1-1-2017

Variants of Indonesian Prepositions as Intra-speaker Variability at PF

Helen Jeoung
University of Pennsylvania

Alison Biggs
University of Pennsylvania

Variants of Indonesian Prepositions as Intra-speaker Variability at PF

Abstract

This paper discusses the formal representation of morphosyntactic intra-speaker variability in a modular grammar. It presents novel data from Indonesian functional prepositions that exhibit unusual variability in form. It is shown that *oleh* 'by' and *dengan* 'with' may either (i) be phonologically unrealized (Preposition-drop), a process sensitive to syntactic structure, or (ii) may be realized as *sama* 'by/with,' a variant speakers use in an informal style. We establish that it is necessary to invoke distinct processes to formally introduce these outputs: Indonesian preposition-drop is argued to reflect a morphophonological variable deletion rule, while style-shifting is modeled as competition in use of lexical inventories.

The variable realization of Indonesian prepositions indicates that it is possible to identify distinct variable processes (competition and variable rules), here operating over a single morphosyntactic item. The appeal to two variable processes follows in part from the nature of the factors (internal or external to the grammar) that condition variability, but we argue that it also follows if implementation additionally obeys constraints imposed by the grammatical architecture. Finally, we find that, in order to maintain modularity, variable processes may have distinct loci in the post-syntactic derivation.

Variants of Indonesian Prepositions as Intra-speaker Variability at PF

Helen Jeoung and Alison Biggs*

1 Introduction

This paper examines the generation of intra-speaker variability in a derivational PF, based on novel data from Indonesian where prepositions have multiple and variable surface forms:[1]

(1) Buku ini   di-baca   (oleh/sama) adik.
    book this PV-read   by   younger.sibling
    'This book was read by little brother.'
(2) Aku kaget   di-tegur (oleh/sama) guru-ku.
    1SG shocked PV-scold by   teacher-1SG
    I was shocked to be scolded by my teacher.'
(3) Tangan-nya di-ikat (dengan/sama) tali   plastik.[2]
    hand-DEF   PV-tie with   string plastic
    'His hands were bound with plastic cord.'

In (1-3) each preposition has three morphological variants: 'by' is optionally realized as *oleh, sama* or null, while 'with' is optionally realized as *dengan, sama* or null. The alternation between the different realizations can be straightforwardly characterized, at least descriptively: the forms *oleh* 'by' and *dengan* 'with' can be regarded as 'base' forms. The base forms vary first with the informal or colloquial form *sama* 'by/with.' In addition, each of the forms *oleh*, *dengan*, and *sama* can be null/unpronounced: we refer to this as variable P(reposition)-drop.

The variant exponents in (1-3) are 'optional' only in the sense that variability in the form of functional Ps in Indonesian does not reflect variation in syntax or semantics, as shown in Section 2. Section 3 shows that the variant Ps in (1-3) are not in free variation. Variable realization of *sama* is conditioned by context: use of *sama*, as the informal form of P, reflecting a particular style. Conditions on intra-speaker variable P-drop are more complex: Section 2 shows that variable P-drop is only grammatical in a certain linguistic structure, which acts as a 'pre-condition' on the possibility of P-drop. The use of null P then also reflects an informal style. Section 3 explores the formal generation of intra-speaker variability in (1-3) in a derivational PF that integrates both types of conditioning factor (linguistic structure and style-shifting). We argue that if we wish to maintain a modular linguistic system – where modular means that linguistic structure is part of linguistic competency, while style-shifting is part of a distinct system of use or performance (Chomsky 1965, Labov 1969, 2001) – then both variable rules and a form of Competing Grammars (Kroch 1989) are needed to generate the variants in (1-3).

2 The Structure of Indonesian P-drop

2.1 Condition 1: Initiator Ps

P-drop is possible only in a limited structural environment in Indonesian. First, P-drop is only possible where the nominal introduced by P is interpreted as an Initiator. Broadly, the Ps *oleh* 'by' and *sama* 'by/with' introduce DP arguments bearing a variety of thematic roles. (4-8) show that

non-pronunciation is possible with Agent (4-5), Causer (6) or Experiencer (7-8) arguments.

(4) Buku ini   dibaca   (oleh/sama) adik.   Agent PP
    book this PV-read   by   younger.sibling
    'This book was read by little brother.'

(5) Trotoar ini   di-tabrak   (oleh/sama) mobil   itu.   Agent PP
    curb   this PV-crash   by   car   that
    'The curb was hit by the car.'

(6) Jendela kamar itu   di-buka   (oleh/sama) angin sejuk.   Causer PP
    window room   that PV-open   by   wind   cool
    'The window was opened by a cool wind.'

(7) bunyi yang bisa di-dengar (oleh/sama) manusia   Experiencer PP
    sound that can PV-hear   by   human
    'a sound that can be heard by humans'

(8) Russell Brand di-benci (oleh/sama) masyarakat Inggris.   Experiencer PP
    Russell Brand PV-hate by   community English
    'The Russell Brand is hated by the English.'

These examples also demonstrate that P-drop is insensitive to the animacy or definiteness of the
DP, which may be animate (4, 7, 8), inanimate (5, 6), definite (5), indefinite (6, 8) or generic (7).
    As shown in (2), Instruments are introduced with *dengan* 'with' or *sama* 'by/with.' However,
there is a split: P may be null in (9-10) but must be overt in (11-12):

(9) Celana jeans saya bisa di-potong (dengan/sama) gunting itu.   Instrument-causer PP
    trouser jeans 1SG can PV-cut   with   scissors that
    'My jeans can be cut with those scissors.'

(10) Tangan-nya di-ikat (dengan/sama) tali   plastik.   Instrument-causer PP
     hand-DEF   PV-tie with   string plastic
     'His hands were bound with plastic cord.'

(11) Anak-ku di-suap-i   *(dengan/sama) sendok.   Pure Instrument PP
     child-1SG PV-feed-LOC with   spoon
     'My child was fed with a spoon.'

(12) Nasi boleh di-makan *(dengan/sama) tangan.   Pure Instrument PP
     rice   may PV-eat   with   hand
     'Rice may be eaten by hand.'

The contrast between the Instruments in (9-10) and those in (11-12) corresponds to an independent
distinction between Pure Instruments and Instrument-causers (Kamp and Rossdeutscher 1994):
Instrument-causers can act on their own, after being introduced by an Agent; Pure Instruments are
directly employed by an Agent but are 'peripheral' to the event which they help to bring about.
The two classes can be diagnosed syntactically: the former, but not the latter, can be the subject of
an active clause. As demonstrated in (13-14), when (11-12) are in the active voice, the Instruments
sendok 'spoon' and *tangan* 'hand' cannot occur as subject because they are Pure Instruments:

(13) *Sendok itu   meny-uap-i   anak-ku.   Pure Instrument PP
     spoon   that AV-feed-LOC child-1SG
     'The spoon feeds my child.'

(14) *Tangan itu   boleh makan nasi.   Pure Instrument PP
     hand   that may   eat   rice
     'The hand may eat rice.'

In contrast, the Instrument-causers *gunting* 'scissors' and *tali plastik* 'plastic cord' may occur as
subject in an active voice clause. These examples also allow P-drop (9-10):[3]

_____

[3](15) is grammatical if some situation renders the scissors able to cut jeans. It is not if the scissors are
construed as directly employed by an Agent. Conversely, *sendok* "spoon," which is a Pure Instrument in (11)

(15) Gunting itu   bisa me-motong celana jeans saya.   Instrument-causer PP
     scissors that can AV-cut   trouser jeans 1SG
     'Those scissors can cut my jeans.'

(16) Tali   plastik itu meng-ikat tangan-nya.   Instrument-causer PP
     string plastic that AV-tie   hand-DEF
     'The plastic cord tied his hands.'

We leave analysis aside, and simply take the contrast between (13-14) and (15-16) as evidence for a (likely structural) alternation between the Instrument types, to which P-drop is sensitive.

In summary, P-drop is grammatical with Agent, Experiencer, Causer, and Instrument-causer arguments, but not Pure Instruments. We label the grammatical set 'Initiator Ps.'

2.2 Condition 2: Passive Voice, Marked by *di-*

Second, Initiator P-drop only occurs where the verb is marked by the passive voice prefix *di-*, and is not possible in other voices.[4] To illustrate, the overt P *dengan* 'with' is required with in Active Voice (with a topicalized object) (17) and in Object Voice (18):[5]

(17) Tangan-nya, aku akan ikat   *(dengan) tali   plastik.   Active voice
     hand-DEF   1SG will AV. tie   with   string plastic
     'His hands, I will bind with plastic cord.'

(18) Tangan-nya akan ku-ikat   *(dengan) tali   plastik.   Object voice
     hand-DEF   will 1SG-OV.tie   with   string plastic
     'I will bind his hands with plastic cord.'

Nor is P-drop possible with Indonesian '*ter-* verbs' that take *oleh* by-phrases (Sneddon et al 2012). 'Accidental' *ter-* indicates a non-intentional event (19), and 'abilitative' *ter-* indicates that the action or event takes place despite being difficult or unexpected (20). P-drop is not possible in either environment with *ter-*:

(19) Pintu   itu   ter-buka   *(oleh/sama) angin.   Accidental *ter-*
     door   that Accid-open   by   wind
     'The door was (accidentally) opened by the wind.'

(20) Pesawat itu   masih ter-lihat *(oleh/sama) Pak Ahmad.   Abilitative *ter-*
     airplane that still Abil-see   by   Mr. Ahmad
     'The plane can still be seen by Mr. Ahmad' (even though it is very high in the sky).

Initiator P-drop, then, is only possible if the verb is marked with the prefix *di-*. We assume *di-* is the spellout of passive Voice^0; and that in the passive, Initiator PPs are adjoined to VoiceP.

2.3 Condition 3: Linear Adjacency Between PP and the Passive Verb Complex

Finally, Initiator PP must be linearly adjacent to the verb complex. (21-23) show that if an Initiator PP occurs elsewhere than immediately to the right of the verb, an overt P is required. (21) shows

---

and (13), can be an Instrument-causer in the coerced context of (i), where it is sharp enough to cut fabric; in this case it may then occur as subject in the active (ii), and may undergo P-drop (i):

 (i) Celana jeans saya bisa di-potong (dengan/sama)   sendok itu.
     trouser jeans 1sg can PV-cut   by   spoon that
     "My jeans can be cut with the spoon."

(ii) Sendok itu   bisa me-motong celana jeans saya.
     spoon   that can   AV-cut   trouser jeans 1sg
     "The spoon can cut my jeans."

[4]Past descriptions note that *oleh* "by" and *dengan* "with" can sometimes be unpronounced after a passive verb (e.g., Macdonald and Darjowidjojo 1967, Dardjowidjojo 1978, Arka and Manning 2008); Sneddon et al. 2012 also briefly lists some conditions under which they must be pronounced.

[5]Object voice is distinct from passive and active (Chung 1976, Guilfoyle et al 1992, Cole et al 2008).

that P-drop is ungrammatical when a *by*-phrase is focused and appears in clause initial position. P-drop is also impossible if any item intervenes between the verb and the PP, such as an adverb (22) or a theme object in a double-object construction (23).

(21) *(Oleh) Pak Rudiyanto buku itu   di-tulis   pada tahun 2001.
       by   Mr Rudiyanto book that PV-write at   year 2001
       'By Mr Rudiyanto that book was written in 2001.'

(22) Aku   kaget   di-tegur kemarin   *(sama) guru-ku.
       1sg   shocked   PV-scold yesterday by   teacher-1SG
       'I was shocked to be scolded yesterday by my teacher.'

(23) Semua anak di-jahit-kan baju *(oleh) Ibu Mindy.
       all   child PV-sew-Ben   shirt by   Ms. Mindy
       'All the children were sewn shirts by Ms. Mindy.'

(21-22) might be taken as an indication that P-drop is unavailable when the Initiator PP has moved out of SpecVoiceP; (23) shows this is an insufficient description. Characterizing the condition as linear adjacency captures the fact that an unpronounced lower copy of the object that raises to SpecTP (e.g., in 1-3) does not block P-drop, in contrast to the overt object in (23). P-drop is ungrammatical only where (phonological) material linearly – not structurally – intervenes between the passive verbal complex and PP. Thus (21-23) suggest that P-drop reflects post-syntactic variability, a point developed in detail in the next section.

2.4 Evidence for a Post-syntactic Account of P-drop

Overt and null forms might correspond to (at least) (a) distinct feature bundles in the lexicon; (b) alternations in syntactic structure; or (c) variation in (morpho)phonology, following transfer from the syntactic computation. As already hinted, the last option appears correct for Indonesian P-drop.

The requirement of linear adjacency described in Section 2.3 might suggest that P-drop is licensed by (pseudo-)incorporation, particularly since P-drop in other languages has been argued to reflect some type of incorporation (e.g., Ioannidou and den Dikken 2009, Gehrke and Lekakou 2013, Myler 2014). In fact, Myhill (1988) proposes that examples such as (1-3) involve N-incorporation into the verb. Close investigation, however, shows Indonesian Initiator P-drop is not consistent with properties of incorporation.

(Pseudo-)incorporation may manifest in different ways in different languages, but Agent/ Initiator incorporation is believed to be otherwise typologically unattested (Mithun 1984). Typically, incorporation has a semantic or syntactic reflex; for example, the nominal might be interpreted as generic, indefinite, non-referential, or non-individuated. None of these apply (4-8). Nor are any other semantic distinctions reported on the DP when P is null or overt: DP is not interpreted as more or less agentive, specific or non-specific; the activity denoted by the verb is not required to be 'conventionalized.' P-drop is insensitive to the syntactic or semantic class of the verb. The DP can be heavy or complex, even embedding a relative (24), and possessed (25):

(24) Di tengah jalan saya di-sapa   (oleh/sama) seseorang yang ternyata   dari Singapura.
       at middle street 1SG PV-greet by   someone   that apparently from Singapore
       'In the middle of the street, I was greeted by someone that is apparently from Singapore.'

(25) Aku di-sapa   (oleh/sama) teman-ku.
       1SG PV-greet by   friend-1SG
       'I was greeted by my friend.'

The data suggest, then, that the syntactic structures in which overt and null Initiator P occur are equivalent. A consequence is that a postverbal preposition-less Initiator should behave as an adjunct, rather than a verbal argument. Arka and Manning (2008) present a contrast in the interpretation of reflexives that shows this is the case: working in the LFG framework, they show that the postverbal Agent is an oblique argument rather than a 'term' or core argument.[6] In (26) the

---

[6]Fortin 2007 and Kroeger 2014 also assume null P in Indonesian, but do not provide an analysis.

object voice Agent (*kau* 'you') can bind a fronted reflexive (*dirimu* 'yourself'). We assume this is possible because the Agent remains in its thematic position (see Cole et al. 2008) and c-commands the reflexive in its base (thematic object) position. In contrast, in the passive voice an Agent (*dia* 'he') cannot bind the reflexive (27), because this Agent is embedded in a PP *by*-phrase, and does not c-command the VP internal object position.[7]

(26) Dirimu   mesti kau serahkan ke polisi. (Arka and Manning 2008, ex 16b, 6a)
     self.2   must 2   surrender to police
     'You must surrender yourself to the police.'

(27) ??Dirinya yang di-ajukan   sebagai calon   oleh dia.
     self.3   REL   PV-nominate as   candidate by   3SG
     'It is himself$_1$ that is nominated as a candidate by him$_1$.'

In (28), P-drop is licensed because the PP immediately follows the verb:

(28) ?*Dirinya selalu di-utamakan ~~dirinya~~ (oleh) Amir. (Arka and Manning 2008, ex 29b)
     self.3   always PV-prioritise self.3   by   Amir
     'Amir$_1$ is always prioritizing himself$_1$.'

Crucially, in (28) the binding facts remain consistent whether P is overt or null. If *Amir* is not embedded in a PP but is rather a verbal argument (as proposed in Nomoto and Kartini 2014) *Amir* should be able to bind *dirinya* 'self' as the Agent *kau* 'you' does in (26). However, *Amir* never binds the reflexive, and so must be embedded in a prepositional adjunct even when P is null.[8, 9]

    In summary, P-drop does not reflect variability in syntactic structure, nor is there any evidence that zero P reflects a different functional item or feature bundle that feeds the syntactic derivation; both should exhibit some syntactic or semantic reflex. Variability in Indonesian Initiator P therefore reflects a post-syntactic alternation, namely morpho-phonological variability.

    It is often suggested that morpho-/syntactic intra-speaker variability is only apparent, and does not (ever) reflect truly equivalent syntactic and semantic structures (Lavandera 1978, Newmeyer 2003), such that variable output forms reflect distinct inputs, unlike the variables typically identified in sociophonetics. This is not an issue for our case study if, as above, the syntactic/semantic object that is the input to the morpho-phonology is invariant.

3 Two Types of Intra-speaker Variability in a Derivational PF

3.1 Internal and External Factors Conditioning Variability in Indonesian P

We now turn to the representation of intra-speaker variability in the realization of Indonesian P in a modular grammar, in which morpho-phonological forms are determined in a derivational PF.

    Our starting point is that variability is conditioned in different ways for variable P-drop and for the alternation between *oleh~sama* or *dengan~sama*. First, we can distinguish factors that condition variability relating to linguistic structure, a kind of internal factor. The linguistic structure that acts as a pre-condition on the possibility of P-drop (Section 2) is an internal factor.

    Factors relating to style-shifting are a kind of external factor. The rate of realization of null vs. overt P is conditioned by factors relating to style (only in the structural environment laid out in Section 2); the rate of null P is lower, for example, in high registers.

    There is no structural 'pre-condition' on the realization of the colloquial form *sama*. Instead, its use is conditioned exclusively by context for style-shifting: *oleh* and *dengan* are strongly preferred in formal written Indonesian, while *sama* is an informal form, characteristic of low register or the colloquial variety. In short, use of *sama* is conditioned only by an external factor.

---

[7]Original glosses are preserved; indices and lower copies added for clarification.

[8]Kroeger 2014 discusses binding of the "short reflexive" in subject position, suggesting that the judgments arise from pragmatics rather than syntactic binding. Regardless of the analysis, the crucial point is the contrast between (26) vs. (27), which does not arise when P is null in (28).

[9]This judgment is from our consultants; it is not made explicit in Arka and Manning 2008.

We simplify discussion by excluding other factors, such as cognitive or sociolinguistic (demographic) factors for which we lack data (see also Section 4 for discussion). Our factors conditioning intra-speaker variability for present purposes, then, are as follows:

Initiator P		Internal factors	External factors
oleh 'by'	*dengan* 'with'	None	--
sama 'by/with'		None	Style, context of use
null (P-drop) 'by/with'		1. DP interpreted as Initiator 2. Passive prefix *di-* 3. Linear adjacency	Style, context of use

Table 1: Summary of conditions on distribution of Indonesian Initiator Preposition forms.

The factor 'style' *probabilistically* conditions realization of P: though the use of *oleh~sama* or *dengan~sama* may be stylistically odd in some context, (non-)realization of *sama* is never ungrammatical. As such we anticipate that the rate of realization of *sama* will co-vary with other informal variables. (The non-categorical nature of 'style' as a conditioning factor is clear in examples such as (3) and (11), where 'formal' *oleh*/*dengan* appears in the same clause as an informal variable, the familiar pronoun *aku*). In contrast, linguistic structure acts as a *categorical* condition on variable P-drop: P-drop is judged as grammatical or non-grammatical dependent on the structure in which it occurs, as shown in Section 2.

Both sets of factors are assumed to form part of an individual's linguistic knowledge, broadly conceived. However, we wish to maintain a modular grammar, where modular means separation between the core computational operations comprising a generative grammar, and a system of use (e.g., Chomsky 1965, Labov 1969, 2001); in particular, in a modular grammar, we assume that the two kinds of factor do not interact (see Mackenzie 2012 for discussion). Following standard assumptions on modularity, then, a factor such as linguistic structure may operate within the grammar proper, while style-shifting is external, part of a usage system. With this background, we examine the implementation of intra-speaker variability in Indonesian P in a derivational PF.

3.2 The Post-Syntactic Derivation: Preliminaries

As in Section 2.4, the surface variants of Indonesian prepositions must have the same syntactic/semantic input. The variable processes giving rise to the surface variants must therefore operate post-syntactically. We examine implementation of these processes in the Distributed Morphology (DM) framework (Halle and Marantz 1993, 1994, Harley and Noyer 1999 *et seq.*).

The variant Indonesian Ps are clearly functional morphemes, rather than lexical roots. In DM, functional morphemes are composed of bundles of features that lack phonological specification. The phonological content of a terminal node is supplied by Vocabulary insertion, where the feature bundles transferred from syntax are mapped to phonological strings. These phonological mappings are stored as Vocabulary items; items for *oleh* and *dengan* are represented as:[10]

(29) P {[agent], [experiencer], [causer]}, [+Arg] ↔ ɔle
     P [instrument], [+Arg] ↔ dəŋɑn

(29) states that P is realized as /ɔle/ in the context of the feature [+Arg(ument)] plus [agent], [experiencer] or [causer]. P is realized as /dəŋɑn/ in the context of the features [+Arg] and [instrument], such that *dengan* is inserted for both Instrument Causers and Pure Instruments.

(29) shows Vocabulary insertion does not and cannot give rise to variability, as it occurs

---

[10]The features of functional prepositions introducing initiators, such as in *by*-phrases, are not well understood cross-linguistically, and so we leave their precise identification to future research. Generally, we take the interpretation of arguments to arise configurationally, consistent with examples (9-16). The thematic features here are therefore for expository purposes only, and are not claimed to be atomic.

deterministically: without additional operations, the featural specification of a morpheme P will always result in /ɔle/ or /dəŋan/. We therefore introduce variability by alternative means.[11]

3.3 Generating Intra-speaker Variability Post-syntactically: Competition and Rules

We examine two possible means of generating intra-speaker variability post-syntactically for the two variants of Indonesian Ps:

    (i)   Competition: a form α or a form β may be used (Kroch 1989 i.a.)

    (ii) Variable rule: α may take the form β (in some environment) (Labov 1969 i.a.)

We show that Indonesian variant Ps reveal differences in implementing (i)-(ii) in the grammar: we propose (i) generates realization of *sama*, and (ii) generates P-drop. We conclude that both (i)-(ii) are required to generate intra-speaker variability in a modular and derivational PF architecture.

    'Competition for use' describes the possibility that a speaker can select which variant of a grammatical object to use, in contexts where more than one variant is possible/grammatical. 'Competition' describes the fact that the choice of one variant necessarily excludes the possible use of the other (Kroch 1989) (see Embick 2008 for detailed discussion of forms of competition).

    Competition for usage could be implemented by weighting each Vocabulary Item that is a member of a List by the probability of their use when the speaker adopts a particular style, such as informal/ informal. This clearly entails a departure from modularity.

    Instead, competition for use can be operationalized in DM over Lists (rather than members of a List) (Embick 2008). To illustrate for the *oleh~sama* and *dengan~sama* alternation: let there be more than one List of Vocabulary items that can be activated for a node P. List 1 (L_1) contains the items *oleh* and *dengan* (i.e., 29), while List 2 (L_2) contains only the item *sama* (30):

  (30) P {[+Arg], …} ↔ sɑmɑ

 The Vocabulary item in L_2 is underspecified so that it can be inserted in multiple P terminal nodes, 'neutralizing' the distinction between 'by' and 'with.' L_2 states that the phonological string /sɑmɑ/ is inserted if P is specified [+Arg].[12] L_1 (29) and L_2 (30) are internally invariant; once a List is selected for use, Vocabulary insertion proceeds deterministically, as before.

    Necessarily, only one List can be activated or used with a single terminal node. 'Competition' is the choice of use of L_1 or L_2, and the choice of use of a List may be conditioned by context of use. (Each List could be regarded as a 'grammar' G: L_1 is 'G_1' and L_2 is 'G_2.') Weighting Lists for probability of use means that no external factor interacts with the computation of the grammar; modularity between a system of use and grammatical competency is retained (Kroch 1989).

    Is variable P-drop also competition between Lists? Suppose null P is a stored item on a third List, L_3. L_3 would contain a vocabulary item (31) whose phonetic form is zero:

  (31) L_3: P [Initiator] ↔ Ø   (*To be rejected*)

    In a restricted linguistic structure (Table 1) a speaker would select between L_1, L_2, L_3. Choice of L_1, L_2, L_3 would be probabilistically conditioned by stylistic factors. Outside of this structure, L_3 would not be in competition, so that L_1 and L_2 are in a two-way competition for activation. As such the weighting of L_1 and L_2 would shift according to linguistic structure. Quantitative investigation would aid in determining whether this is correct.

    The formal issue of interest here is how to activate $List_3$ in the restricted linguistic structure defined in Section 2. Vocabulary insertion is strictly local, applying only to terminal nodes and the features that node bears. For example, competition for activation of L_1 and L_2 was triggered by the presence of a P node. In contrast, the linguistic condition on P-drop references structure that is non-local to the P node, the passive prefix *di-*. Architectural considerations therefore strongly

---

[11]Variability in P is not phonological, given the unrelatedness of the forms: *sama~oleh~dengan*.

[12]*Dengan* also occurs in manner adverbial PPs and comitative PPs, e.g., *dengan teman-ku* "with my friend." *Sama* alternates with *dengan* in both these environments, as well as Pure Instrument and Initiator Ps. Manner adverbials exhibit P-drop but comitative PPs do not. Due to reasons of space we cannot explore these domains, but these facts may shed light on the formal features at issue (cf. fn.9).

weigh against an analysis of P-drop in Indonesian as competition for activation of a List (cf. Tamminga et al. to appear, Embick 2015).

A second formal implementation of variability is via variable rule. Rules classically reference linguistic conditions on their application (e.g., Labov 1969 i.a.). To this end we propose a variable morpho-phonological deletion rule for P-drop:

(32) P deletion rule: $P \rightarrow \emptyset$ / [V$_{passive}$] $\frown$ [___ DP$_{Initiator}$]

(32) states the observations from Section 2, that when P introduces an argument interpreted as an Initiator, and PP is linearly adjacent to a passive verb complex, P deletes.[13] If (32) (variably) applies late in the derivation, following Vocabulary insertion, it can apply over the inserted forms *oleh, dengan,* or *sama*. Evidence for late application includes the linear adjacency requirement, which locates the application of (32) at least post-linearization. The relative ordering of linearization and Vocabulary insertion has not been independently established for Indonesian, but if Vocabulary Insertion precedes linearization, we have a neat account of the linear intervention data in (21-23). Late application therefore seems correct, but should be investigated further.[14]

Finally, P-drop is also sensitive to stylistic factors. We could say that (32) applies variably according to context of style, but this means an integration of style into the grammar. An alternative it is to say that (32) applies such that it derives two outputs (Form 1: P$_{OVERT}$, Form 2: P$_{NULL}$). For expository purposes, ands in the absence of quantitative data, the two outputs can be weighted at chance (p=0.50). We suggest that the weight of the two outputs is then modified by style-related conditions on use, which determines the overt variant (cf. Mackenzie 2012). Stylistic effects then only apply to output forms, separate from the grammar, and modularity is maintained.

Given a variable rule seems the appropriate characterization of variable P-drop in Indonesian, we might question whether the first formal means of generating variability that we discussed – competition for use – is also necessary. The answer seems to be that it is necessary for this data set. To see why, we start again with the derivation set up in Section 3.2. Without competition, there is a single List, so that the item *sama* (33) is a member of the same List as *dengan* and *oleh* in (29):

(33) P [+Arg] ↔ sama   (added to List (29); *To be rejected*)

As noted in Section 3.2, this single List will not capture variability, because in DM insertion is deterministic. As the single Vocabulary item realized as *sama* is underspecified for insertion in a wide range of P terminal nodes, it will never be realized: under the Subset Principle (Halle and Marantz 1994), the Vocabulary item matching the greater number of features will always be chosen over one with fewer matching features. As such, without further operations, we will have uniform insertion of *dengan* or *oleh* for every P node.

In order to prevent this, we can add an operation that derives different bundles of features, and applies variably, so that the features of P variably correspond to those of (29) and (33). This can be achieved by a (variable) Impoverishment rule. Impoverishment is a deletion operation that acts on feature bundles post-syntactically. Here, it would delete features on a terminal node P:

(34)   Variable P Impoverishment: P [θ] → [Ø] / {[informal, …]}   (*To be rejected*)

(34) states that its application deletes thematic ([θ]) 'features' such as [agent], [instrument] etc. on

---

[13]In (32) brackets show a PP and verb complex as adjacent ($\frown$) (notation from Embick 2007). The data require that some hierarchical structure be "visible" to the morphological component (post-linearization) to allow P to "see" the passive prefix *di-*; we leave this to future research. Embick (2015) suggests that such non-local conditions are often observed with feature deletion (Impoverishment), and not with allomorphy.

[14]If deletion applies "late", following Insertion, it cannot refer to the phonological string /ɔle/ or /dəŋɑn/: this would result in overapplication, since the form *dengan* undergoes P-drop with Instrument-causers, but the homophonous *dengan* that introduces Pure Instruments does not (Section 2.1). A consequence of late application, though, is that a variable process conditioned only by usage factors (*sama* realization) precedes another that applies categorically (P-drop). Quantitative investigation would aid determining whether this is correct. This conclusion does not breach our goal of modularity in the architecture laid out here.

P nodes. This would leave the specification of P as {[+Arg], …}. Following Impoverishment, *oleh* or *dengan* would therefore not be inserted, as they are specified for additional features, and *sama* would be inserted.

However, (34) states that this Impoverishment 'rule' or operation applies variably, according to [informal] contexts. Impoverishment, as an operation on linguistic structure, and its (non-) application otherwise results in grammatical or ungrammatical objects. Adding factors relating to a system of use ('informal style') to this grammatical operation therefore entails a departure from modularity. We therefore reject an analysis of *sama* as derived by a variable Impoverishment rule.

In sum, intra-speaker variability in the various realization of Indonesian P in part reflects a form of competition in use (variable *sama* realization), and in part rule(s) whose application results in deletion (P-drop). Both representations of variability are necessary to capture the full range of data, and both were integrated into a post-syntactic derivation. The formal representation of intra-speaker variability, and its locus in the PF derivation, was shown to depend in part on the nature of the factors conditioning variability, combined with consideration of a linguistic system that shows a modular relation between grammar and use.

4   Conclusions and Extensions

Although the three possible realizations of P (e.g., *oleh~sama~Ø*) are realizations of one morpheme in a particular syntactic position, close investigation suggests that *oleh, sama* and null P are not three 'variants' of one (morpho-)syntactic 'variable,' except in a superficial sense. Instead, surface 'optionality' masks multiple variable processes operating in the grammar, whose interaction may have non-trivial consequences for their respective rates of realization. In particular, realization of *sama* is conditioned by style, a condition on usage. We suggested *sama* is generated via competition for selection of Vocabulary Lists at PF (Embick 2008; cf. Kroch 1989). Section 2 showed that variable P-drop occurs only in a limited structural context; due to architectural considerations, we argued variable P-drop is generated by a morpho-phonological deletion rule.

Including only the factors of style and linguistic structure as conditioning variability (Table 1) is, we assume, a simplification; these are the factors that we have data for, and that we can clearly distinguish in a modular system. We anticipate that other factors will further mediate the rate of P-drop, and possibly realization of *sama*, and that these will include further internal factors (e.g., phonological weight, animacy, definiteness, etc. of DP) and external factors (demographic factors such as speaker's education, age, dialect, speech rate, etc.) Interestingly, there seems to be no prosodic motivation for P-drop: the overt prepositions *oleh, dengan* and *sama* do not otherwise undergo phonological reduction, even in fast speech. And although the weight of the DP may condition variable rates of P-drop, it will not exclude it; see (24-25). In sum, we anticipate that quantitative study will allow for a more nuanced picture.

Moving beyond intra-speaker variability to typological variation, several aspects of Indonesian P-drop are cross-linguistically unusual. Besides languages closely related to Indonesian (e.g., Madurese), we are not aware of other examples of P-drop where P introduces Initiator arguments: we believe all previous cross-linguistic work on P-drop examines *spatial* Ps.[15] In addition, where spatial P-drop has always been shown to be a reflex of alternations in the syntax or semantics (e.g., Ioannidou and den Dikken 2009, Terzi 2010, Gehrke and Lekakou 2013 on Greek; Collins 2007, Myler 2013 and Biggs 2015 on (varieties of) English; Nchare and Terzi 2014 on Shupamem), we have shown that Indonesian P-drop involves a uniform syntactic structure, with P-drop a reflex of the PF grammar. Other types of P may also be null in Indonesian, including spatial P, temporal P and manner adverbial P. We believe that extending the proposed analysis of Initiator P-drop to other types of P may profitably capture their distribution, and shed light on the questions raised here.

Finally, this study highlights the need for close analysis of surface variants, prior to quantitative investigation, in order to determine the structural distribution of the variants, the formal processes underlying variability, and their locus in a derivational architecture.

---

[15]Spatial PPs in Indonesian also allow P-drop, but for reasons of space data is not included here.

References

Arka, I. Wayan and Christopher Manning. 2008. Voice and grammatical relations in Indonesian: a new perspective. *Voice and Grammatical Relations in Austronesian Languages*. Stanford: CSLI Publications.

Biggs, Alison. 2015. A new Case for A-movement in Northwest British English. *Proceedings of the 32nd West Coast Conference on Formal Linguistics (WCCFL 32)*, 218-227.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chung, Sandra. 1976. On the subject of two passives in Indonesian. In *Subject and Topic,* ed. C. N. Li, 57-98. New York: Academic Press.

Cole, Peter, Gabriella Hermon and Yanti. 2008. Voice in Malay/Indonesian. *Lingua* 118:1500-1553.

Collins, Chris. 2007. Home sweet home. *NYU Working Papers in Linguistics* 1, 1-34.

Dardjowidjojo, Soenjono. 1978. *Sentence Patterns of Indonesian*. Honolulu: University of Hawaii Press.

Embick, David. 2007. Linearization and local dislocation: Derivational mechanics and interactions. *Linguistic Analysis* 33:303-336.

Embick, David. 2008. Variation and morphosyntactic theory: Competition fractionated. *Language and Linguistics Compass* 2: 59-78.

Embick, David. 2015. *The Morpheme: A Theoretical Introduction*. Berlin: Mouton de Gruyter.

Fortin, Catherine. 2007. Indonesian sluicing and the P-stranding generalization. In *Proceedings of the 14th Annual Meeting of the Austronesian Formal Linguistics Association*.

Gehrke, Berit and Marika Lekakou. 2013. How to miss your preposition. *Studies in Greek Studies* 33:92-106.

Guilfoyle, Eithne, Henrietta Hung and Lisa Travis. 1992. Spec of IP and Spec of VP: Two subject positions in Austronesian languages. *Natural Language & Linguistic Theory* 10:375-414.

Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger,* ed. K.L. Hale and S.J. Keyser, 111-176. Cambridge, MA: MIT Press.

Halle, Morris and Alec Marantz. 1994. Some key features of Distributed Morphology. In *MIT Working Papers in Linguistics* 21:275-288.

Harley, Heidi and Rolf Noyer. 1999. Distributed morphology. *Glot International* 4:3-9.

Ioannidou, Alexandra and Marcel Den Dikken. 2009. P-drop, D-drop, D-spread. In *Proceedings of the 2007 Workshop in Greek Syntax and Semantics*, ed. C. Halpert, J. Hartman, and D. Hill, 393-408. Cambridge, MA: MIT Press.

Kamp, Hans and Antje Rossdeutscher. 1994. Remarks on lexical structure and DRS construction. *Theoretical Linguistics* 20:97-164.

Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199-244.

Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45:715-762.

Labov, William. 2001. *Principles of Linguistic Change: Social Factors*. Malden, MA: Blackwell.

Lavandera, Beatriz R. 1978. Where does the sociolinguistic variable stop? *Language in Society* 7:171-182.

Macdonald, R. Ross and Soenjono Darjowidjojo. 1967. *Indonesian Reference Grammar*. Washington, DC.: Georgetown University Press.

Mackenzie, Laurel. 2012. Locating Variation above the Phonology. Doctoral dissertation, UPenn.

Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60:847-894.

Myhill, John. 1988. Nominal agent incorporation in Indonesian. *Journal of Linguistics* 24:111-136.

Myler, Neil. 2013. On coming the pub in the North West of England: Accusative unaccusatives, dependent case, and preposition incorporation. *The Journal of Comparative Germanic Linguistics* 16:189-207.

Nchare, Abdoulaye Laziz and Arhonto Terzi. 2014. Licensing silent structure: The spatial prepositions of Shupamem. *Natural Language & Linguistic Theory* 32:673-710.

Newmeyer, Frederick J. 2003. Grammar is grammar and usage is usage. *Language* 79:682-707.

Nomoto, Hiroki and Kartini Abd Wahab. 2014. Person restriction on passive agents in Malay: Information structure and syntax. In *NUSA: Linguistic Studies of Languages in and around Indonesia* 57:31-50.

Sneddon, James Neil, Alexander Adelaar, Dwi Noverini Djenar and Michael C. Ewing. 2012. *Indonesian Reference Grammar*. New York: Routledge.

Tamminga, Meredith, Laurel Mackenzie, and David Embick. To appear. The dynamics of variation in individuals. *Linguistic Variation.*

Terzi, Arhonto. 2010. On null spatial Ps and their arguments. *Catalan Journal of Linguistics* 9:167-187.

Department of Linguistics,
University of Pennsylvania
Philadelphia, PA 19104–6305
jeoungh@sas.upenn.edu
biggsa@ling.upenn.edu