Analisis Koleksem Khas dan potensinya untuk kajian sinonim konstruksional di Bahasa Indonesia

Gede Primahadi Wijaya Rajeg & I Made Rajeg

# Pengantar

This is a first paragraph.

Body texts go here.

# Data dan metodologi

## Berkas korpus dan pencarian kolokasi

Makalah ini menggunakan lima berkas korpus (lihat [Tabel 1](#corpus-text-size)) yang menjadi bagian dari Korpus Leipzig Bahasa Indonesia (Biemann, Heyer, Quasthoff, & Richter, [2007](#ref-biemann_leipzig_2007); Quasthoff & Goldhahn, [2013](#ref-quasthoff_indonesian_2013)). Jumlah keseluruhan kata untuk kelima korpus ini adalah sekitar 29,343,544 juta kata.

Tabel 1 Berkas korpus Leipzig yang digunakan beserta ukurannya.

|  |  |
| --- | --- |
| Berkas korpus | Ukuran (token kata) |
| ind\_news\_2008\_300K-sentences.txt | 5,875,376 |
| ind\_news\_2009\_300K-sentences.txt | 5,868,276 |
| ind\_news\_2010\_300K-sentences.txt | 5,874,158 |
| ind\_news\_2011\_300K-sentences.txt | 5,852,211 |
| ind\_news\_2012\_300K-sentences.txt | 5,873,523 |

Dengan menggunakan fungsi colloc\_leipzig(), yang terdapat dalam modul (*package*) R collogetr (G. P. W. Rajeg, [2018a](#ref-rajeg_collogetr_2018)), pencarian kolokasi *membesarkan* dan *memperbesar* dibatasi pada lingkup satu kata di sebelah kanan kedua verba tersebut (atau disebut dengan kolokat R1). Pilihan ini dilandasi atas dua asumsi. Pertama, satu kata yang tepat mengikuti verba transitif tersebut diasumsikan sebagai objek langsungnya. Kedua, objek langsung yang secara kuantitatif (§[2.2](#dcaintro)) khas mengikuti verba tersebut dipandang sebagai salah satu cara mencirikan perbedaan semantis kedua verba tersebut.

## Pengenalan ringkas terhadap Analisis Koleksem Khas

Analisis Koleksem Khas (atau *Distinctive Collexeme Analysis*; selanjutnya disingkat DCA) (Gries & Stefanowitsch, [2004](#ref-gries_extending_2004)) merupakan salah satu ragam dari gugusan metode linguistik korpus kuantitatif yang disebut Analisis Kolostruksional (*Collostructional Analysis*[[1]](#footnote-1), selanjutnya CA) (Hilpert, [2014](#ref-glynn_collostructional_2014); Stefanowitsch, [2013](#ref-hoffmann_collostructional_2013); Stefanowitsch & Gries, [2003](#ref-stefanowitsch_collostructions_2003), [2009](#ref-stefanowitsch_corpora_2009)). DCA dikembangkan untuk mengkaji perbedaan semantis dua konstruksi, yang mirip secara semantis/fungsional, berdasarkan unsur leksikal (disebut **koleksem**) yang memiliki keterkaitan khas secara statistik terhadap satu dari kedua konstruksi tersebut. Sebagai contoh, DCA dapat digunakan membedakan konstruksi Datif Preposisional dengan Ditransitif; konstruksi *Will + infinitive* dengan *be going to + infinitive*; konstruksi Aktif dengan Pasif (periksa Gries & Stefanowitsch, [2004](#ref-gries_extending_2004) untuk ulasan mendalam).

Makalah ini bertujuan melihat potensi DCA untuk menemukan perbedaan pemakaian sepasang kata pencontohan dua konstruksi kausatif morfologis pada BI, yaitu [*AJEKTIVA + -kan*] dan [*per- + AJEKTIVA*], yang perbedaan maknanya tidak terlalu diperhatikan oleh kebanyakan penutur BI (periksa §[1](#intro)). Analisis statistik dalam DCA didasari atas tabulasi kekerapan dua dimensi seperti pada [Tabel 2](#xtab)] yang mencontohkan distribusi nomina *jumlah* sebagai kolokat R1 untuk *memperbesar* dan *membesarkan* dalam kelima berkas korpus pada [Tabel 1](#corpus-text-size).

Tabel 2 Tabulasi silang sebagai masukan DCA

|  |  |  |  |
| --- | --- | --- | --- |
|  | *memperbesar* | *membesarkan* | Total |
| *jumlah* | 9 (5.77) | 1 (4.23) | 10 |
| *¬jumlah* | 408 (411.23) | 305 (301.77) | 713 |
| Total | 417 | 306 | 723 |

Post table text is here.

R codes below.

# buat matrix tabulasi silang  
xtab <- matrix(c(9, 1, 408, 305), nrow = 2, byrow = TRUE)  
  
# lakukan fisher exact test dan ambil p-value  
p\_fye <- fisher.test(xtab, alternative = "greater")$p.value  
  
# periksa hasil  
p\_fye

## [1] 0.03306132

# ubah p-value menjadi kadar kekhasan  
collstr <- -log10(p\_fye)  
round(collstr, digits = 2)

## [1] 1.48

DCA sejauh ini telah diterapkan pada data Bahasa Indonesia untuk kajian interaksi antara metafora dan sinonim emosi (G. P. W. Rajeg, [2014](#ref-rajeg_metaphorical_2014), [2016](#ref-rajeg_collostructional_2016), [2018b](#ref-rajeg_happyr_2018)) dan konstruksi negasi dengan *tak* dan *tidak* (G. P. W. Rajeg, Denistia, & Rajeg, [2018](#ref-rajeg_working_2018)).

# Hasil dan pembahasan

Results discussion goes here.

Tabel 3 Koleksem R1 khas untuk *membesarkan* (*collstr* > 1.30103 atau *p*Fisher Exact < 0.05)

Post table text goes here.

Tabel 4 Koleksem R1 khas untuk *memperbesar* (*collstr* > 1.30103 atau *p*Fisher Exact < 0.05)

# Penutup

# Daftar pustaka

Biemann, C., Heyer, G., Quasthoff, U., & Richter, M. (2007). The Leipzig Corpora Collection: Monolingual corpora of standard size. In M. Davies, P. Rayson, S. Hunston, & P. Danielsson (Eds.), *Proceedings of the Corpus Linguistics Conference*. University of Birmingham, UK. Retrieved from <http://ucrel.lancs.ac.uk/publications/CL2007/paper/190_Paper.pdf>

Gries, S. T., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective on ’alternations’. *International Journal of Corpus Linguistics*, *9*(1), 97–129.

Hilpert, M. (2014). Collostructional analysis: Measuring associations between constructions and lexical elements. In D. Glynn & J. A. Robinson (Eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy* (pp. 391–404). Amsterdam: John Benjamins Publishing Company.

Quasthoff, U., & Goldhahn, D. (2013). *Indonesian corpora* (Technical report series on corpus building No. 7). Leipzig, Germany: Abteilung Automatische Sprachverarbeitung, Institut für Informatik, Universität Leipzig. Retrieved from <http://asvdoku.informatik.uni-leipzig.de/corpora/data/uploads/corpus-building-vol7-ind.pdf>

Rajeg, G. P. W. (2014). Metaphorical profiles of five Indonesian quasi-synonyms of ANGER: Multiple distinctive collexeme analysis. In *Proceedings of the International Congress of the Linguistic Society of Indonesia 2014* (pp. 165–170). Bandar Lampung, Sumatra, Indonesia: Masyarakat Linguistik Indonesia (MLI). doi:[10.4225/03/58578ddba1fd2](https://doi.org/10.4225/03/58578ddba1fd2)

Rajeg, G. P. W. (2016, July). *Collostructional analysis meets metaphors: A study on Indonesian near-synonyms of HAPPINESS*. Presented at the Twentieth International Symposium on Malay/Indonesian Linguistics (ISMIL 20), The University of Melbourne, Australia. doi:[10.4225/03/592e2ba2e6ea7](https://doi.org/10.4225/03/592e2ba2e6ea7)

Rajeg, G. P. W. (2018a). Collogetr: Collocates retriever and Collocation association measure (Version 1.0.2). doi:[10.26180/5b7b9c5e32779](https://doi.org/10.26180/5b7b9c5e32779)

Rajeg, G. P. W. (2018b). Happyr: The accompanying R package for Rajeg’s (2018) PhD thesis titled “Metaphorical profiles and near-synonyms: A corpus-based study of Indonesian words for Happiness” (Version 0.1.0). doi:[10.5281/zenodo.1436331](https://doi.org/10.5281/zenodo.1436331)

Rajeg, G. P. W., Denistia, K., & Rajeg, I. M. (2018). Working with a linguistic corpus using R: An introductory note with Indonesian negating construction. *Linguistik Indonesia*, *36*(1), 1–36. doi:[10.4225/03/5a7ee2ac84303](https://doi.org/10.4225/03/5a7ee2ac84303)

Stefanowitsch, A. (2013). Collostructional analysis. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford handbook of Construction Grammar* (pp. 290–306). Oxford: Oxford University Press. doi:[10.1093/oxfordhb/9780195396683.013.0016](https://doi.org/10.1093/oxfordhb/9780195396683.013.0016)

Stefanowitsch, A., & Gries, S. T. (2003). Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics*, *8*(2), 209–243.

Stefanowitsch, A., & Gries, S. T. (2009). Corpora and grammar. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics: An international handbook* (Vol. 2, pp. 933–951). Berlin: Mouton de Gruyter.

1. Unsur *collostruction* pada *Collostructional Analysis* (CA) adalah perpaduan antara kata *collocation* dan *construction* (Stefanowitsch & Gries, [2003](#ref-stefanowitsch_collostructions_2003)). CA sejauh ini merupakan salah satu metode kuantitatif yang lazim digunakan dalam kajian linguistik yang dilandasi atas Pendekatan Konstruksional, seperti Tatabahasa Konstruksi (*Construction Grammar*) (Stefanowitsch, [2013](#ref-hoffmann_collostructional_2013)). [↑](#footnote-ref-1)