)>'

It should be noted from (5) and (6) that the value of the attribute PRED is the same as the form of the item represented in the lexical entry, which in this case is the derivative form rather than the base for reasons that will be explained later. Secondly, the lexical entry for *menangis* (6) does not contain any information about the tense or the form of the verb for the reason mentioned earlier in this section. Therefore, the feature TENSE and V-FORM, which are common in LFG representation of many other languages, do not appear in the f-structure (5). This explains the previous argument about the unnecessary assumption of the past action in relation with the interpretation of (1) and (2).

Another observable point in (5) is that in dealing with the subject feature (SUBJ), we have used the traditional version of LFG, i.e. by assigning separate attribute-value pairs of [NUM SING] and [PERS THIRD] rather than putting them together under the category-valued feature agreement (AGR), which is supposed to indicate the subject-verb agreement. This feature is not applicable in the f-structure, as there is no such grammatical phenomenon as agreement in number between a subject NP or any other NPs and the verb in Indonesian, although there are other circumstances where agreement between a noun and an indefinite article classifier

and between reflexive pronouns and their antecedents takes place. Following this argument, the lexical entry for *anak* is as simple as described below.

(7) anak: N, (
$$\uparrow$$
 PRED) = 'ANAK'  
( $\uparrow$  NUM) = SING  
( $\uparrow$  PERS) = THIRD

For sentence (2), however, we need to take account of the fact that the order of articles depends on their definiteness:

(8) a. NP 
$$\rightarrow$$
 D N  
( $\uparrow$ DEF) = -  $\uparrow = \downarrow$   
b. NP  $\rightarrow$  N D  
 $\uparrow = \downarrow$  ( $\uparrow$ DEF) = +

The constraints assigned to the determiners would rule out NPs in the wrong order. Applying (8.b), (2) can be straightforwardly generated, resulting in the f-structures as described in (9).

## Verbal Sentences with Two NPs

Having looked at the LFG descriptions of the simplest structures of Indonesian sentences such as (1) and (2), we now go on to deal with sentences with two NPs functioning as objects which are immediately dominated by VP as exemplified in (10). (10) Pria itu membelikan gadis itu sebuah cincin.

man the BEN-buy girl the a ring

'The man bought the girl a ring'.

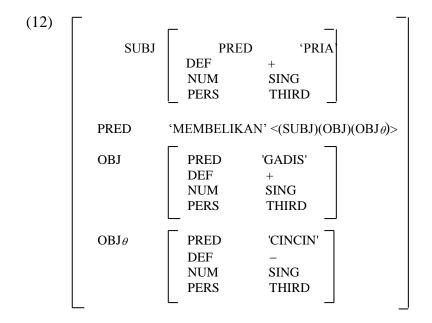
As is indicated by the gloss, *membelikan* is a benefactive verb that can be translated 'buy for'. It may be useful here to describe some characteristics of Indonesian verbs before we continue with the LFG account of (10).

Apart from their independence of tense and number, Indonesian verbs are much more complex than English verbs. They are so much characterized by affixation, and the meaning of a verb has much to do with the affixes attached. The affixation carries with it some semantic elements in the sense that the prefix and suffix attached to the verb determine the sub-categorization frame that it is assigned to, or more specifically, the arguments that it takes, functionally and semantically. The verb *membelikan* is derived from the base form *beli*, which without affixation cannot operate syntactically in the formal, standard Indonesian, although there is the case of independent verbs, which can operate syntactically in their base forms. The prefix *me*- attached to a verb implies that the verb is active rather than passive. In addition to *membelikan*, some other verbs can be derived from the base form *beli* such as membeli 'buy', dibeli 'be bought'(PASSIVE), terbeli 'can be bought/have been bought', and dibelikan (BENEFACTIVE PASSIVE). Although morphological details are ignored here, such characteristics explain why the derivative forms are preferred to the base forms as the lexical items. As far as LFG is concerned, it would be too complicated to develop lexical entries on the basis of the base form. Applying the base forms in the lexical entries will result in the obligation of building up nonoperational entries for the base forms that cannot operate syntactically without affixation and a conflict of double entries between these base forms and their derivative active forms, which semantically have the same meaning.

To generate the sentence in (10) we can apply the VP rule with the double dominated NPs offered by Kaplan and Bresnan (1982:185) to handle the similar construction in English, as is presented in (11).

(11) 
$$VP \rightarrow V$$
 [NP] [NP]  
 $\uparrow = \downarrow$  ( $\uparrow OBJ$ ) =  $\downarrow$  ( $\uparrow OBJ_{\theta}$ ) =  $\downarrow$ 

jsh Jurnal Sosial Humaniora, Vol 6 No. 1, Juni 2013



The f-structure resulting from applying (11) to (10) is illustrated in (12).

An alternative to (10) is a construction with a prepositional phrase following the object NP, which is similar to that of English as exemplified in (13).

(13) Pria itu membeli sebuah cincin untuk gadis itu.man the buy a ring for girl the 'The man bought a ring for the girl'.

However, as far as the meaning is concerned, the word "buy" has two equivalents in Indonesian. Whilst in English "buy" is a ditransitive verb, its meaning entails two Indonesian verbs of two different subcategorization frames, i.e., *membelikan*, which is a benefactive verb, and *membeli*, which is a monotransitive verb.

In terms of the Lexical Mapping Theory (see Bresnan and Karneva, 1989), the difference between the two verbs is captured by their different thematic and argument structures and their mappings to the relevant grammatical functions as presented in (14) and (15) respectively. (14) *membelikan* < agent beneficiary theme> [- 0] [- r] [+0]**SUBJ** OBJ  $OBJ_{\theta}$ (15) *membeli* < agent theme goal > [- 0] [-r] [- 0] **SUBJ** OBJ **OBL**<sub> $\theta$ </sub>

As it is seen from (14) and (15), the difference in the thematic and argument structures of the two verbs consequently results in different grammatical functions they subcategorize for. With *membelikan* (14), the second thematic role - beneficiary - is mapped to [- r] by intrinsic classification and to OBJ thanks to its unrestrictedness, whilst the third thematic role - theme - is mapped to [+ o] and is assigned to OBJ<sub> $\theta$ </sub>. With *membeli* (5), however, it is the theme that is mapped to [- r] and consequently to OBJ, whilst the third thematic role – goal – is mapped to OBL<sub> $\theta$ </sub> owing to its non-objective function.

There is apparently a difference in the grammatical function assigned to *gadis itu* in (10) and that in (13). The distinction between the grammatical functions assigned to this argument in the two constructions can be explained in terms of whether or not it is a nucleus/core argument or marked with a preposition (cf. Falk 2001, Mathews 1997, Trask 1993, Brown and Miller 1991). A non-core argument that is realized as an object to a preposition showing a path reference is an OBL*goal*. Thus in (13) *gadis itu* is an OBL*goal* because it is marked with such a preposition, i.e., *untuk* 'for' and it cannot be assigned to the OBJ of the core clause. In (10), however, *gadis itu* is obviously a core argument realized as an OBJ, and since this argument benefits from the action *membelikan* 'buy-for' rather than being directly part of it, it is definitely a beneficiary OBJ.

Under the mapping patterns of the two words as described above, the appropriate equivalent of "buy" to fit in (13) is *membeli* rather than *membelikan*. Such mapping patterns will obviously rule out the following constructions.

- (16) a. \* Pria itu membelikan sebuah cincin gadis itu.man the buy a ring girl the 'the man buy a ring the girl'
  - b. \* Pria itu membeli gadis itu sebuah cincin.man the buy girl the a ring'the man buy the girl a ring'

As for the mapping to the f-structure, again, the phrase structure rules offered by Kaplan and Bresnan (1982), cited below (17), fit well to account for (13).

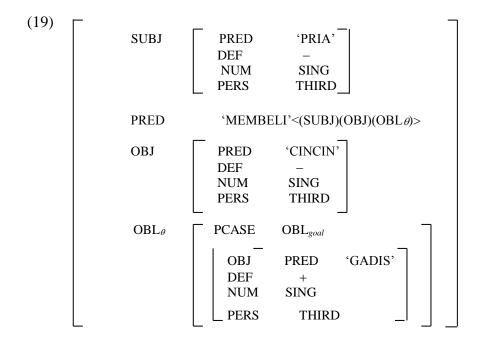
(17) a. 
$$VP \rightarrow V$$
  $\begin{bmatrix} NP \\ \uparrow = \downarrow & (\uparrow OBJ) = \downarrow & (\uparrow OBJ_{\theta}) = \downarrow & (\uparrow (\downarrow PCASE)) = \downarrow \\ b. PP \rightarrow P & NP \\ \uparrow = \downarrow & (\uparrow OBJ) = \downarrow \end{bmatrix}$ 

By applying these rules and mapping the lexical entry for *untuk* and *membeli* given in (18.a) and (18.b), the f-structure for (13) can be built as illustrated in (19).

(18) a. untuk: P, ( $\uparrow$  PCASE) = OBL<sub>goal</sub>

b. membeli: V, ( $\uparrow$  PRED) = 'MEMBELI' <(SUBJ)(OBJ)(OBL $_{\theta}$ )>

Following the suggestion of Kaplan and Bresnan (1982), the GADIS f-structure is accessible as an OBL, and it is correctly mapped onto the goal thematic argument of MEMBELI by the semantic form for *membeli* in (18), as well as in (19).



In Indonesian, however, a PP may occur before the subject as well as between the subject and the verb, as is shown by the alternatives of (20) below.

(20) a. Untuk gadis itu pria itu membeli sebuah cincin.

For girl the man the buy a ring 'For the girl, the man bought a ring.'

b. Pria itu untuk gadis itu membeli sebuah cincin.man the for girl the buy a ring'The man bought a ring for the girl.'

For reasons of convenience, in what follows we will deal with these constructions whilst at the same time accounting for the construction with adjuncts, as in many cases a PP element serving as an adjunct may take the same position as an oblique object.

### Verbal Constructions with Adjuncts

Our rules have so far proved to be adequate to handle the nuclear clause components such as NP, VP and PP. A sentence, however, may have other elements that do not serve as arguments to lexical predicates, such as adjuncts. To handle

jsh Jurnal Sosial Humaniora, Vol 6 No. 1, Juni 2013

sentences with adjuncts, the rule in (17.a) needs a further expansion. To simplify, we shall extend the sentence in (13) to (21).

(21) Pria itu membeli sebuah cincin untuk gadis itu pada hari Minggu.man the buy a ring for girl the on day Sunday'The man bought a ring for the girl on Sunday'.

This sentence (21) can be accounted for by expanding (17.a) to (22), which is exactly the rule offered by Kaplan and Bresnan (1982) to capture the similar construction in English.

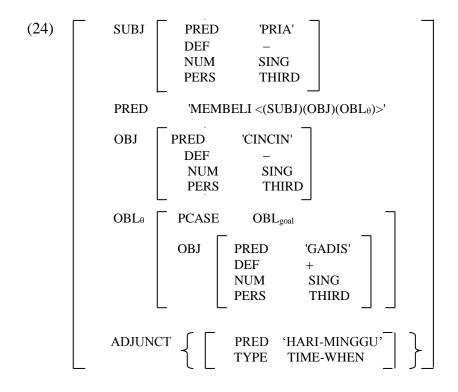
(22) 
$$VP \rightarrow V$$
  $\left( \begin{array}{c} NP \\ \uparrow = \downarrow \end{array} \right) \left( \begin{array}{c} OBJ \\ (\uparrow OBJ) = \downarrow \end{array} \right) \left( \begin{array}{c} OBJ \\ (\uparrow OBJ_{\theta}) = \downarrow \end{array} \right) \left( \begin{array}{c} (\uparrow (\downarrow PCASE)) = \downarrow \\ \downarrow \in (\uparrow ADJUNCTS \end{array} \right)$ 

The membership schema  $\downarrow \in (\uparrow ADJUNCTS)$  placed in the braces indicates the adjunct possibility as an alternative to the oblique-object PP and shows that the value of ADJUNCTS is a set containing the PP's f-structure as one of its elements. The optional schemata assigned to the PP with a Kleene star also permit the appearance of any number of both adjuncts and oblique objects in the same sentence in any order of occurrence. Thus, the positions of PP-adjuncts and PP-arguments can be arbitrarily switched round as shown in (23).

- (23) a. Tono memberi uang kepada Dodi di sekolah pada hari Senin Tono give money to Dodi at school on day Monday 'Tono gave some money to Dodi at school on Monday'
  - b. Tono memberi uang kepada Dodi pada hari Senin di sekolah
    Tono give money to Dodi on day Monday at school
    'Tono gave some money to Dodi at school on Monday'
  - c. Tono memberi uang di sekolah kepada Dodi pada hari Senin.
    Tono give money at school to Dodi on day Monday
    'Tono gave some money at school to Dodi on Monday'

d. Tono memberi uang di sekolah pada hari Senin kepada Dodi Tono give money at school on day Monday to Dodi 'Tono gave some money to Dodi at school on Monday'

Applying (22) to (21) results in the f-structure as is illustrated in (24) overleaf.



Following what Kaplan and Bresnan (1982) note, the element of the set value of the attribute ADJUNCT is indicated by the enclosing braces. The braces used here are distinct from those in c-structure rules that indicate alternative expansions. In (24), ADJUNCT has a value of a set with one member containing attribute-value pairs of [PRED HARI-MINGGU] and [TYPE TIME-WHEN].

However, as with the case of obliques, Indonesian adjuncts have more freedom in their position in the sense that they can take any position in the sentence but between a transitive verb and its object, and several adjuncts may occur together in different positions in a sentence. Thus, all sentences in (25) are equally wellformed.

- (25) a. Pada hari Minggu Lisa membaca koran di rumah.on day-Sunday Lisa read newspaper at home'On Sunday Lisa read a newspaper at home.'
  - b. Lisa pada hari Minggu membaca koran di rumah.
    Lisa on day Sunday read newspaper at home
    'Lisa read a newspaper at home on Sunday
- c. Lisa membaca koran di rumah pada hari Minggu.
  Lisa read newspaper at home on day Sunday
  'Lisa read a newspaper at home on Sunday.'

But (26) is ungrammatical.

(26) \*Lisa membaca pada hari Minggu koran di rumah.Lisa read on day Sunday newspaper at home'Lisa read a newspaper at home on Sunday.'

Coincidentally, in all of the cases above the adjunct takes exactly the position where an oblique PP can slot in. It is therefore necessary to add the constraint on the PS rule that the positions of the subject NP and PP are alternatives, and that the adjunct possibility is an alternative to the oblique PP, meaning that a sentence-initial PP can be an OBL. Secondly, the VP rule should allow more freedom for a PP to occur in the positions as described in (20.b). To cope with these suppositions, the PS rule in (3.a) should be redefined as (27).

$$\begin{array}{ccc} (27) \ S \ \rightarrow \ \left\{ \begin{array}{c} PP* \\ (\uparrow (\downarrow PCASE)) = \downarrow \end{array} \right\} & NP \\ \downarrow \in (\uparrow ADJUNCTS) \end{array} \begin{array}{c} NP \\ (\uparrow SUBJ) = \downarrow \ \left\{ \begin{array}{c} (\uparrow (\downarrow PCASE)) = \downarrow \\ \downarrow \in (\uparrow ADJUNCTS) \end{array} \right\} & \uparrow = \downarrow \\ \downarrow \in (\uparrow ADJUNCTS) \end{array}$$

Rule (27) and rule (22) will allow all sentences in (21) and (25) to be generated. Under these rules, any oblique objects as well as adjuncts occurring before the verb or at the final position can be well generated. Rule (27) also allows an adjunct to come in front of an oblique or the other way round, which is considered as grammatical despite its rare use in formal expressions.

### Conclusions

As far as the Indonesian verbal structures are concerned, apart from some lesser features such as plurality, tenses, and the definiteness and indefiniteness of articles, the syntax can be clearly described in LFG representations with only a few adjustments. Given the similarity between the structures of English simple sentences and those of verbal sentences in Indonesian, some formulations of the earlier version of LFG developed by Kaplan and Bresnan (1982) are still usable for such verbal sentences. In dealing with the SUBJ feature, for example, a concept of the earlier version of LFG, i.e., the assignment of separate numbers (NUM) and person (PERS) attribute-value pairs rather than putting them together under the category-valued feature agreement has been used due to the lack of subject-verb agreement in Indonesian. For a similar reason, the feature for tense is not required in the fstructure, as Indonesian verbs are not subject to tense.

In the account of some derivational verbs such as the benefactive verb *membelikan* 'buy for', which is derived from the verb *membeli* 'buy', the representation has benefited from the Lexical Mapping Theory in that the thematic and argument structures formulated by this theory have helped to capture the difference between the two verbs. Under the theory's intrinsic classification, such a difference is captured by mapping the thematic roles to the correct grammatical functions via the restricted and unrestricted and the objective and non-objective functions formulated therein.

Despite its typological SVO characteristic, Indonesian is freer than English in terms of the ordering rules. This is particularly true as far the position of obliques and adjuncts is concerned. Indonesian obliques and adjuncts can take any position in the clause except for that between a transitive verb and its object, and several adjuncts may occur together in different positions in a sentence. To cope with this ordering freedom of the two constituents, the S-structure rule normally used for the similar phenomenon in English has been slightly changed by introducing the optional PP annotated with equations indicating an alternative between an oblique and adjuncts in two positions, i.e., before the SUBJ-NP and between the SUBJ-NP and the VP.

### **Refferences:**

- Bresnan, J. (Ed) (1982a) *The Mental Representation of Grammatical Relations*. MIT, Massachusetts.
- Bresnan, J. (1982b) 'The Passive in Lexical Theory'. In *The Mental Representation* of *Grammatical Relations*, J. Bresnan, Ed. MIT, Massachusetts.
- Bresnan, J. et al. (1982c) 'Cross-serial Dependencies in Dutch'. *Linguistic Inquiry*, Vol 13.
- Bresnan, J., and J. Kanerva (1989) 'Locative Inversion in Chichewa: A Case Study of Factorization in Grammar'. *Linguistic Inquiry*, 20.
- Bresnan, J., and A. Zaenen (1990) 'Deep Unaccusativity in LFG'. In *Grammatical Relations*, K. Dziwirek et.al., Eds. Center for the Study of Language and Information.
- Brown, K., and J. Miller (1991) Syntax A Linguistic Introduction to Sentence Structure, London: Routledge.
- Cholisi, F. (2010) 'Lexical-Functional Grammar: A computationally Tractable Model of Grammar Representation', *Jurnal Sosial Humaniora*, Vol 3 No.1, Juni 2010.
- Cook, V.J., and M. Newson (1988) Choamsky's Universal Grammar An Introduction. Blackwell Publishers, Oxford.
- Dalrymple, M., R.M. Kaplan, J.T. Maxwell III, and A. Zaenen (Eds) (1995) *Formal Issues in Lexical-Functional Grammar*, The Center for The Study of Language and Information, Leland Stanford Junior University.
- Falk, Y.N. (1983) 'Constituency, Word Order, and Phrase Structure Rules'. In *Linguistic Analysis 11*, pp. 331 360
- Friedman, J. (1971) A Computer Model of Transformational Grammar. American Elsevier, New York.
- Grimshaw, J. (1990) Argument Structure. The MIT Press, Cambridge, Massachusetts
- Halvorsen, Per-Kristian (1983) 'Semantics for Lexical-Functional Grammar'. In *Linguistic Inquiry Vol. 14, No.4.*
- Kaplan, R, and J. Bresnan (1982) 'Lexical-Functional Grammar: A Formal System for Grammatical Representation'. In *The Mental Representation of Grammatical Relations*, J. Bresnan, Ed. MIT, Massachusetts.

- Kaplan, R. (1994) 'The Formal Architecture of Lexical-Functional Grammar'. In Formal Issues in Lexical-Functional Grammar, L. Dalrymple et.al., Eds. Center for The Study of Language and Information, Leland Stanford Junior University.
- Matthews, P.H. (1997) *The Concise Oxford Dictionary of Linguistics*, Oxford: Oxford University Press.
- Mintz, M.W. (1994) A Student's Grammar of Malay and Indonesian. Singapore: EPB Publishers.
- Neidle, C. (1994) 'Lexical-Functional Grammar'. In *Encyclopedia of Language and Linguistics*, Pergamon Press, New York.
- Purwo, B.K. (ed) (1994) Analisis Klausa, Pragmatik Wacana, Pengkomputeran Bahasa, PELLBA 7, Jakarta: Lembaga Bahasa Unika Atmajaya.
- Roberts, I. (1997) Comparative Syntax. London: Arnold.
- Rosen, S.T. (1990) Argument Structure and Complex Predicates. New York: Garland Publishers.
- Sadler, L. (1996) *New Developments in Lexical Functional Grammar*. University of Essex, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.57.1088
- Sneddon, J.N. (1996). Indonesian A Comprehensive Grammar. London: Routledge.
- Soenjono, D., Bambang KP., Harimurti K., Anton M.M., M. Ramelan, Samsuri, Sudaryanto, Mangasa S., D.P. Tampubolon., and Henry G. Tarigan (1993). *Tata* Bahasa Baku Bahasa Indonesia, Jakarta: Balai Pustaka. Trask, R.L. (1993). A Dictionary of Grammatical Terms in Linguistics. London: Routledge.
- Trask, R.L. (1993). A Dictionary of Grammatical Terms in Linguistics. London: Routledge.
- Zaenen, A., and J. Maling (1983). 'Passive and Oblique Case'. In *Papers in Lexical-Functional Grammar*, L. Levin, M. Rappaport, A. Zaenen (eds), 159-91.
   Bloomington, Indiana: Indiana University Linguistics Club.