



Abstrak sudah dikirim pada **Korpus Linguistik** dengan nomor booking **Kor965**

Bidang:

Korpus Linguistik

Penulis:

Gede Primahadi Wijaya Rajeg, Karlina Denistia

Judul:

Distinctive Collexeme Analysis of Indonesian Causative Rival Affixes per- and -kan

Abstak:

Indonesian has a pair of rival affixes, namely *per-* and *-kan*, to derive semantically causative transitive verbs from adjective (ADJ) bases. While the deadjectival verbs with *per-* and *-kan* similarly express causative meaning, Sneddon et al. (2010, p. 103) suggests nuances between formatives of the same base: ADJ-*kan* implies that the direct objects are caused to possess ADJ property they previously lack (e.g. *besar* ‘big’ - *besarkan baju* ‘make shirt big’), meanwhile *per-*ADJ increases the ADJ property of the direct objects (e.g. *besar* ‘big’ - *perbesar baju* ‘make shirt bigger’). Besides, Sneddon et al. (2010, p. 103) mentions formal and distributional constraints on the affixes regarding the ADJ bases they occur with: “[w]ith some bases one affix or the other does not occur or is very infrequent”.

This paper addresses the aforementioned constraints on *per*-ADJ and ADJ-*kan* using *Distinctive Collexeme Analysis* (DCA) (Gries & Stefanowitsch, 2004). DCA contrasts two semantically/functionally similar grammatical constructions in terms of lexemes showing statistical preference for one member of the contrasted constructions. We extend DCA into Indonesian rival affixes to investigate which bases strongly prefer, or only occur in, one of the affixes. From the *Indonesian Leipzig Corpora* (Goldhahn et al., 2012) (180,769,204 word-tokens), we retrieved the candidate deadjectival verbs following *memper-*, *diper-*, *meN-/-kan*, and *di-/-kan* patterns. The verbs were then morphologically parsed using *MorphInd* (Larasati et al., 2011) before extracting the adjective-tagged bases. We used R programming language (R Core Team, 2020) for data pre-processing and statistical analyses.

Table 1. Top-20 distinctive bases from the outputs of *Distinctive Collexeme Analysis* (<https://tinyurl.com/gpwr-kd-kolita19>)

Table 1 demonstrates that the distribution of the bases is asymmetric across affixes. The distinctive bases for *per*-ADJ can also occur with ADJ-*kan* (Table 1a) whereas many distinctive bases for ADJ-*kan* are significantly absent from *per*-ADJ (cf. the zeros in *N_{PER-}* column in Table 1b), pointing to (i) negative evidence (Stefanowitsch, 2008) for the bases to occur in *per*-ADJ; (ii) the entrenchment of the morphologically complex words with those bases in ADJ-*kan*; and (iii) the extent to which ADJ-*kan* and *per*-ADJ differ in their formal and distributional constraints (e.g. *-kan* attaches to more types than *per-* does).

Given the asymmetric distribution of the bases as revealed by DCA, we further investigated if the distinctive bases might have to do with gradability by comparing the total token frequencies of the distinctive bases in (morphological and periphrastic) Degree Constructions, expressing the equative, comparative and superlative functions. Figure 1 below shows that the total frequency of the distinctive adjectives for *per*-ADJ is significantly greater in the equative, comparative, and periphrastic superlative constructions (i.e., *paling* ADJ). Meanwhile, the total frequency of ADJ-*kan*’s distinctive adjectives is only higher in the morphological superlative constructions (i.e., *ter*-ADJ). Overall, the distinctive bases for *per*-ADJ are predominantly linked to degree constructions and gradable semantics as opposed to ADJ-*kan*, giving some support to the textbook characterisation of *per*-ADJ with (increased) gradability of the direct objects’ ADJ property.

Figure 1. Summed token frequency of distinctive adjective bases for ADJ-*kan* and *per*-ADJ in Degree Constructions (<https://tinyurl.com/gpwr-kd-kolita19>)

Kata Kunci:

Distinctive Collexeme Analysis, Construction Morphology, Indonesian Morphological Causatives, Rival Affixes, Quantitative Corpus Linguistics

Daftar Pustaka:

Goldhahn, D., Eckart, T., & Quasthoff, U. (2012). Building large monolingual dictionaries at the Leipzig Corpora Collection: From 100 to 200 languages. *Proceedings of the 8th Language Resources and Evaluation Conference (LREC) 2012*, 759–765. http://www.lrec-conf.org/proceedings/lrec2012/pdf/327_Paper.pdf

Gries, S. Th., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective on “alternations.” *International Journal of Corpus Linguistics*, 9(1), 97–129.

Larasati, S. D., Kubo?, V., & Zeman, D. (2011). Indonesian Morphology Tool (MorphInd): Towards an Indonesian Corpus. *Systems and Frameworks for Computational Morphology*, 119–129. https://doi.org/10.1007/978-3-642-23138-4_8

R Core Team. (2020). *R: A language and environment for statistical computing* [Manual]. <https://www.R-project.org/>

Sneddon, J. N., Adelaar, A., Djenar, D. N., & Ewing, M. C. (2010). *Indonesian reference grammar* (2nd ed.). Allen & Unwin.

Stefanowitsch, A. (2008). Negative entrenchment: A usage-based approach to negative evidence. *Cognitive Linguistics*, 19(3), 513–531. <https://doi.org/10.1515/COGL.2008.020>