Rajeg, G. P. W., & Denistia, K. (2021, May 21). *A study in productivity of Indonesian causative* per- *and* -kan [Paper]. The 24th International Symposium on Malay/Indonesian Linguistics (ISMIL 24), Online. https://doi.org/10.6084/m9.figshare.14633133

A study in productivity of Indonesian causative *per-* and *-kan*

Gede Primahadi Wijaya RAJEG & Karlina DENISTIA
Universitas Udayana & Universitas Gadjah Mada
primahadi_wijaya@unud.ac.id & karlinadenistia@ugm.ac.id

21 May 2021
The 24th International Symposium on Malay/Indonesian Linguistics (ISMIL 24)

Project update:
http://bit.ly/per-kan-idn

Outline

- Overview of *per-* and *-kan*
  - Why studying productivity of *per-* and *-kan*
- Materials and methods
- Results
  - Realised Productivity
  - Potential Productivity
- Conclusion
  - lingering questions

Overview of causative *per-* vs. *-kan*

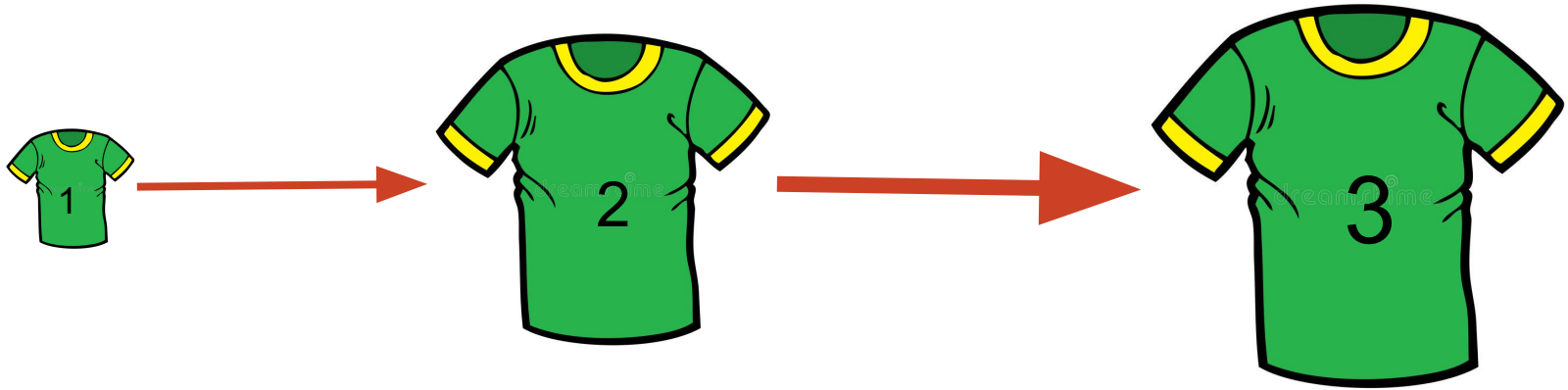
- *-kan* and *per-* are rival affixes (cf. Szymanek 2005: 438; Aronoff & Lindsay 2014: 243; Lindsay 2012)

Overview of causative *per-* vs. *-kan*

- *-kan* and *per-* are rival affixes (cf. Szymanek 2005: 438; Aronoff & Lindsay 2014: 243; Lindsay 2012)
- *-kan* causes the direct object to have the ADJ quality from a non-existence characteristic (Sneddon et al. 2010: 103)
- *per-* increases the object's ADJ quality (Sneddon et al. 2010: 103)

Overview of causative *per-* vs. *-kan*

- *-kan* and *per-* are rival affixes (cf. Szymanek 2005: 438; Aronoff & Lindsay 2014: 243; Lindsay 2012)
- *-kan* causes the direct object to have the ADJ quality from a non-existence characteristic (Sneddon et al. 2010: 103)
- *per-* increases the object's ADJ quality (Sneddon et al. 2010: 103)



*besar**kan** baju* '**make** shirt big'

***per**besar* *baju* '**make** shirt bigg**er**'

Overview of causative *per-* vs. *-kan*

- *per-* is historically older
  - Reflex of PAN causative prefix **pa-* (Adelaar 1992: 49; Ogloblin 1998: 182)
  - *-kan* is thought to develop from preposition *akan* ‘towards’ (Ogloblin 1998: 182)

Overview of causative *per-* vs. *-kan*

- *per-* is historically older
  - Reflex of PAN causative prefix **pa*- (Adelaar 1992: 49; Ogloblin 1998: 182)
  - *-kan* is thought to develop from preposition *akan* 'towards' (Ogloblin 1998: 182)
- *per-* vs. *-kan* is less studied
  - Semantic ambiguity/generalisation of *per-* and *-kan*
    - "many people making no distinction between the function of the two affixes" (Sneddon et al. 2010: 103)
  - Extensive works on *-kan* (esp. in comparison to *-i*)

Why studying productivity of *per-* and *-kan*

- Hint for the semantic ambiguity by “many people” (Sneddon et al. 2010: 103)
- Attempt to explain the results using the notion of **entrenchment** from the usage-based, cognitive linguistic perspective (cf. Langacker 1988; Hilpert & Diessel 2016; Stefanowitsch & Flach 2016)
- Rate of growth of the affixes and their extent of “saturation” (Baayen 2009)
  - Potential productivity of the affixes
- Attempt to generate hypothesis for further experimental study
  - e.g., lexical processing and storage

Materials and methods

- ★ Indonesian Leipzig Corpora (13 files; >180M word tokens) (Goldhahn, Eckart & Quasthoff 2012)
  - Regular expressions with *meN-*, *di-*, unprefixed
- ★ *MorphInd* (Indonesian morphological parser) (Larasati, Kuboň & Zeman 2011)
- ★ *MALINDO Morph* (Nomoto, Choi, Moeljadi & Bond 2018)
- ★ *Kamus Besar Bahasa Indonesia* (KBBI) (https://kbbi.kemdikbud.go.id/)
- ★ R (R Core Team 2020) and the `tidyverse` R package (Wickham et al 2019)

Snippet of the database

Word	Word Translation	Frequency	Affix	Base	Base Translation
birukan	make something blue	9	*-kan*	*biru*	blue
bangkrutkan	make something bankrupt	39	*-kan*	*bangkrut*	bankrupt
empukkan	make something soft; tender	12	*-kan*	*empuk*	soft; tender
peranyar	make something newer	2	*per-*	*anyar*	new
perburuk	make something worse	1,232	*per-*	*buruk*	bad
peraktif	make something more active	1	*per-*	*aktif*	active

Results: REALISED PRODUCTIVITY

Analysis on token frequency, types, and hapaxes

Results: REALISED PRODUCTIVITY

Rank	Token Freq	Word	Translation
1	96,957	*nyata**kan***	make something real
821	1	*zalim**kan***	make someone feel hurt

-kan has 820,370 tokens; 821 types; 173 hapaxes

Results: REALISED PRODUCTIVITY

Rank	Token Freq	Word	Translation
1	96,957	*nyata**kan***	make something real
821	1	*zalim**kan***	make someone feel hurt

-kan has 820,370 tokens; 821 types; 173 hapaxes

Rank	Token Freq	Word	Translation
1	18,957	***per**kuat*	make something stronger
169	1	***per**yakin*	make something more certain

per- has 78,896 tokens; 169 types; 50 hapaxes

Results: REALISED PRODUCTIVITY

Affix	Total tokens	Total types	Total hapaxes
-kan	820,370	821	173
per-	78,896	169	50

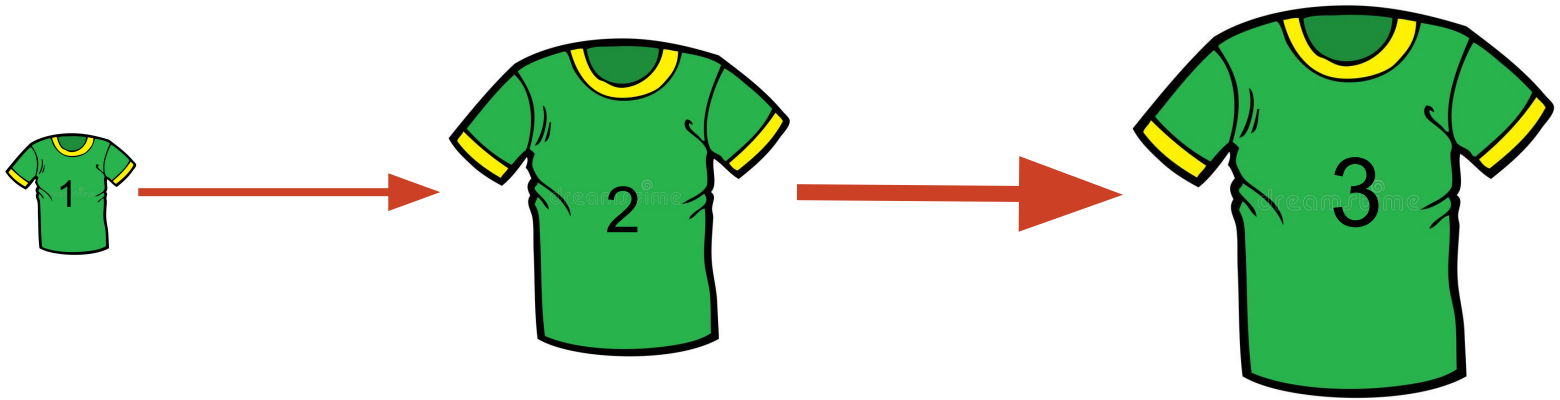
-kan is more productive than *per-*

-kan is more productive than *per-*

- *-kan* is more entrenched as causative construction (cf. Stefanowitsch & Flach 2016; Hilpert & Diessel 2016)

-kan is more productive than *per-*

- *-kan* is more entrenched as causative construction (cf. Stefanowitsch & Flach 2016; Hilpert & Diessel 2016)
- *-kan* causes the direct object not only to have the ADJ quality from a non-existence characteristic, but also to increase the object's ADJ quality



*besar**kan** baju* '**make** shirt bigg**er**'

Results: POTENTIAL PRODUCTIVITY

Analysis on the ratio of hapaxes over tokens for the vocabulary growth rate of a morphological pattern

(cf. Baayen 2009: 902-905)

Results: POTENTIAL PRODUCTIVITY

per- has higher potential productivity than *-kan*

Affix	**Token Freq**	Type Freq	**Hapax**	**HTR**
per+ADJ	**78,896**	169	**50**	**0.0634**
ADJ+*kan*	**820,370**	821	**173**	**0.0211**

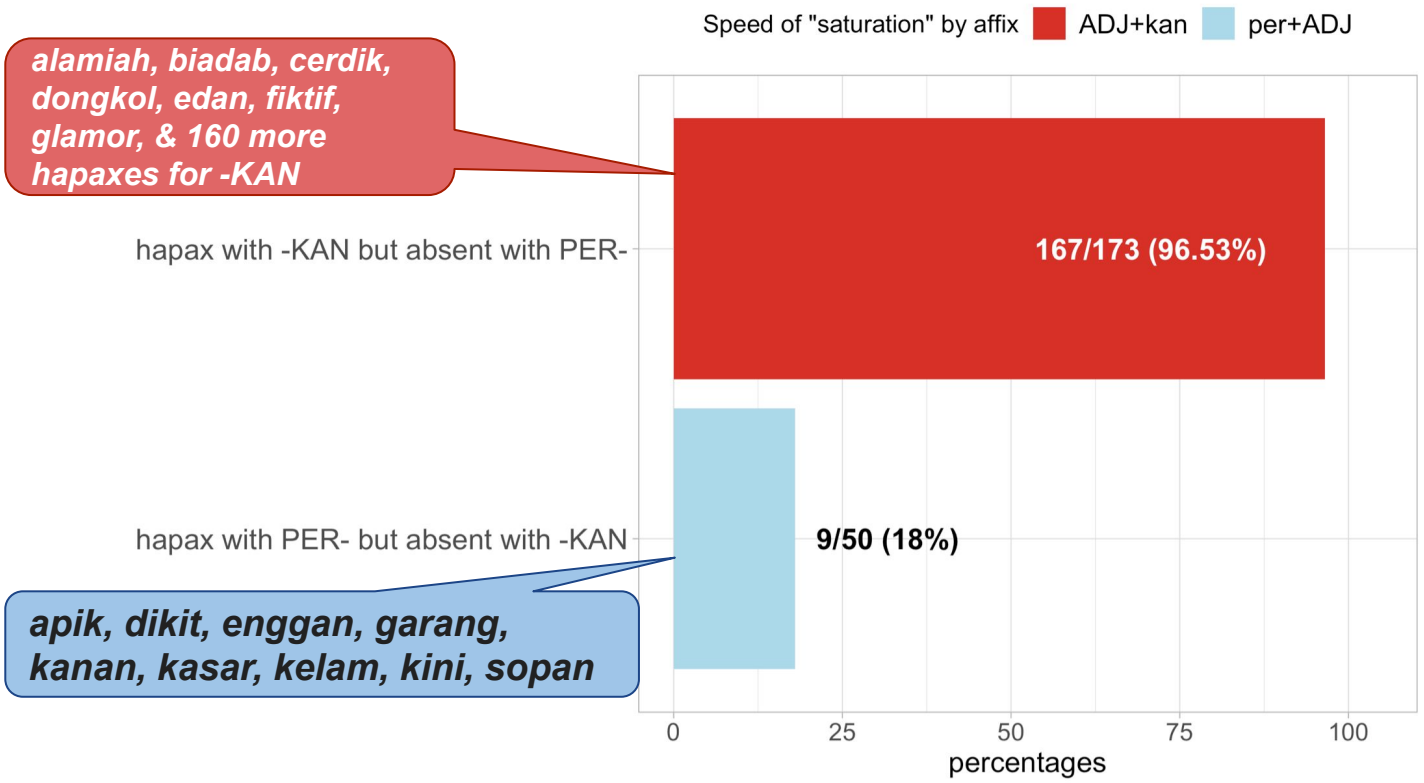
Results: POTENTIAL PRODUCTIVITY

-- Quantifying the saturation rate of *per-* and *-kan* --

1. Proportion of ADJ hapax with *-kan* but unattested (i.e., zero token) with *per-*

2. Proportion of ADJ hapax with *per-* but unattested (i.e., zero token) with *-kan*

Rate of “saturation” by affix



96.53% of thet total number of ADJ hapax with -KAN are unattested with PER-.
Only 18% of the total number of ADJ hapax with PER- that are unattested with -KAN.

Conclusions

- ADJ+*kan* is more productive than *per*+ADJ in usage (**realised productivity**)
  (cf. Ogloblin 1998: 179)
  - ADJ+*kan* as a morphological form-meaning pairing is potentially more entrenched
  - One motivation for crippling and levelling the semantics/function of *per*+ADJ
- *per*+ADJ shows higher **potential productivity**
  - High extent of “saturation” for ADJ+*kan*’s “vocabulary/onomasiological market”
- Future works:
  - Is it really (and to what extent) that many people do not recognise the semantic/functional difference between the affixes?
  - Distributional constraints of the bases with two affixes (cf. Sneddon et al. 2010: 103) (our KOLITA 19 plenary talk on 15th July 2021)

References

- Aronoff, Mark & Mark Lindsay. 2014. Productivity, Blocking, and Lexicalization. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Derivational Morphology*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199641642.013.0005 (16 August, 2017).
- Baayen, R. Harald. 2009. Corpus linguistics in morphology: Morphological productivity. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics: An international handbook*, vol. 2, 899–919. Berlin: Mouton de Gruyter.
- Goldhahn, Dirk, Thomas Eckart & Uwe Quasthoff. 2012. Building large monolingual dictionaries at the Leipzig Corpora Collection: From 100 to 200 languages. In *Proceedings of the 8th Language Resources and Evaluation Conference (LREC) 2012*, 759–765. Istanbul. http://www.lrec-conf.org/proceedings/lrec2012/pdf/327_Paper.pdf (6 March, 2014).
- Hilpert, Martin & Holger Diessel. 2016. Entrenchment in Construction Grammar. In Hans-Jörg Schmid (ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge*, 57–74. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110341423-004 (6 June, 2017).
- Langacker, Ronald W. 1988. A usage-based model. In Brygida Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics*, 127–161. Amsterdam: John Benjamins Publishing Company.
- Larasati, Septina Dian, Vladislav Kuboň & Daniel Zeman. 2011. Indonesian Morphology Tool (MorphInd): Towards an Indonesian Corpus. In *Systems and Frameworks for Computational Morphology*, 119–129. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-23138-4_8 (12 August, 2017).
- Lindsay, Mark. 2012. Rival suffixes: synonymy, competition, and the emergence of productivity. In Angela Ralli, Geert Booij, Sergio Scalise & Athanasios Karasimos (eds.), *Mediterranean Morphology Meetings - Morphology and the architecture of grammar*, vol. 8, 192–203. Patras: University of Patras. https://doi.org/10.26220/mmm.2431.
- Nomoto, Hiroki, Hannah Choi, David Moeljadi & Francis Bond. 2018. MALINDO Morph: Morphological dictionary and analyser for Malay/Indonesian. In *Proceedings of the LREC 2018 Workshop "The 13th Workshop on Asian Language Resources,"* 36–43. http://lrec-conf.org/workshops/lrec2018/W29/pdf/8_W29.pdf.
- Ogloblin, Alexander K. 1998. From Inert to Actional Causative. In Leonid Kulikov & Heinz Vater (eds.), *Typology of Verbal Categories*, 179–184. Berlin, Boston: DE GRUYTER. https://doi.org/10.1515/9783110913750.179. http://bit.ly/from-inert-to-actional-causative (30 December, 2020).
- R Core Team. 2020. *R: A language and environment for statistical computing*. Manual. Vienna, Austria. https://www.R-project.org/.
- Sneddon, James Neil, Alexander Adelaar, Dwi Noverini Djenar & Michael C. Ewing. 2010. *Indonesian reference grammar*. 2nd edn. Crows Nest, New South Wales, Australia: Allen & Unwin.
- Stefanowitsch, Anatol & Susanne Flach. 2016. The corpus-based perspective on entrenchment. In Hans-Jörg Schmid (ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge*, 101–128. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110341423-006 (6 June, 2017).
- Szymanek, Bogdan. 2005. The latest trends in English word-formation. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of Word-Formation* (Studies in Natural Language and Linguistic Theory v. 64), 429–448. Dordrecht, The Netherlands: Springer.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy McGowan, Romain François, Garrett Grolemund, et al. 2019. Welcome to the Tidyverse. *Journal of Open Source Software* 4(43). 1686. https://doi.org/10.21105/joss.01686.
- Green T-shirt picture source:
  https://www.dreamstime.com/royalty-free-stock-photo-green-t-shirt-cartoon-illustration-yellow-can-be-used-as-flashcards-babies-toddlers-language-image38814285

Thank you &
Really looking forward to any feedback

Project update:
http://bit.ly/per-kan-idn