A LINGUISTIC AND COMPUTATIONAL MORPHOSYNTACTIC ANALYSIS FOR THE APPLICATIVE -I IN INDONESIAN

I Wayan Arka♦, Mary Dalrymple♣, Meladel Mistica♦, Suriel Mofu♣,
Avery Andrews♦, Jane Simpson♥
♦ Australian National University, ♣ Oxford University, ♥ University of Sydney

Abstract

We present a precise LFG-based analysis of the suffix –*i* in Indonesian, addressing the issues of applicative-causative polysemy of the suffix and its alternation with –*kan*. We also show how the analysis can be integrated into an implementation of an existing computational grammar. Our computational implementation of applicativisation and related phenomena is the first of its kind for Indonesian, and also provides evidence for the robustness of LFG as a theory and XLE as a computational implementation of the theory in handling this linguistically complex phenomena.

Building on work on predicate composition (Alsina 1996; Butt 1995) and using the restriction operator (Kaplan and Wedekind 1993; Butt, King, and Maxwell III 2003; Butt and King 2006), we demonstrate that our novel unified a-structure based approach to verbal derivation in Indonesian can handle not only valence changing –*i*, giving rise to applicative-causative polysemy, but also valence-preserving –*i*. We argue that different types of –*i* (and also –*kan*) result from different possibilities of argument fusion: double or single fusion. Double fusion, which is typical for -*i*, results in applicativisation, whereas single fusion of –*i* results in causativisation.

1 Introduction

Indonesian is one of the most extensively studied Austronesian languages (Chung 1976; Myhill 1988; Purwo 1989, 1995; Macdonald 2001; Musgrave 2001, among others), yet the precise linguistic analysis of the suffix –*i* has not been investigated in detail. Linguistically, the analysis must address the issues of applicative-causative polysemy or homonymy of –*i*,[1] and, from a computational point of view, we must address how the analysis can be integrated into an implementation of an existing computational grammar. The suffix –*i*, like the suffix –*kan* (Arka 1993), can appear as applicative or causative, as seen in (1). There are other uses of –*i* for which the proper analysis as applicative or causative is not clear (as we discuss in section 2, examples (5)-(6)).

(1)   a. Applicative -*i*

     i. X_{SUBJ} *datang*   [ke Y]$_{OBL}$ 'X come to Y'   (intransitive)

     ii. X_{SUBJ} *datang-i* Y_{OBJ} 'X come-APPL Y'   (transitive)

   b. Causative –*i*:

     i.   X   *panas*   'X is hot'   (intransitive)

     ii.   Y   *panas-i*   X   'Y heat up X'   (transitive)

Traditional grammars of Indonesian (Moeliono and Dardjowidjojo 1988; Sneddon 1996, among others) typically simply list the uses of –*i* without explicit argumentation as to whether there is one (polysemous) –*i*, or more than one (homonymous) –*i*. Our proposal addresses this issue; we claim that Indonesian –*i* is polysemous, as further explicated in section 3.

The discussion of morphological applicativisation and causativisation in the literature has typically focused on clear cases with verbal/adjectival stems, i.e. stems like *datang* or *panas* as in (1). Such stems are argument-taking predicates, and causativisation/applicativisation can be clearly identified by checking argument alternations of the stems. However, the same causative/applicative affix may also take stems of other categories, and in those cases the analysis of the form as causative or applicative is not straightforward. The suffix –*i* can be productively used with a noun stem, e.g. *kantong-i* 'pocket-*i*=put X in (own) pocket' and *garam-i* 'salt-*i* = put salt in X'. In these cases the noun stem is not understood as an argument-taking predicate, but rather as an argument (a location or theme) at some underlying semantic level. Again, this kind of derivation is often only mentioned in passing, and not given a precise analysis. Its significance is often overlooked as part of a wider family of transitivising processes that include applicativisation and causativisation. In section 3, we propose an argument-structure based analysis of –*i* which can be easily extended to account for –*i* verbs formed with noun stems.

---

[1] Similar cases of applicative/causative homophony are found in other languages, for example Australian Aboriginal languages (Austin 2005 [1996]) .

From a computational point of view, we need to capture the syntactic, semantic and aspectual characteristics of –*i* so that our computational grammar can produce correct parses of sentences headed by verbs with -*i*.

This paper proposes a novel unified a-structure based approach to verbal derivation, which can handle not only valence changing –*i*, giving rise to applicative-causative polysemy, but also valence-preserving –*i* (discussed later in this paper). We analyse -*i* as carrying its own PRED(ICATE) argument structure. Word-formation with –*i* involves predicate composition of the PRED of the suffix with the PRED of its stem, similar to complex predicate formation as described in Alsina (1996) and Butt (1995). We adopt the LFG-based predicate composition approach of complex predicate formation (Alsina 1996, Butt 1995), and extend it to handle Indonesian data. The implementation makes use of the restriction operator (Kaplan and Wedekind 1993; Butt, King, and Maxwell 2003; Butt and King 2006). The implementation of the analysis is discussed in section 4.

Computational morphosyntactic treatments addressing the issues of applicativisation and related phenomena have not been previously proposed for Indonesian. Applicativisation (and its complex interaction with other kinds of word-formation such as voice selection and reduplication) has not been previously implemented in XLE either. The implementation of the LFG analysis of –*i* in XLE therefore provides the first evidence for the robustness of XLE and LFG in handling this linguistically complex phenomena.

2   Indonesian verbal morphology and properties of -i

The Indonesian suffix–*i*, like –*kan*, has traditionally been described as a valence-increasing morpheme. The suffix –*i* has been given less attention than its –*kan* counterpart, as it is often regarded as 'simpler' than –*kan* (Vamarasi 1999). However, as we shall show, its behaviour is equally complex, as we see in the following sections, which outline Indonesian verbal morphology and describe the basic properties of –*i*.

2.1   Verbal template in Indonesian

Indonesian verbs can be morphologically simple or complex. The verbal template in Indonesian consists of a root, possibly with one or more affixes: (Prefix*-)Root(-Suffix*). The outermost prefix is a voice-related prefix and the outermost suffix is typically a transitiviser suffix –*kan/-i*. Between the outermost affix and the root, there may be another affix, e.g., the causative prefix *per-* as in *memperlihatkan* '*show (<(lit.) 'cause X to be seen')*, or a loan suffix, e.g. –*isasi* as in *memfungsionalisasikan* '*functionalize'*. Reduplication may also add complexity.

The verbal root can be free or bound. A free root such as *datang* 'come' and *pergi* 'go' can appear in its affixless form in syntax. A bound root, however, must be affixed to appear in syntax. The bound root often has a vague meaning and no clear grammatical category in isolation, which has led to the claim that it is 'precategorial' (Verhaar 1984). The root only gets a specific meaning and specific grammatical category when it is affixed; e.g., -*alih*- 'change position or course' → *mengalih* (V) 'change to (a different position, topic, etc.), *alihkan* (V) 'distract, shift', *peralihan* (N) 'transfer, transition', and *pengalihan* (N) 'diversion'.

2.2   Basic properties of *-i*

The suffix –*i* is a derivational suffix with the following properties. Firstly, -*i* can be affixed to stems of different categories. The following table shows that –*i* can appear with a noun, an adjective or a verb. The verb stem can be intransitive or transitive.

Roots	*Derived -i verbs*	*Roots*	*Derived -i verbs*
air (N) 'water'	*air-i* 'water'	*lompat* 'jump' (V_{ITR})	*lompat-i* 'jump over'
kulit (N) 'skin'	*kulit-i* 'peel'	*tidur* 'sleep' (V_{ITR})	*tidur-i* 'sleep on'
gula (N) 'sugar'	*gula-i* 'put sugar in'	*diam* 'stay' (V_{ITR})	*diam-i* 'dwell in'
ketua '(N) chair (of an organization)	*ketua-i* 'chair or lead in a meeting/organization'	*tulis* 'write' (V_{TR})	*tulis-i* 'write on something.'

panas (A) 'hot'	*panas-I* 'heat (water)'	*kirim* 'send' (V$_{TR}$)	*kirim-i* 'send' (ditrns)
basah (A) 'wet'	*basah-i* 'dampen'	*siram* 'spray' (V$_{TR}$)	*siram-i* 'spray with'
lengkap (A) 'complete'	*lengkap-i* 'complete'	*cium* 'kiss' (V$_{TR}$)	*cium-i* 'kiss repeatedly'
jauh 'far' (A)	*jauh-i* 'make oneself far from'	*pegang* 'hold' (V$_{TR}$)	*pegang-i* 'hold tightly'

Secondly, affixation with *-i* may or may not result in a change of valence of the stem. The examples in (2)-(4) show that *-i* can increase the syntactic valence of the stem by promoting an oblique to Object:

(2) a. *Ia   duduk   di   kursi   itu*   (loc) (Vitr→Vtr)
        3s   sit   LOC   chair   that
        'S/he sat on the chair'

     b. *Ia   menduduk-i   kursi   itu*
        3s   AV.sit-i   chair   that
        'S/he was sitting on the chair'

(3) a. *Ia   melempar   batu   ke   saya*   (go) (Vtr→Vtr)
        3s   AV-throw   stone   to   1s
        'S/he threw stones at me.'

     a. *Ia   melempar-i   saya   dengan   batu.*
        3s   AV.throw-i   1s   with   stone
        'S/he pelted me with stones.'

(4) a. *Ayah   mengirim   uang   kepada   {dia|=nya}* (go) (Vtr → Vdtr)
        father   AV.send   money   to   3s =3s
        'Father sent money to him/her.'

     b. *Ayah   mengirim-i   {dia|=nya}   uang*
        father   AV.send-i   3s =3s   money
        'Father sent her/him money.'

The examples in (5) and (6), however, show no valence increase with *–i*. Rather *-i* merely adds some aspectual meaning (repetition, intensity).

(5) a. *Ia   memukul   saya*
        3s   AV.hit   1s
        'S/he hit me'

     b. *Ia   memukul-i   saya*
        3s   AV.hit-i   1s
        'S/he was hitting me'

(6) a. *Ia   memegang   pencuri   itu.*
        3s   AV.hold   thief   that
        'S/he held the thief.'

     b. *Ia   memegang-i   pencuri   itu.*
        3s   AV.held-i   thief   that
        'S/he was holding the thief tightly.'

It is useful to compare the syntax and semantics of *–i* to its *–kan* counterpart. As noted by Kaswanti (1995), Sneddon (1996) and Kroeger (2007), *-i* alternates with *-kan* to provide different possibilities for object linking, similar to the spray-load alternation in English (Levin 1993). Verbs affixed with *-i* have a locative/goal object, whereas those with *–kan* have what Kroger (2007) calls a displaced theme/patient object:

(7) a. *Buruh   itu   memuat-kan   beras   ke   kapal.*
worker   that   AV.load-*kan*   rice   to   ship
'The workers loaded the rice onto the ship.'

b. *Buruh   itu   memuat-i   kapal   dengan beras.*
worker   that   AV.load-*i*   ship   with   rice
'The workers loaded the ship with rice.'

(8) a. *Anak-anak   menempel-kan   poster   ke   tembok.*
child-REDUP   AV.stick-kan   poster   to   wall
'The children stuck a picture to the wall.'

b. *Anak-anak   menempeli tembok dengan   poster*
child-REDUP AV.stick-*i* wall   with   poster
'The children stuck pictures all over the wall.'

(9) a. *Ia   memuntah-kan   darah   segar.*
3s   AV.vomit-kan   blood   fresh
'S/he vomited fresh blood.'

b. *Ia   memuntah-i   baju-nya.*
3s   AV.vomit-*i*   shirt-3s
'S/he vomited on his/her shirt'

The following examples also show the different effects of –*i* and –*kan*:

(10) a. *Polisi   datang*
police   come
'The police arrived/came.'

b. *Mereka   men-datang-kan   polisi*
3p   AV-come-kan   police
'They arrived with the police'   (comitative-applicative -*kan*)
'They called for/made the police come.'   (causative -*kan*)

c. *Mereka   mendatang-i   polisi*   (applicative -*i*)
3p   AV.come-i   police
'They came to/approached the police'

(11) a. *Dia   men-jatuh-kan   buku   saya.*   (causative -*kan*)
3s   AV-fall-*kan*   book   my
'S/he dropped my book.' (Lit: S/he made my book fall)

b. *Dia   men-jatuh-i   buku   saya.*   (applicative -*i*)
he   AV-fall-*i*   book   my
'S/he fell on my book.'

In these examples, the (causative) locative –*i* alternates with the causative displaced-theme -*kan*.

(12) a. *Buku-nya   menumpuk   /   bertumpuk*
book-3s   AV.pile.up   *ber*.pile.up
'His/her books piled up.'

b. *Mereka menumpuk-kan   buku   di   meja.*   (causative –*kan*)
3p   AV.pile.up-kan   book   on   table
'They piled up books on the table.'

c. *Mereka menumpuk-i   meja   itu   dengan   buku.*   (causative –*i*)
3p   AV.pile.up-i   table   that   with   book
'They were piling up books on the table.'

(13)   (from (Kroeger 2007); adapted from Dardjowidjojo 1971)

      a.   *Air   itu   sedang   meng-alir   ke   sawah.*
          water   that   PROG   AV.flow   to   rice.field
          'The water is flowing to the rice field.'

      b.   *Dia   meng-alir-kan   air   itu   ke   sawah=nya.*
          3s   AV.flow-*kan*   water   that   to   rice.field=3sg
          'S/he caused the water to flow to his/her rice field.'

      c.   *Dia   meng-alir-i   sawah=nya   dengan   air   itu.*
          3s   AV.flow-*i*   rice.field=3sg   with   water   that
          'S/he flooded his/her rice field with the water.'

The following shows that causative-permissive *–i* alternates with benefactive *-kan*. That is, *-i* encodes the object as borrower with the A/subject understood as the source/owner of the thing borrowed. With *–kan*, the A is the borrower and s/he borrows it from someone else for the benefit of the U *saya*.

(14)   a. *Ia   meminjam-i   saya   uang.*
         3s   AV.borrow-i   1s   money
         'S/he lent me money.'

      b. *Ia   meminjam-kan   saya   uang.* (*saya*=benefactive; source is not the actor)
         3s   AV.borrow-kan   1s   money
         'S/he borrowed money for me.'

When the root is a noun, the root is often understood as a '(displaced) theme' associated with the location designated by the object (in a real or metaphorical sense). The location can be understood as part of a 'static' situation as in (15a), or the goal or source of action as in (15b-d). The stem itself can be also understood as the location as in (15e).

(15)   a. *Sungai   ini   membatas-i   Malang   dan   Lumajang*   (root = theme)
         river   this   AV.border-i   Malang   and   Lumajang
         'This river becomes the border of Malang and Lumajang.'

      b. *Mereka   mengair-i   sawah-nya.*
         3p   AV.water-i   rice.field-3s
         'They were watering their rice-field.'

      c. *Dia   mengulit-i   pisang   itu.*
         3s   AV.skin-i   banana   that
         'S/he peeled the bananas (Lit. removed the skin from the bananas)

      d. *Dia   mengulit-i   buku   itu.*
         3s   AV.skin-i   book   that
         'S/he added a cover to the book.' (adding cover to the book)

      e. *Pihak   China   hendak   memenjara-i   Hồ Chí Minh …*
         side   Chinese   want   AV.prison-i   Hồ Chí Minh
         'The Chinese wanted to imprison Hồ Chí Minh.'

Though *–i* may show both applicative and causative functions, some roots do not allow both functions. The following patterns are observed. Firstly, there are roots which allow only the applicative function, as exemplified with *datang* 'come' → *datangi* (10c) and *jatuh* 'fall' → *jatuhi* (10b). The verb *datang-i* can only mean 'come to X[loc]', not *'make X come'. To derive a causative meaning with these verbs, the suffix *–kan* is used, *datang-kan* 'make X come' (10b), *jatuhkan* (11a). Secondly, denominal *–i* verbs with noun roots conceptualised as displaced themes typically have only an applicative function. The Object is understood as a location; e.g., *kutu-i* 'louse-i= delouse X', *gula-i* 'put sugar in X', etc. The *–kan* form is typically not attested; e.g., **kutu-kan,*gula-kan*. Finally, there are roots that allow both applicative and causative functions, such as the root *takut* 'afraid', as in (16a)

for the applicative *–i* reading and the causative *–i* in (16b). Unlike the second pattern, *-kan* is also often commonly used for this third type of root.

(16) a. *Sebut   1   hewan   yang   kamu   takut-i*!
         mention   one   animal   REL   2   afraid-i
         'Name one animal that you are scared of'

     b. *jikalau   tiada   yang   menakuti   mereka*[2]
         if   NEG   REL   AV.afraid-i   3p
         'if there is no one/thing that makes them afraid'

To sum up, the suffix *–i* can take roots of different categories with applicative and/or causative functions. The analysis of the precise nature of these functions is outlined in the next section.

3   Analysis

3.1   Patterns of alternations

The following properties must be captured by the analysis of *–i*:

(17) a. Syntax
         i.   valence increasing:
                      intransitive → transitive (examples (2), (10)),
                      monotransitive → ditransitive (example (4)).
         ii.   no valence changing effect (examples (5)-(6))

     b. Semantics
         i.   Locative applicative and causative
         ii.   Iterative/intensifying/progressive

     c. Related to (a) and (b), alternations with *–kan*.

The first question is whether we have one *–i* or more than one *–i*. We argue that we have one *–i*, and that the different properties as summarized in (17) are predictable from the interaction of the core information carried by *–i* and the information carried by the stem. We begin with a characterisation of the facts in theoretical terms, appealing to a Jackendoff-style semantic structure (Jackendoff 1990) to express our analysis. This forms the theoretical underpinning of our computational analysis, although our implementation is different and simpler than the analysis we present here in that it does not appeal to a separate level of argument structure.

Central to the semantics of *–i* is its locative meaning component in the state of affairs (SOA) it encodes. This can be informally represented as in (18).[3]

(18)   A AFFECT U$_i$ ({TO|FROM}) BE.AT([LOC]$_i$)

AFFECT (henceforth, AFF; see Jackendoff 1990) is a general semantic primitive which is intended to cover different degrees of affectedness associated with causativisation and applicativisation. The exact interpretation of AFFECT is determined by the semantics of the stem.[4]

The locative meaning which is always added by *-i* can in certain cases be thematically interpreted as part of PATH (i.e. TO or FROM in (18)). Thus it can sometimes be interpreted as 'goal' or 'source', depending on the meaning of the stem and world knowledge. For example, *menguliti* (< 'AV.skin.APPL') can mean 'remove the skin FROM' for an object understood to have skin to begin

---

[2] http://sabda.org/sabdaweb/bible/verse/?b=19&c=53&v=5&version=bis
[3] Kroeger (2007) analyses *–i* as having the following LCS (Lexical Conceptual Structure):
Locative object (V-i): [x ACT] CAUSE [z BECOME [FULL?]$_{STATE}$ WITH-RESPECT-TO y] [LOAD manner]$_{MANNER}$]
While this is the right LCS representation for many *–i* verbs, it does not capture the broad range of the semantics of all *–i* verbs: in particular, those where the object is understood as a location (i.e. source) from which something (i.e. y) is taken away, as in e.g. example (15c).
[4] For psychological verbs, REACT instead of AFFECT is used (Jackendoff 1990).

with (e.g. a crocodile). Alternatively, it can mean 'to put skin ON': e.g. for 'books' because books are created with skin (i.e. cover) added in the production process; see 3.4.4.

The aspectual meaning — iterative, progressive, intensifying; see examples (5)-(6) — is arguably also related to the locative meaning of *-i*. That is, the locative U is conceptually understood as having a spatial surface to which the action is applied. Affectedness applied to an unbounded space leads to a repetitive or progressive meaning, e.g. *mengecat* X 'paint X' → *mengecati* X 'paint all over the surface of X'. Other cases in which the same marking is used for locative alternation and aspectual distinctions such as telicity have been noted in the literature (Levin and Rappaport Hovav 1998).

The representation of the semantics of *–i* as shown in (18) is sufficient for the purposes of this paper. The locative and aspectual meaning of the *-i* verb, which is tied to OBJ, will be simply represented by its realisation as OBJ linked to a locative role.

3.2 *–i* as a head PRED verbalizer

In line with work on causativisation and applicativisation (Alsina and Joshi 1991; Alsina 1992; Butt 1995; Austin 2005 [1996]), we propose an a-structure-based analysis with the following key points.

First, we claim that there is one *–i* (i.e., a polysemous suffix). We analyse *–i* as a three-place predicate with its own argument structure, as shown in (19). Affixation with *–i* involves complex predicate composition, with argument fusion of the matrix and embedded arguments. Central to the analysis of *–i* is that the matrix's second argument (ARG2) is thematically a locative (LOC)-related argument (i.e., possibly Goal or Source, in addition to Locative). ARG2 fuses with the LOC argument of the base wherever possible. ARG1 is thematically higher than ARG2, though not necessarily an agent. This representation, as we shall see, allow us to capture both causative and applicative uses of *-i*, as well as other uses.

(19)   A-structure of *–i* and the associated semantic roles
         'PRED$_1$ < ARG$_1$, ARG$_2$,   PRED$_2$ < _ , ...>>'
              (A)   (U:$_{LOC}$)   where argument(s) of PRED$_1$
                                             fuse(s) with arguments of PRED$_2$

The second key point of our analysis is underspecified fusion. While the ARG2 of *–i* is thematically specified as LOC-related, the overall fusion is underspecified, constrained by the semantics of the root, possibly with lexicalisation for certain verbs. For example, the verb *baui* (<*bau* 'odour'-*i*' (lit.) cause the odour of X to be in Y') is lexicalised to mean 'Y smells X'.

We argue that this underspecified a-structure allows two different types of argument fusion, single or double, which then gives rise to different applicative and/or causative effects. We propose a general rule for *–i* composition: arguments of thematically similar types tend to fuse. Thus, the Actor-like ARG1 of the matrix PRED tends to fuse with the actor-like ARG1 of the embedded PRED. Likewise, the Undergoer-like ARG2 of the matrix PRED fuses with the Undergoer-like ARG2 of the embedded PRED. However, since PRED2 may be a one-place predicate, ARG2 of the matrix PRED may fuse with the sole argument (i.e. ARG1) of the embedded PRED. Different possibilities are further discussed and illustrated in 3.4 below.

The third key point is that the derived a-structure is constrained by a set of a-structure well-formedness properties: core arguments outrank non-core arguments, and within these groups arguments are ordered thematically (Manning 1996; Arka 2003). This constraint determines the derived syntactic transitivity, and also possible *–i* and *–kan* alternations, as we show in the next subsection.

Due to space limitations, we cannot outline the full details of the linking mechanism in this paper. We adopt a version of a-structure-based linking as described in Arka (2003:148-158), which is applicable to Indonesian (Arka and Manning 2008). An argument in the a-structure is represented as ARG or simply a "_" within angle brackets. If necessary, core status is indicated by nested bracketed groupings, with core arguments on the left group, and associated thematic roles placed within brackets underneath; see (21a) below.[5] Voice alternations may or may not alter argument structure. Actor Voice (AV) maps the most prominent core argument (ARG1 for *–i*) to SUBJ and the second core

---

[5] Thematic roles are shorthand labels associated with positions with prominence in the Lexical Conceptual Structure.

argument (ARG2) to OBJ whereas the Undergoer Voice (UV) has a reverse mapping. Passive Voice alters the a-structure associated with the Actor. This is further discussed in 4.1.

3.3 Transitiviser: -i and –kan alternation

Our proposal that *–i* introduces two core arguments (ARG1, ARG2) as shown in (19) is fully consistent with the traditional analysis of *–i* as a transitiviser. Moreover, given our theory of a-structure (core > non-core; matrix > embedded; thematically-based ordering within core/non-core), we also predict the valency of the resulting structure: this is determined by the thematic role of the argument of the embedded predicate that is not fused with a matrix argument (ARG1 or ARG2). The chart in (20) gives the subcategorisation frames resulting from combining *–i* and *–kan* with predicates with different inventories of thematic roles, where one argument of the predicate remains unfused. The unfused argument is labelled ARG3 in the linking configuration of *–i* and *–kan* in (20). It is the least prominent argument in the derived structure, and it can be either core (OBJ2) (20a) or non-core (OBL) (20b).

(20) Subcategorisation frames and associated thematic roles for verbs with *–i* and *–kan*

(a)

Ditransitive		
NP$_{SUBJ}$	NP$_{OBJ}$	NP$_{OBJ2}$
ARG1	ARG2	ARG3
-i: agt	go/src/loc	th
-kan: agt	ben/rec	th

(b)

Monotransitive		
NP$_{SUBJ}$	NP$_{OBJ}$	PP$_{OBL}$
ARG1	ARG2	ARG3
-i: ag	go/src/loc	instr
-kan: ag	pt/th	go/src/loc

The patterns in (20) are expected because the two matrix arguments, ARG1 and ARG2, are linked to the two arguments that are highest on the grammatical function hierarchy, namely SUBJ and OBJ respectively. The actor-like argument (e.g. Agent) is linked to ARG1/SUBJ by default. Each of the two suffixes *–i* and *–kan* locks the linking of ARG2 to a certain role: the suffix *–i* links ARG2 to a locative-related role (including goal/source), whereas *–kan* links it to beneficiary/recipient or theme.

The unfused argument (ARG3) is therefore the third in the list. When ARG3 is a displaced entity (i.e. a Theme) as in (20a), it is realised as OBJ2 in both *–i* and *–kan* verbs. A theme is ranked low in the thematic hierarchy (Bresnan and Kanerva 1989; Bresnan and Moshi 1990; Simpson 1991), and fits well as the least prominent core argument among the three core arguments that make up a ditransitive structure.

However, when the unfused argument (ARG3) is not a theme, it is realised as an oblique and the structure is monotransitive, as in (20b). Non-theme arguments (instrument, goal, source, locative) cannot naturally appear as (second) objects (i.e. third least-prominent core arguments) in Indonesian. This constraint can be taken as reflecting the cross-linguistically common generalization that a locative role is intrinsically classified as a non-objective function (Bresnan and Kanerva 1989). The instrument role can be analysed as a causer-like argument which is also classified as a non-objective function.

The proposed analysis, capturing similarities and differences between *–i* and *–kan* as depicted in (20), accounts for the fact that *–kan* and *–i* can alternate with certain stems. A key aspect of the analysis is the idea that *–i* and *–kan* are both matrix predicates with their own a-structure, including matrix ARG1 and ARG2 (cf. (19)), and that they lock the linking of ARG2. The suffix *-i* links ARG2 to a Goal/Loc-related role, whereas *–kan* links it to a Benefactive/Recipient or theme. The resulting (ditransitive or monotransitive) structures follow from independent principles. For instance, *kirim* 'send' is a monotransitive verb stem with the a-structure shown in (21a). This is precisely what *–kan* also specifies; i.e. ARG2 is a theme (see (20b)). Hence, it is not surprising that *–kan* is optionally present in this type of structure and does not change the basic argument linking, indicated by *–kan* in brackets in (21b).

(21) a. kirim 'send' <<ARG1, ARG2> <ARG3>>
                     (ag)   (th)   (go)

b. *Ayah   mengirim(-kan)   uang   kepada   saya*   *(monotransitive)*
   father   AV.send(-kan)   money   to   1s
   'Father sent money to me.'

Since *kirim* 'send' also has a displaced theme and a Goal argument, it also allows the ditransitive structure with *–i*, as in (20a). This is shown in (22):

(22) a. *Ayah   mengirim-i   saya   uang*   (ditransitive)
        father   AV.send-i   1s   money
        'Father sent me money.'

      b. *Ada   dua   pihak   yang   rutin   mengirim-i   saya   SMS*
         exist two   side   REL   regular   AV.send-i   1s   short.message
         'There are two groups of people who regularly send me short messages.'

Note that in this case the Goal is a human being, *saya*. It is therefore also interpreted as a Beneficiary/Recipient. For this reason, the ditransitive beneficiary structure (i.e., that of *–kan*, (20) a) is also expected to be available with this verb. This is correct, and in fact attested with high frequency on the Internet; a naturally-occurring example is shown in (23)[6].

(23)   *Jika   anda   bermaksud   mengirim-kan   saya   satu   pesan,*
       if   2   intend   AV.send-kan   1s   one   message

       Klik   "Contact"   di   atas
       click   contact   at   above
       'If you want to send me a message, click "Contact" above'

The alternation of *–i* and *–kan* (*mengirimi* vs. *mengirimkan*) in some ditransitive structures exhibits aspectual differences. The suffix *–i* signals progressive, iterative, or plural events. Thus, *mengirimi* in (22) implies that the act of sending took place on numerous occasions. In (22b), it is also signalled by the adverbial *rutin* 'regularly'. In contrast, *–kan* is typically associated with one-off events. In (23), the object/ARG2 is explicitly coded with *satu* 'one'. Replacing *mengirimkan* with *mengirimi* in this example would mean sending the same message over and over again to the same recipient.

3.4   Types of fusion

In our analysis, different results of combining predicates with *–i* (and also *–kan*) result from different possibilities of argument fusion: double or single fusion.

A derivation with *–i* typically involves double fusion. That is, both matrix arguments (ARG1 and ARG2) are fused with arguments of the embedded predicate. Double fusion may result in applicativisation, in which an embedded non-core argument is promoted to core status due to its fusion with ARG2, or simply additional meaning without valence-changing effect.

In single fusion, only one of the matrix arguments is fused with an argument of the embedded predicate. Single fusion typically results in causativisation.

In the following, we discuss and exemplify different types of derived *–i* structures.

3.4.1   Type 1

Type 1 involves derived monotransitive *–i* verbs undergoing a valence-changing applicativisation effect. With a two-place intransitive base (with Goal/Locative second argument) such as *jatuh* 'fell (on)to X', *datang* 'come to X', and *lewat* 'pass at X', the result is a strictly monotransitive *–i* verb.[7] This is exemplified by (24a-b), which can be represented as (24c).

---

[6] This attested example is from http://dart1202.wordpress.com/2009/04/07/review-bmw-t-shirt-bmw-city-pants, accessed 27 November, 2009.

[7] There is evidence that Goal/Locative of *jatuh* 'fall' or *datang* 'come' is an oblique-like argument (i.e., associated with the conceptual unit of [PATH] of the verbs) although it is not required to be overtly present on the surface syntax. A (general) Goal/Locative adjunct cannot typically take -i in Indonesian:

(24) a. *Mangga   yang   besar   jatuh   ke   rumah-nya*
          mango   REL   big   fall   to   house-3s
          'A big mango fell onto his house.'

     b. *Mangga   yang   besar   men-jatuh-i   rumah-nya*   (*menjatuhkan*)
          mango   REL   big   AV-fall-i house-3s
          'A big mango fell onto his house.'

     c.   'mango' 'house'
                 SUBJ   OBJ
          –i   <ARG1,   ARG2   'jatuh < _ (_) >'
                         (U:loc)

The same promotion effect is also observed for certain two-place bound verbal roots, e.g. *–kunjung*-
'visit (to) X' and *–saing*- 'compete (with) X'.[8] These roots are 'bound' because they must be affixed in
order to appear as verbs in a clause. The second argument (ARG2) is realised either as an OBL (non-
core) or OBJ (core) depending on the affix. For example, affixed with the intransitive *ber*-, the verb
berkunjung comes with its Goal argument as an OBL marked by *ke* (25a). Affixed with *–i*, the Goal is
OBJ in (25b), exhibiting the applicativisation effect. The derived a-structure of *kunjungi* is shown in
(25c).

(25) a. *Mereka   ber-kunjung   ke   rumah sakit*
          3p   ber-visit   to   hospital
          'They visited the hospital' (Lit. they paid a visit to the hospital')

     b. *Mereka   mengunjungi   rumah sakit*
          3p   AV.visit-i   hospital
          'They visited the hospital.'

     c.   'they' 'hospital'
                 SUBJ   OBJ
          –i   <ARG1, ARG2   'visit < _ , _ >)'
                        (U:go)

3.4.2 Type 2

This type is associated with three-place predicates with a displaced theme such as *kirim* 'send'.
There are two sub-types (Type 2a and Type 2b), depending on the core status of the displaced theme.

In Type 2a, the displaced theme is unfused and realised as OBJ2 (cf., (20a)). The example given
earlier in (22) is *mengirimi* (AV.*kirim-i*), formed out of a free root, *kirim* 'send'. A bound root can also
appear in this pattern; e.g., *serah*- 'transfer', *sodor*- 'offer', and *suguh*- 'serve'. For this type of root,
the affix marking transitivity (*-i* or *-kan*) is obligatory. Consider the AV verb *menyuguhi* in (26a). Note
that for this root, there is no alternative monotransitive structure with the verb stem without a
transitiviser, as seen by the unacceptability of *menyuguh* (26b) where the displaced theme is ARG2-
OBJ. If this structure is intended, then *–kan* must be used (26c). The a-structures for (26a) and (26c)
are given in (26a') and (26c') respectively. The alternation of *–i* (26a) and *–kan* (26c) is expected for

---

i)   a. Ia   tinggal   di   Jakarta   b. * Ia   meninggal-i Jakarta
        3s   live   LOC   Jakarta   3s   AV.live-i   Jakarta
        'S/he lives in Jakarta.'   FOR "s/he lives in Jakarta.'
ii) a. Ali menangis di   kamar   b. * Ali menangisi   kamar
        Ali AV.cry   LOC room   Ali AV.cry-I   room
        'Ali cried in the room'   FOR: 'Ali cried in the room'

[8] We analyse that verbal bound roots carry argument structures. However, the syntactic core status of the
arguments are specified when the roots are affixed, e.g. by the intransitive prefix ber- or transitive *–i/-kan*.
Thus, in this analysis,–*kunjung* 'visit', for example, is a predicate with two arguments (the 'visitor' and the 'thing
visited'). The second argument is understood as Goal, which then fuses with ARG2 of *–i* in the derived *–i* verb.

reasons discussed in 3.3. That is, given the subcategorisation frame of –*i* shown in (20), the (unfused) theme can only be mapped onto OBJ2, because the other two arguments are fused with ARG1 and ARG2, realised as SUBJ and OBJ respectively.

(26)   a. *Engkau   menyuguh-i   aku   minuman lezat*
        2s   AV.serve-i   1s   drink   tasty
        'You served me a very tasty drink.'

    b. * *Engkau   menyuguh   minuman   lezat   kepada aku*
          2s   AV.serve   drink   tasty   to   1s
          FOR 'You served a very tasty drink to me.'

    c. *Engkau   menyuguh-kan   minuman   lezat   kepada   aku*
         2s   AV.serve-kan   drink   tasty   to   1s
        'You served a very tasty drink to me.'

a'.   '2s' '1s'   'tasty drink'
        SUBJ OBJ   OBJ2
          | | |
  –*i*   <ARG1, ARG2 'serve < _ , _ , _ >'
           (U:go)   (ag) (go) (th)

c'.   '2s' 'tasty drink'   '1s'
        SUBJ OBJ   OBL
          | | |
 –*kan*   <ARG1, ARG2 'serve < _ , _ , _ >'
           (U:th)   (ag) (go) (th)
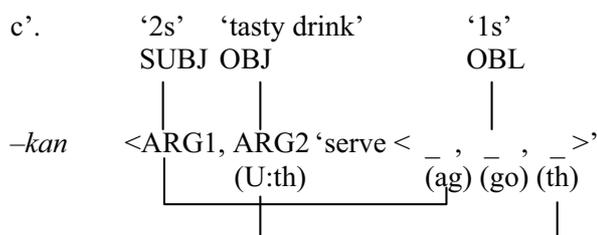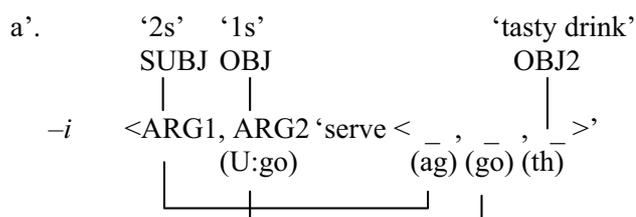
Type 2b is when the displaced theme of the underlying three-place predicate is non-core and hence surfaces as OBL. This is the type shown in (20b). The verb *menyuguhi* can be ditransitive (26a), or monotransitive (26c). Again, for reasons discussed in 3.3, these alternative structures are expected in our a-structure based analysis. The theme shows up with a prepositional instrumental marker *dengan*, as seen in example (27a) with the bound root -*suguh*- . More examples of this pattern, taken from the internet, are given in (27b-c)[9].

(27)   a. *Engkau   menyuguh-i   aku   dengan   minuman   lezat*
        2s   AV.serve-i   1s   with   drink   tasty
        'You served a very tasty drink to me'

    b. … *menyodor-i Juno   dengan   pertanyaan   tentang   nasib   sang   bayi ..*
         AV.offer-i Juno   with   question   about   fate   ART   baby
        ' …asked Juno questions about the fate of the baby …'

    c. *Dia   menyerah-i   saya   dengan   tugas   yang   berat*
        3s   AV.give-i   1s   with   duty   REL   heavy
        'He burdened me with heavy duties'

Our proposed analysis predicts that a sentence like (28) is unacceptable. This sentence is headed by –*i* but has an a-structure with ARG2/OBJ linked to a non-LOC argument, in violation of the restrictions imposed by –*i*.

(28)   * *Engkau   menyuguhi minuman lezat   aku*   (ARG2=theme)
         2s   AV.serve-i drink   tasty   1s
        (FOR: 'You served me with a very tasty drink')

---

[9] These examples were found on http://old.rumahfilm.org/esai/esai_juno.htm, and an online version of an Indonesian language learners' guide `Beginning Indonesian Through Self-Instruction' by Wolff et al (1992:890), accessed 27 November, 2009.

3.4.3   Type 3

Type 3 shows no valence change in –*i* derivation, exemplified by *pukul* 'hit'➔ *pukuli* (29a-b). Verbs that are of the same type as 'hit' can be represented as having an a-structure shown in (29a'), involving a 'hitter' (agent) and a 'hittee' understood as the patient/target (pt/go). The hittee, *saya* in (29), is the affected participant (i.e., patient) that is also the target of (i.e., in contact with) an understood movable instrument. The second argument of the stem meets the specification for –*i* in that it is a goal, and hence can fuse with the matrix ARG2. Double fusion of *memukuli*, as seen in (29b'), produces no valence change effect in the derivation.

(29) a. *Ia   memukul   saya*   (base verb: monotransitive)
           3s   AV.hit   1s
           'S/he hit me

      b. *Ia   memukul-i   saya*   (derived –*i* verb: monotransitive)
           3s   AV.hit-i   1s
           'S/he was hitting me, s/he hit me repeatedly'

      a.'   *pukul*   'hit < _ , _ >'
                              (ag)(pt/go)

      b'   '3s'   '1s'
                          SUBJ OBJ
                            | |
              –*i*   <ARG1, ARG2 'hit < _ , _ >'
                          (U:go)   (ag) (pt/go)

While not involving a change in transitivity, -*i* affixation does bring about a change in aspect; it indicates that the event is in progress or completed with iterative meaning as seen in the translation of (29b). This aspectual property of –*i* is not surprising on the analysis that ARG2/OBJ is semantically linked to a locative-goal role. That is, the –*i* verb leads to the interpretation that the action is applied to a surface of an object.   For inherently punctual verbs like 'hit', actions affecting the surface of an object are given a repetitive interpretation. For other verbs where OBJ measures event completion (Tenny 1992, 1994), e.g., *bunuh* 'kill' (where the object being dead measures the event), –*i* also gives rise to pluralisation or individuation of the object. For example, the object of *membunuh* in (30a) is by default singular, though plural is possible. However, the object of *membunuhi* in (30b) must be plural (i.e., reading i). For mass noun objects, -*i* gives rise to individuation; e.g., *tanah* 'land' sold in (31) is understood as 'fractions' of the land.

(30) a. *Dia membunuh   binatang   itu*
           3s   AV.kill   animal   that
           'S/he killed the animal(s)'

      b. *Dia membunuhi   binatang   itu.*
           3s   AV.kill-I   animal   that
           i) 's/he killed the animals one by one'
           ii) * 's/he killed the animal'

(31)   *Ia   menjual-i   tanah   orang tua-nya*
        3s   AV.sell-i   land   person old-3sPOSS
        'He sold his parents' land, bit by bit.'

3.4.4   Type 4

This is a single fusion type where ARG2, like the other types discussed earlier, fuses with a LOC argument of the base wherever possible. However, ARG1 is newly introduced in Type 4. This –*i* affixation then results in causativisation. Consider the bound root –*alir*- 'flow' in (32). It appears as an intransitive verb in (32a) with the locative/goal *sawah* as an oblique. The derived –*i* verb can be monotransitive (32b) (with the displaced theme showing up as an oblique instrument marked by *dengan*) or ditransitive (32c) (with the displaced theme being OBJ2). These two –*i* structures involve

single fusion as depicted in (32d), the only difference being the realisations of the unfused embedded displaced theme.[10]

(32)   a.   *Air   itu   sedang   meng-alir ke   sawah.*   SUBJ OBJ
               water   that   PROG   AV-flow to   rice.field   'flow < _ , _ >'
               'The water is flowing to the rice field.'   (th) (loc)

       b.   *Dia   meng-alir-i   sawah=nya   dengan air   itu.*
           3s   AV-flow-*i*   rice.field=3sg   with   water that
           'S/he flooded his/her rice field with the water.'

       c.   *Dia   meng-alir-i   sawah=nya   air   itu.*
           3s   AV-flow-*i*   rice.field=3sg   water   that
           'S/he flooded his/her rice field with the water.'

       d.   '3s'   'rice.field'   'water'
                SUBJ   OBJ   OBJ2/OBL
                 | | |
          –*i*   <ARG1, ARG2   'flow' < _ , _ >'
                 (U:go)   (th) (loc)

Denominal –*i* verbs with noun roots such as *sinar* 'light', *air* 'water' and *kutu* 'louse' are common examples of Type 4 –*i* verbs. The referent of the stem is understood as a displaced theme. The issue here that the embedded predicate of –*i*, i.e., the predicate that assigns 'displaced theme' to the stem, is not overtly realised. We posit an unexpressed predicate 'be.at' as part of the predicate a-structure of the verb. For example, the verb *mengatapi* (33a) can be interpreted as depicting an event where *atap* 'roof' is the displaced theme (i.e., placed on the house).

(33)   a. *Mereka mengatap-i   rumahnya*
          3p   AV.roof-i   house-3s
          'They roofed the house.'

       b.   '3p'   'house'
              SUBJ   OBJ
               | |
         –*i*   <ARG1, ARG2   'be.at' < 'house' , _ >'
               (U:go)   (th)   (loc)

Adjective roots (e.g., *sakit* 'sick', *panas* 'hot', and *kotor* 'dirty') can also derive –*i* verbs with causative meaning. This is exemplified in (34a). The fusion of the theme-locative argument shown in (34b) captures the meaning that *jalan* 'road' is understood as the surface of the road.

(34)   a. *Jangan kotor-i jalan   itu*
          NEG   dirty-i   road   that
          'Don't (you) make (the surface of) the road dirty.'

       b.   'you'   'road'
              SUBJ   OBJ
               | |
         –*i*   <ARG1, ARG2   'dirty' < _ >'
               (U:loc)   (th)

---

[10] In fact, the a-str of the type (32d) allows for double fusion if ARG1 is not filled in with an agent. Thus, the following is acceptable. The water flows because of its natural force.

    Air   itu   mengalir-i   sawahnya
    water that   AV.flow-I   rice.field
    'The water flooded his/her rice field.'

3.4.5 More than one type

A derived –*i* verb may be of more than one fusion type. This is exemplified by the –*i* verb derived from the root *tumpuk* 'pile up'. First consider the intransitive verb *bertumpuk* (35), in which the locative shows up as an oblique marked by *di*:

(35) *Buku-nya   ber-tumpuk   di   meja.*
       book-DEF   ber-pile.up   on   table
       'The books piled up on the table.'

The –*i* verb *menumpuki* can appear in three structures. It can be a three-place monotransitive verb, with the displaced theme appearing as an instrument. This is a derivation of Type 4 with single fusion, exemplified in (36a), whose fusion is represented in (36b).

(36) a. *Mereka   menumpuk-i   meja   itu   dengan   buku.*
          3p   AV.pile.up-I   table   that   with   book
          'They piled the table with books.'

       b.'   '3p'   'table'   'book'   (single fusion)
                          SUBJ   OBJ   OBL
                            | | |
              –i   <ARG1, ARG2 'pile.up < | , _ >'
                         (U:go)   (th) (loc)
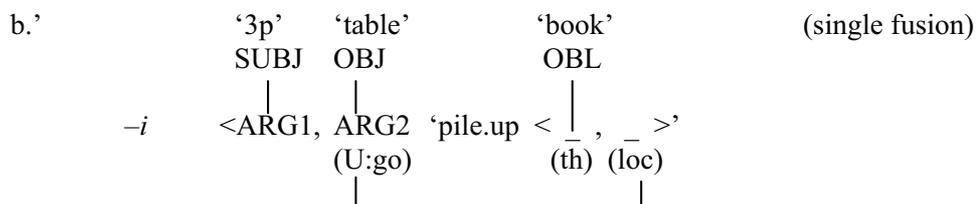
Menumpuki can also involve double fusion, with two variants. The first variant is a two-place monotransitive transitive verb, exemplified in (37a). ARG1 and ARG2 fuse with the embedded displaced theme and locative arguments respectively.

(37) a. *Buku itu   menumpuk-i   meja   itu*
          book that   AV.pile.up-i   table   that
          'The books piled up on the table.'

       c.'   'book'   'table'   (double fusion)
                          SUBJ   OBJ
                            | |
              –i   <ARG1, ARG2   'pile.up < _ , _ >'
                         (U:go)   (th) (loc)
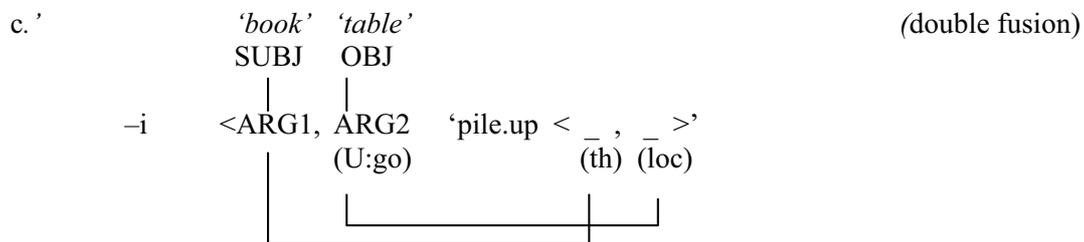
The second variant is just like the first variant, the difference being that double fusion gives rise to a reflexive reading (38). That is, ARG1/SUBJ instantiated by a human-denoting argument *mereka* 'they' invokes a volitional/agent reading, producing the meaning 'cause themselves to pile up'.

(38) a. *Mereka   menumpuk-i   meja   itu*
          3p   AV.pile.up-i   table   book
          'They piled themselves up on the table.'

       d.'   '3p'   'table'   (double fusion)
                          SUBJ   OBJ
                            | |
              –i   <ARG1, ARG2 'pile.up < _ , _ >'
                         (U:go)   (th) (loc)
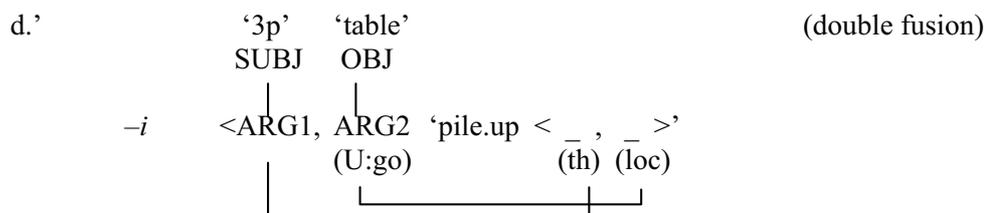
4   Implementation

We implement the analysis using XLE, [11] an LFG-based grammar development environment for large-scale grammars. Applicativisation as predicate composition has not been implemented in XLE before. Our analysis and implementation serve as evidence that (classic) LFG can adequately handle all types of –*i*, including both applicative and causative properties. It can robustly deal with derivation with or without valence increase, as well as with noun stems with no clear a-structure.

The relevant components of the system include a tokenizer and morphological analyser, as well as phrase structure and sublexical rules, which include annotation with relevant constraints as defined in templates for reusability. We omit description of the morphological analyser here; see Pisceldo et al. (2008), and Mistica et al. (2009) for full description and discussion. For present purposes, the relevant function of the morphological analyser is to decompose the words of a sentence into a stem plus morphological tags, often representing morphemes, which are analysed by sublexical phrase structure rules which are annotated in the same way as standard LFG phrase structure rules. For example, the morphological analyser breaks the verb *menduduki* into the four component parts: AV+ (Actor Voice), the stem *duduk,* +I (the –*i* suffix), and the part of speech tag +Verb. A sublexical phrase structure rule combines these components into a lexical V constituent.

4.1   Annotated phrase structure and sub-lexical rules

Our syntactic rules for Indonesian consist of phrase structure rules (c-str) and sublexical rules. The relevant (partial, somewhat simplified) c-str rules needed for –*i* verbs are shown in (39)-(40). The c-structure rules regulate clausal structure, and the sublexical rules are used to analyse the upper-side output of the morphological analyser, and hence regulate word-internal hierarchical structures. They are annotated with grammatical functions and templates. The templates for –*i* are those regulating VOICE and predicate composition (APPL and CAUS). Template calls are indicated by @ in the rules, as is standard.

(39)   Clausal c-str rules:

   a.   S → NP VP

   b.   VP → V' PP

   c.   V' → (NP) V (NP)

(40)   Sublexical rules:

 a. V   → V_VOICE_BASE V_STEM' V_SFX_BASE.
 b. V_STEM' → {V_STEM-APPL_I   _   V_I_BASE
        @(VOICE @(APPL_I VApp_I))
        |
        V_STEM-CAUS_I   V_I_BASE
        @(VOICE @(CAUS_I VCaus_I)) }.

It should be noted that the sublexical structure is not flat, as is often the case in XLE grammars. This is linguistically motivated by word formation patterns in Indonesian, and is also practically useful, e.g. the rule for V_STEM' (used in the analysis of the suffix -*i* (with the tag +I; see (44e))) is specified only once for all types of voice selections. Following the notation and structure in XLE, affix positions are represented in (40) as V_VOICE_BASE for the verbal prefix (e.g., *meN-*) and V_I_BASE for the applicative/causative suffix –*i*. The V_SFX_BASE is also added to provide a slot for verbal category information. These sublexical positions are filled in by the tags of the relevant affixes; e.g., tag AV+ for the AV prefix meN- and +I for the causative/applicative –*i* (see (44) sample entries with tags).

The template for Voice, given in (41), specifies three voice types in Indonesian: ACTOR-VOICE, UNDERGOER-VOICE, and PASSIVE-VOICE. Each comes with a template (not shown here) and imposes changes in the input subcategorisation frame, formulated here as classic lexical rules in LFG (Bresnan 1982). For example, PASSIVE-VOICE replaces OBJ by SUBJ, and specifies three

---

options for how the underlying SUBJ is realised[13] (suppressed as NULL, OBL, or OBJ). Note that ACTOR-VOICE does not change the subcategorisation frame of the base, so if the verb is recognised as AV (prefixed with *meN-*), the grammar just passes back the subcategorisation frame as is.

(41) Templates for Voice and Voice types in Indonesian
$$VOICE(\_SCHEMATA) =$$
$$\{ \_SCHEMATA$$
$$@ACTOR\text{-}VOICE$$
$$| \_SCHEMATA$$
$$@UNDERGOER\text{-}VOICE$$
$$(\uparrow OBJ) \rightarrow (\uparrow SUBJ)$$
$$(\uparrow SUBJ) \rightarrow (\uparrow OBJ)$$
$$| \_SCHEMATA$$
$$@PASSIVE\text{-}VOICE$$
$$(\uparrow OBJ) \rightarrow (\uparrow SUBJ)$$
$$\{ SUBJ) \rightarrow NULL$$
$$| (\uparrow SUBJ) \rightarrow (\uparrow OBL)$$
$$| (\uparrow SUBJ) \rightarrow (\uparrow OBJ) \} \}.$$

The template in (42) encodes the contribution of the applicative *–i* affix:

(42) APPL_I =
$$\{(\uparrow PRED) = \text{'V\_Appl\_i} <(\uparrow SUBJ) (\uparrow OBJ) \%PRED3>\text{'}$$
$$\uparrow \backslash PRED \backslash GF = \downarrow \backslash PRED \backslash GF$$

$\{ (\downarrow SUBJ) = (\uparrow SUBJ)$	**Type 1: IntrRoot → Vtr**
$(\downarrow OBL\text{-}LOC) = (\uparrow OBJ) \quad$ |	
$(\downarrow SUBJ) = (\uparrow SUBJ)$	
$(\downarrow OBL\text{-}LOC) = (\uparrow OBJ)$	**Type 2: TrRoot → Vtr**
$(\downarrow OBJ) = (\uparrow OBL\text{-}INST)$	
$(\uparrow OBL\text{-}INST\ CASE) =c\ obl\text{-}inst \quad$ |	
$(\downarrow SUBJ) = (\uparrow SUBJ)$	
$(\downarrow OBJ) = (\uparrow OBJ)$	**Type 3: TrRoot → Vtr**

$$(\uparrow TNS\text{-}ASP\ PROG) = +$$
$$\sim(\uparrow OBL\text{-}INST) \quad \text{"just for the iterative meaning of } –i\text{"} \}$$
$$(\downarrow PRED) = (\uparrow PRED\ ARG3) \quad |$$
$$(\uparrow PRED) = \text{'V\_Appl\_i} <(\uparrow SUBJ)(\uparrow OBJ) (\uparrow OBJ2) \%PRED4>\text{'}$$
$$\uparrow \backslash PRED \backslash GF = \downarrow \backslash PRED \backslash GF$$

$(\downarrow SUBJ) = (\uparrow OBJ)$	
$(\downarrow OBL\text{-}LOC) = (\uparrow OBJ)$	**Type 4: IntrRoot → Vtr**
$(\downarrow OBJ) = (\uparrow OBJ2)$	

$$(\downarrow PRED) = (\uparrow PRED\ ARG4) \}$$
$$(\uparrow APPLICATIVE) = +.$$

This template implements the analysis discussed in section 3.2, using the restriction operator (Butt, King, and Maxwell 2003; Butt and King 2006). The implementation is somewhat constrained by the current setup of XLE. Given that subcategorisation/a-structure frames are represented by GFs in XLE, rather than having a separate argument structure representation encoding thematic roles, ARG1 and ARG2 in our earlier representation in section 3 correspond to SUBJ and OBJ respectively in (42). Additionally, rather than using underspecification, each type of combination of *–i* with a predicate is treated separately by a set of equations. %PRED represents the base PRED, which is treated as the rightmost ARG of the V_Appl_I predicate; hence %PRED3 or %PRED4. When *–i* appears as an affix

---

[13] It should be noted that a *di*-passive verb may have its Actor realised as an enclitic, e.g. *di-pukul=nya* 'di-hit=3s', in which case the actor behaves more like a core argument than an oblique (see Musgrave 2001; Arka and Manning 2008).

to a verb, this results in predicate composition. For example, simplifying the sublexical tree somewhat, we show the structure of the verb *duduki* 'sit.APPL' in (43). The effect of predicate composition is shown in the box. In addition to supplying the information that the *–i* verb is applicative (i.e. ($\uparrow$ APPLICATIVE)= +), the template specification ($\downarrow$PRED)=($\uparrow$ PRED ARG3) in the template results in the embedded PRED being the third (subordinate) argument of the applicative verb. The equation ($\downarrow$ SUBJ) = ($\uparrow$SUBJ) of the template captures the fusion of the subordinate argument with the matrix argument, which is indicated by a connecting line in the box.

(43)

($\uparrow$ PRED) = 'V_App_I<($\uparrow$ SUBJ)($\uparrow$OBJ) 'duduk<__>'>'

($\uparrow$ APPLICATIVE)= +

V

V_STEM
@APPL_I

Suffix

duduk
($\uparrow$ PRED)= 'duduk<($\uparrow$ SUBJ)>'

-i
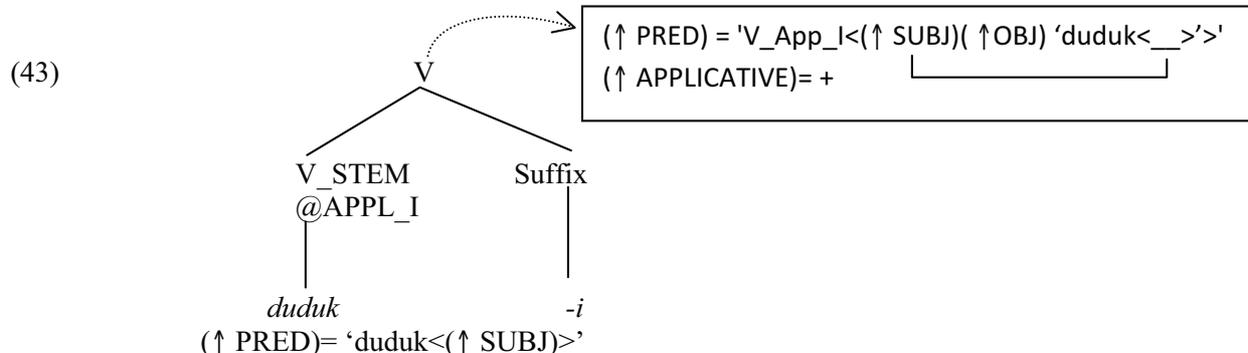
As shown in (43), the restrictions associated with *–i* (the template shown in (42)) appear in the sublexical c-structure rule introducing *–i* as seen in (43). That achieves the desired effect: when *–i* is affixed, the composition affects the stem's a-structure specification. It should be noted that the template is actually specified as @(VOICE @(APPL_I VApp_I) in the sublexical rule shown in (40b). The internal bracketing of @(VOICE @(APPL_I VApp_I)) determines how applicativisation interacts with voice selection: Predicates are first composed by applicativisation, then a particular VOICE-TYPE is selected.

4.2   Tags and lexical entries

Apart from the annotated rules just described, at the heart of the grammar is the lexical entries. Information specified in the lexical entries is of different kinds. This includes information about the grammatical category (N, V, Affix, etc.) needed by c-structure or sublexical rules to place the relevant item in the correct position, semantic information (i.e. the presence of PRED) and other functionally related information (e.g. voice specification bundled in templates as seen in (41)). Sample lexical entries are given in (44). (The functional specifications associated with *–i* (i.e., @APPL_i shown in (42)) are annotated in the sublexical rule of *–i* (41b), and are therefore not shown in the lexical entry.)

(44)   Sample entries: free forms

a.	kursi	N	@(CN chair).
b.	dia	PRON	@(PPRO 3 sg).
c.	duduk	V	{@(INTRANS sit)| @(INTRANS_SEM_PP sit loc) }
d.	di	P	@(PREP in loc);
		Pcase	@(PCASE to obl-dir dir)

Sample entries: bound forms

e.	+I	V_I .	
f.	AV+	V_VOICE	@(VOICE-TYPE AV).
g.	UV+	V_VOICE	@(VOICE-TYPE UV).
h.	PASSdi+	V_VOICE XLE	@(VOICE-TYPE PASSIVE).

Note that affixes are listed as lexical entries for their abstract tags, rather than morphological forms. That is, in the two-level morphology adopted here where an affix form such as *meN-* is treated as a string realised as the prefix *meN-* (lower-side) and corresponding to the morphological tag AV+, it is the tag AV+ and not the morphological form *meN-* that is listed in the lexical entry. It is then associated with the AV restriction as specified by the template @(VOICE-TYPE AV).

Within the templates CN, PPRO, INTRANS, INTANS_SEM_PP, PREP and PCASE are defined features related to the lexical entry they annotate.

4.3   Parsing and sample parses

To illustrate how sentences with *–i* verbs are parsed, we provide a few output parses with a brief description of their analyses.

For (45a), the input string is first broken into tokens by the tokenizer. The output is then fed into the morphological analyser so that morphologically complex words such as *menduduki* and *kursi* can be analysed and assigned morpheme and category tags as in (45b). Since the relevant tags and forms are listed in the lexical entries, e.g. AV+ (for *meN-* and *–i* (see (44)), the XLE parser is able to recognise the tags, and use the information to assign the word a hierarchical structure on the basis of the sublexical rules as formulated in (40). In addition, given the functional constraints carried by the morphemes and the structures (cf. template calls signalled by @ in the entries and in the sublexical rules), the parser can also build functional structures involving predicate composition for the *–i* verb based on the defined grammar rules. The output c- and f- structures are displayed in (45c)..

(45)   Tokenizing and Morpheme identification:
   a.   Input string: *Ia menduduki kursi*
   b.   Morphologically analysed string: Ia AV+ *duduk* +I +Verb kursi

   c.
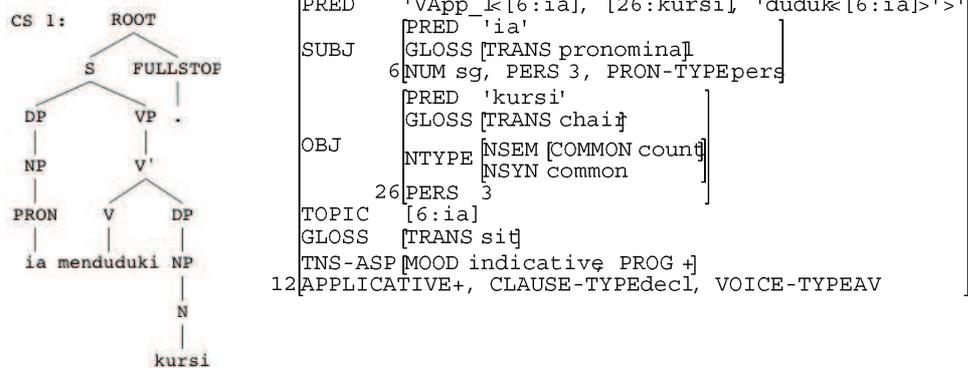
Example (46) shows a sentence with the monotransitive base *pukul* 'hit'. The derivation does not change the transitivity, but marks progressive aspect (indicated by [PROG +] in the f-str).

(46)   a. *Mereka   memukul-i   kami*
         3p   AV.hit-i   1p.ex
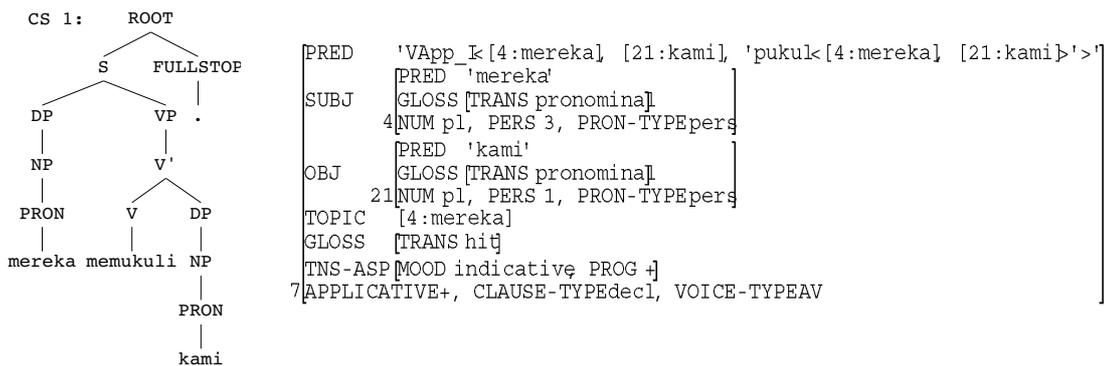
         'They were hitting us.'

5   Concluding remarks

We have presented an a-str based analysis for the applicative-causative polysemy of *–i*. We have shown that different properties of -i — valency-increasing or no valence changing effects as well as the related interative/progressive meaning — and the alternation of *–i* with *–kan*, are predictable from the interaction of the core information carried by the suffix and the information carried by the stem.

Building on earlier work (Alsina 1996; Butt 1995), we present a predicate composition analysis involving underspecified argument structure, which allows two types of argument fusion: single and double. We claim that, at least for the Indonesian –*i*, the fusion follows a natural general rule in which arguments of thematically similar types tend to fuse.

We have demonstrated the implementation of the analysis in XLE. The computational grammar we are developing shows promise how predicate composition can handle verbal derivation with –*i* as part of verb formation in Indonesian. The grammar can correctly identify a range of different aspects of –*i*, in particular its interaction with the VOICE system and the different possibilities of the syntax of –*i*.

6   References

Alsina, A. 1992. On the argument structure of causatives. *Linguistic Inquiry* 23:517–555.

———. 1996. *The role of argument structure in grammar*. Stanford: CSLI.

Alsina, Alex, and Smitha Joshi. 1991. Parameters in causative constructions. In *CLS* 27:2–16.

Arka, I Wayan. 1993. The -kan causative in Indonesian. MPhil Thesis, University of Sydney, Sydney.

———. 2003. *Balinese morphosyntax: a lexical-functional approach*. Canberra: Pacific Linguistics.

Arka, I Wayan, and Christopher Manning. 2008. Voice and grammatical relations in Indonesian: a new perspective. In *Voice and grammatical relations in Austronesian Languages*, edited by P. K. Austin and S. Musgrave, 45-69. Stanford: CSLI.

Austin, P. 2005 [1996]. Causatives and applicatives in Australian Aboriginal languages, edited by K. Matsumura. Tokyo: Hitsuji Shobo.

Beesley, Kenneth R., and Lauri Karttunen. 2003. *Finite State Morphology*. Stanford: CSLI.

Bresnan, Joan. 1982. The passive in lexical theory. In *The Mental Representation of Grammatical Relations* edited by J. Bresnan, 3–86. Cambridge, MA.: The MIT Press.

Bresnan, Joan, and Jonni M. Kanerva. 1989. Locative inversion in Chichewa: a case study of factorization in grammar. *Linguistic Inquiry* 20:1-50.

Bresnan, Joan, and L. Moshi. 1990. Object asymmetries in comparative Bantu Syntax. *Linguistic Inquiry* 21 (2):147-185.

Butt, Miriam. 1995. *The Structure of Complex Predicates in Urdu.* Stanford: CSLI Publications.

Butt, Miriam, and Tracy Holloway King. 2006 Restriction for morphological valency alternations: the Urdu causative. In *Intelligent Linguistic Architectures: Variations on Themes.* , edited by R. M. Kaplan. Stanford: CSLI.

Butt, Miriam, Tracy Holloway King, and John T Maxwell III. 2003. Complex predicates via restrictions. In Proceedings of the LFG'03 Conference, edited by M. Butt and T. H. King. CSLI, http://csli- publications.stanford.edu/LFG/8/lfg03.html.

Chung, Sandra. 1976. On the subject of two passives in Indonesian. In *Subject and topic*, edited by C. N. Li, 57-99. NewYork: Academic Press.

Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, Massachusetts: MIT Press.

Kaplan, Ronald M., and Jürgen Wedekind. 1993. Restriction and correspondence-based translation. In Proceedings of the Sixth European Conference of the Association for Computational Linguistics.

Kroeger, Paul. 2007. Morphosyntactic vs. morphosemantic functions of Indonesian -kan. In *In Architectures, Rules, and Preferences: Variations on Themes of Joan Bresnan*, edited by A. Zaenen, J. Simpson, T. H. King, J. Grimshaw, J. Maling and C. Manning, 229–251. Stanford, CA: CSLI Publications.

Levin, B, and M Rappaport Hovav. 1998. Morphology and Lexical Semantics. In *The Handbook of Morphology*, edited by In, A. Spencer and A. M. Zwicky. Oxford: Blackwell.

Levin, Beth. 1993. *English verb classes and alternations : a preliminary investigation*. Chicago University of Chicago Press.

Macdonald, R. Ross and Dardjowidjojo, Soenjono. 2001. *A student's reference grammar of modern formal Indonesian*. Georgetown: Georgetown University Press.

Manning, Christopher D. 1996. *Ergativity: argument structure and grammatical relations*. Stanford: CSLI.

Mistica, Meladel, Wayan Arka, Timothy Baldwin and Avery Andrews. 2009. Double Double, Morphology and Trouble: Looking into Reduplication in Indonesian, Paper presented at the 2009 Australasian Language Technology Workshop (ALTW 2009), Sydney, Australia

Moeliono, Anton M, and Soenjono Dardjowidjojo, eds. 1988. *Tata bahasa baku Bahasa Indonesia* Jakarta: Balai Pustaka.

Musgrave, S. 2001. Non-subject arguments in Indonesian. PhD thesis, Melbourne University.

Myhill, J. 1988. Agent incorporation in Indonesian. *Journal of Linguistics* 24 (1):111-136.

Pisceldo, Femphy, Rahmad Mahendra, Ruli Manurung, and I Wayan Arka. 2008. A Two-Level Morphological Analyser for Indonesian.
http://rspas.anu.edu.au/linguistics/projects/iwa/IndonParGram/papers/Alta.morpho.pdf

Purwo, Bambang Kaswanti. 1989. Voice in Indonesian : A Discourse Study. In *Serpih -serpih telaah pasif bahasa Indonesia*, edited by B. K. Purwo, 344-442. Jogyakarta: Kanisius.

———. 1995. The Two Proto-types of Ditransitive Verbs: The Indonesian Evidence. In *Discourse grammar and typology*, edited by W. Abraham, T. Givon and S. A. Thompson, 77-99. Amsterdam: John Benjamins Publishing Company.

Simpson, J. 1991. *Warlpiri Morphosyntax*. Dordrecht: Kluwer.

Sneddon, James. 2006. *Colloquial Jakartan Indonesian*. Canberra: Pacific Linguistics.

Sneddon, James N. 1996. *Indonesian reference grammar*. St Leonards: Allen and Unwin.

Tenny, C. 1992. The Aspectual Interface Hypothesis. In *Lexical Matters*, edited by I. Sag and A. Szabolcsi, 1-28. Stanford: CSLI.

———. 1994. *Aspectual Roles and the Syntax-Semantics Interface*. Dordrecht: Kluwer.

Vamarasi, Marit Kana. 1999. *Grammatical relations in Bahasa Indonesia*. Canberra: Pacific Linguistics.

Verhaar, John W. M. 1984. The Categorial System in Contemporary Indonesian: Verbs. In *Towards a Description of Contemporary Indonesian: Part I*, edited by K. Purwo, 27-64. Jakarta: NUSA , Universitas Atma Jaya.

Wolff, John U., Dede Oetomo, Daniel Fietkiewicz. 1992. *Beginning Indonesian Through Self-Instruction: Book 3, Lessons 16-25*. New York: Cornell Southeast Asia Program Publications.