

Bahasa expletive =*nya* and the height of agents

Background. Bahasa (Malay, Indonesian) has an apparent expletive subject (Expl), =*nya*, which encliticizes onto lexical verbs; formally it is a third-person pronoun. Constructions with Expl are thetic, meaning they are comments on an event, not on a topical subject (cf. Lonzi 1986, Sasse 1987); this is shown by the contrast between (1b) and (2b):

- | | |
|--|--|
| (1) Q: What happened at the concert? | (2) Q: What did Ali do at the concert? |
| a. <i>Ali naik ke atas pentas.</i>
A. climb.up to on stage | a. <i>Ali naik ke atas pentas.</i>
A. climb.up to on stage |
| b. <i>Naik=nya Ali ke atas pentas.</i>
climb.up=nya A. to on stage
'Ali climbed on the stage.' | b. <i>*Naik=nya Ali ke atas pentas.</i>
climb.up=nya Ali to on stage
'Ali climbed on the stage.' |

Note that both thetic and categorical sentences can answer the question 'What happened?' (Lonzi 1986, Lambrecht 1987), so it is not a surprise that (1) has two possible answers. Note that the facts that =*nya* participates in thetic constructions, and that it is a third-person pronoun, make it typologically likely that it is an expletive (cf. French presentational *il*, etc.).

I argue that Expl's distribution demonstrates that Bahasa's transitive agents are higher than its unergative agents. The discussion is based on a number of recent findings in the EPP literature (see Richards & Biberauer 2005, Richards 2007, Deal 2009) claiming that Expl is not base-generated in Spec-T, but rather in the highest Spec of a phase: either Spec-*v* or Spec-C, depending on the language and Expl. This view is meant to derive, for instance, that English Expl can't co-occur with material in Spec-*v*:

- | | |
|--|--|
| (3) <i>*There laughed many children.</i> | (4) <i>*There had some friends a picnic.</i> |
|--|--|

Expl occurs with unergatives, but not transitives. Bahasa accepts thetic sentences with both unaccusatives (5) and unergatives (6).

- | | |
|--|---|
| (5) <i>Datang=nya se=ekor orangutan.</i>
come=3SG one=CL orangutan.
'An orangutan came.' | (6) <i>Jerit=nya orangutan-orangutan ni.</i>
shout=nya orangutan.RED this
'These orangutans shouted.' |
|--|---|

The above verbs can be shown to be unergative or unaccusative in Bahasa based on quantifier float:

- | | |
|---|--|
| (7) a. <i>(Semua) anak (semua) sudah (?semua) datang (semua)</i>
(all) child (all) PERF (?all) come (all)
'All the children have come.' | b. <i>(Semua) anak (semua) sudah (?semua) jerit (*semua)</i>
(all) child (all) PERF (?all) shout (*all)
'The children have shouted.' |
|---|--|

The contrast in postverbal quantifier float shows subjects originate below the verb in unaccusatives, but originate higher than the verb in unergatives.

While unaccusatives and unergatives can co-occur with Expl, transitives can't (Sommerlot 2018). To be sure, thetic transitives are marginally possible, but they need comma-intonation (8a) or dislocation (8b); and even with these fixes, the resulting phrases are degraded.

- | | |
|---|---|
| (8) a. <i>Letak=nya laki, buku ke atas meja.</i>
place=nya man, book to on table
'A man put a book on the table.' | b. <i>Makan=nya ais krim laki tu.</i>
eat=nya ice cream man that
'That man ate some ice cream.' |
|---|---|

Such sentences have equivalents in English presentationals (which also generally bar transitives):

- (9) There entered the room a strange man. (Chomsky 2001:21)

As such, to the extent that English bars transitive Expl constructions, so does Bahasa.

Analysis. Richards & Biberauer (2005) find that Expl is always generated in the highest Spec of a phase, rather than Spec-T (as usually assumed in the EPP literature). In languages that only allow Expl with unaccusatives (like English), this is Spec-*v*. In languages (like Icelandic) that allow Expl with all verb types, this is Spec-C. As a result, either Expl is in complementary distribution with the Agent/Experiencer, or it isn't.

Bahasa provides a new piece of empirical evidence for Richards & Biberauer's analysis: specifically, Bahasa actually has Expl surface in the verbal domain, not in Spec-T. Indeed, Bahasa verbs are lower than T, as shown by the adverb-verb order in (10), and *=nya* always cliticizes on the verb, never an auxiliary, as shown in (11).

- (10) *Aku tak akan **sentiasa makan** nasi.* (11) a. *Sudah sampai=nya budak.*
 I not will always eat rice PERF arrive=nya child
 'I will not often eat rice.' 'A child arrived.'
 b. **Sudah=nya sampai budak.*
 c. **Sampai=nya sudah budak.*

This is strong evidence against an analysis of Expl as (universally) resulting from a property of T.

But despite apparently providing new and clear evidence for Richards & Biberauer's theory, Bahasa is unlike other languages in allowing Expl with unergatives but not transitives. Recall that according to Richards & Biberauer's theory, *=nya* would be merged either in Spec-C (in which case it should allow transitives) or in Spec-*v* (in which case it should bar unergatives).

Putting aside complications from cliticization, I argue the PERSON features (Chomsky 2001) of the Expl are base-generated in Spec-Voice in Malay. I follow standard assumptions in taking Voice to be above *v*, and in taking it to be a phase head – meaning that *=nya* is indeed generated in the highest Spec of a phase. I only depart from standard assumptions in the following way: I argue the agents of transitives are generated in (or must move to) Spec-Voice, whereas those of unergatives are generated in (or, can stay in) Spec-*v*:

- (12) [_{VoiceP} (Agent-Tr) [_{Voice}⁰ [_{vP} (Agent-Unerg) [_v⁰ [VP]]]]]

Combined, these hypotheses make it possible to follow Richards & Biberauer (2005) in accounting for the distribution of Expl structurally: Expl and transitives' agents compete for the same Spec, namely Spec-Voice; but Expl and unergatives' agents are in different Specs (Spec-Voice and Spec-*v* respectively) and can therefore co-occur:

- (13) [_{VoiceP} Expl [_{Voice}⁰ [_{vP} Agent-Unerg [_v⁰ [VP]]]]]

Higher agents in transitives: a general property? Based on data from causatives, Massam (2009) argues that the agents of Niuean transitives are higher than those of unergatives. I have just shown that the Expl data from Bahasa provide a new argument for the same thesis. The next logical step is therefore to speculate this might be a broader property of the world's languages, which I argue can be derived from Harley's (2017) argument that Voice and *v* are bundled in some languages but not others. The agents of unergatives and transitives only occur in the same Spec in languages (e.g. English) where Voice and *v* are bundled. But as Bahasa and Niuean show, in languages where Voice and *v* are disjoint, the difference in agent height becomes visible.

References. Chomsky, N. 2001. Derivation by Phase. • Deal, A. R. 2009. The Origin and Content of Expletives: Evidence from "Selection". • Harley, H. 2017. The 'bundling' hypothesis and the disparate function of little *v*. • Lambrecht, K. 1987. Aboutness as a Cognitive Category. • Lonzi, L. 1986. Pertinenza della struttura Tema–Rema per l'analisi sintattica. • Massam, D. 2009. The Structure of (Un)Ergatives. • Richards, M. 2007. Object Shift, Phases, and Transitive Expletive Constructions in Germanic. • Richards, M., & T. Biberauer. 2005. Explaining Expl. • Sasse, H.-J. 1987. The Thetic/Categorical Distinction Revisited. • Sommerlot, C. J. 2018. A Presentational Construction in Indonesian.