

Pemanfaatan Bank-data Digital Dwibahasa dalam Kajian Terjemahan

Studi kasus padanan bahasa Indonesia untuk verba sinonim bahasa Inggris ROB & STEAL

Gede Primahadi Wijaya Rajeg  <https://orcid.org/0000-0002-2047-8621>

I Made Rajeg  <https://orcid.org/0000-0001-8989-0203>

Putu Dea Indah Kartini

I Gede Semara Dharma Putra

Prodi Sastra Inggris, Fakultas Ilmu Budaya, Universitas Udayana

Urgensi Penelitian

Tema peta jalan penelitian Prodi Sastra Inggris 2021 & RIP Univ. Udayana

**RIP UNIV. UDAYANA
Kluster A**

**A.40 Peningkatan
Partisipasi
Pendidikan**

**TEMA PENELITIAN
PRODI SASTRA INGGRIS**

**Pemanfaatan IT dalam
meningkatkan Pembelajaran
Ketrampilan Berbahasa Inggris**

2021

**Pengembangan model
pembelajaran alih bahasa bagi
mahasiswa sastra Inggris**

Latar Belakang (1)

Kemajuan Teknologi dan Informasi

- *Big Data* (untuk berbagai bidang)
- *Big Language/Textual Data* — **KORPUS BAHASA**
 - Untuk penelitian dan pengajaran bahasa
 - Kajian multidisipliner yang melibatkan pemakaian bahasa
 - Paralel Dwibahasa
 - ▶ Korpus Terjemahan
 - ▶ Kumpulan kalimat dari bahasa sumber (BSU) (mis. Bahasa Inggris [BIng]) dan terjemahannya di bahasa target (BT) (mis. Bahasa Indonesia [BI])

Latar Belakang (2)

Sumber Data Korpus Paralel

- **Korpus Paralel Dwibahasa** *OpenSubtitles (OSub)* *v2018* (Lison & Tiedemann 2016)
 - 9,7 juta korespondensi kalimat BIng & BI
 - ▶ 72,8 juta kata Bing
 - ▶ 60,9 juta kata BI
 - 1.31GB

Latar Belakang (2)

Sumber Data Korpus Paralel

The screenshot displays an XML file named 'en-id-OpenSubtitles.tmx' in a text editor. The XML structure consists of multiple <tu> (translation unit) elements. Each <tu> element contains one or more <tuv xml:lang='en'> (English source) and <tuv xml:lang='id'> (Indonesian target) segments. The first <tu> element (lines 12396-12399) shows a parallel sentence pair: 'I'm going to tell her to watch out, I'm going to steal the guy.' in English and 'Aku akan menyuruhnya untuk berhati-hati, aku akan mencuri pria itu.' in Indonesian. The word 'steal' in the English segment is highlighted with a blue box. Annotations with arrows point to the language tags: a yellow arrow points to '<tuv xml:lang="en">' with the label 'Kalimat Bahasa Sumber (Inggris)', and a teal arrow points to '<tuv xml:lang="id">' with the label 'Kalimat Bahasa Target (Indonesia)'. A search bar at the bottom indicates 3563 results for the regex '\bsteal\b'.

```
<tu>
  <tuv xml:lang="en"><seg>I'm going to tell her to watch out, I'm going to steal the guy.</seg></tuv>
  <tuv xml:lang="id"><seg>Aku akan menyuruhnya untuk berhati-hati, aku akan mencuri pria itu.</seg></tuv>
</tu>
<tu>
  <tuv xml:lang="en"><seg>The guy?</seg></tuv>
  <tuv xml:lang="id"><seg>Pria itu?</seg></tuv>
</tu>
<tu>
  <tuv xml:lang="en"><seg>I'm going to flirt with him and declare war to make her cry her eyes out.</seg></tuv>
  <tuv xml:lang="id"><seg>Aku akan menggodanya dan menyatakan perang untuk membuat gadis itu menangis.</seg></tuv>
</tu>
<tu>
  <tuv xml:lang="en"><seg>You don't think I'll be able to do it?</seg></tuv>
  <tuv xml:lang="id"><seg>Menurutmu, aku tak mampu melakukannya?</seg></tuv>
</tu>
<tu>
  <tuv xml:lang="en"><seg>I'm too young to be a father.</seg></tuv>
  <tuv xml:lang="id"><seg>Aku terlalu muda untuk menjadi ayah.</seg></tuv>
</tu>
```

3563 results found for `\bsteal\b`

Find: `\bsteal\b` 1 of 3563

Replace in current buffer

Replace All

Find All

Replace

Replace All

LF UTF-8 XML GitHub Git (0)

Latar Belakang (3)

Topik, Isu, Tujuan, dan Rumusan Masalah

- **kesepadanan** (*equivalence*) (Catford 1965; Kenny 2001) — konsep kunci dalam terjemahan
- Hubungan antara BSU & BT untuk memandang BT sebagai terjemahan BSU



Latar Belakang (3)

Topik, **Isu**, Tujuan, dan Rumusan Masalah

Kesepadanan (dalam konteks terjemahan komparatif)



Tipe dan Derajat
kesepadanan yang
ditemukan antara
BSU & BT

Toury (dikutip dari Kenny 2001: 80)



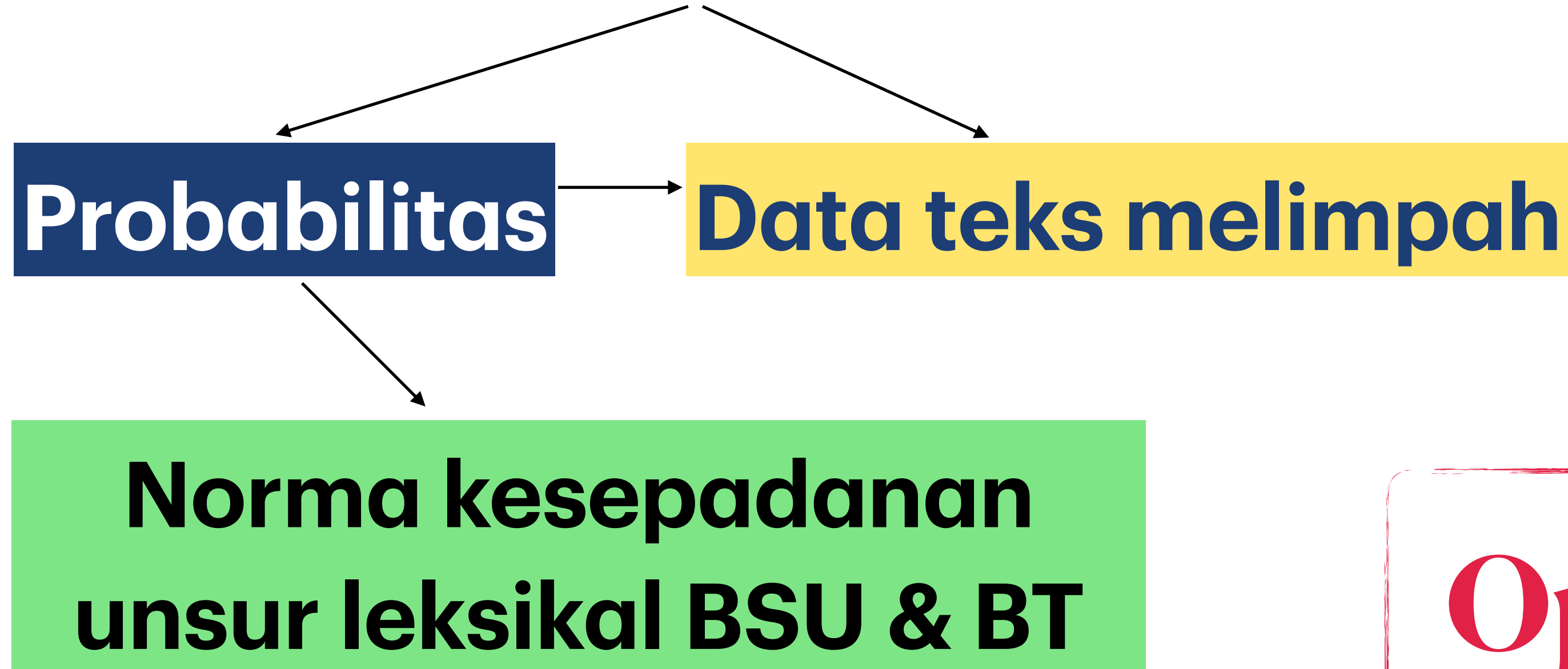
**Apakah BT dan BSU
sepadan atau tidak?**

Latar Belakang (3)

Topik, Isu, Tujuan, dan Rumusan Masalah

Catford (1965: 30-31)

**Tipe & derajat
kesepadanan leksikal/kata**



OpenSubtitles v2018

Latar Belakang (3)

Topik, Isu, Tujuan, dan Rumusan Masalah

1. *Jesse robbed the rich (of all their money)*

pencuri
SUBJ

target
OBJ

barang
KET

2. *Jesse stole the money (from the rich)*

pencuri
SUBJ

barang
OBJ

target
KET

Sinonim ROB & STEAL



Goldberg (1995: 45)

Latar Belakang (3)

Topik, Isu, Tujuan, dan Rumusan Masalah

- Apa variasi bentuk leksikal dalam sampel korpus sebagai terjemahan BI dari ROB & STEAL (R&S)?
- Bentuk leksikal mana yang secara statistik dapat dipandang sebagai norma untuk padanan R&S dalam BI?

Lihat makalah prosiding untuk rincian lebih lanjut metodologi analisis statistik

Data, Salindia/Slides, Grafik Statistik, & Kode R terbuka untuk analisis statistik:

<https://github.com/gederajeg/rob-steal-parallel-corpora> OR <https://doi.org/10.17605/OSF.IO/Y6ESA>

Hasil & Pembahasan

Kelas kata terjemahan verba ROB & STEAL

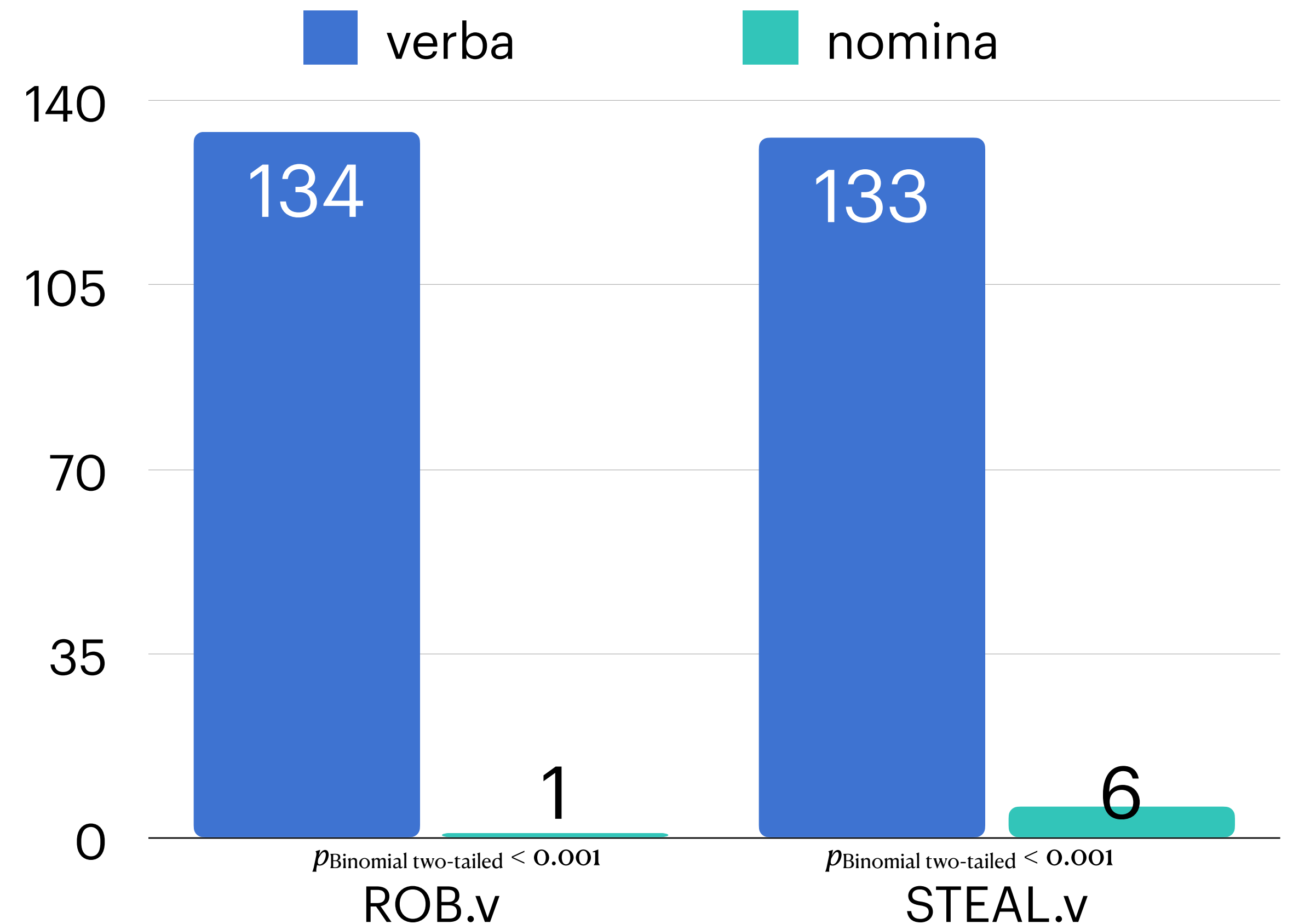
150 sampel acak ROB+150 terjemahan paralel BI

135 butir data relevan ROB sebagai verba

150 sampel acak STEAL+150 terjemahan paralel BI

139 butir data relevan STEAL sebagai verba

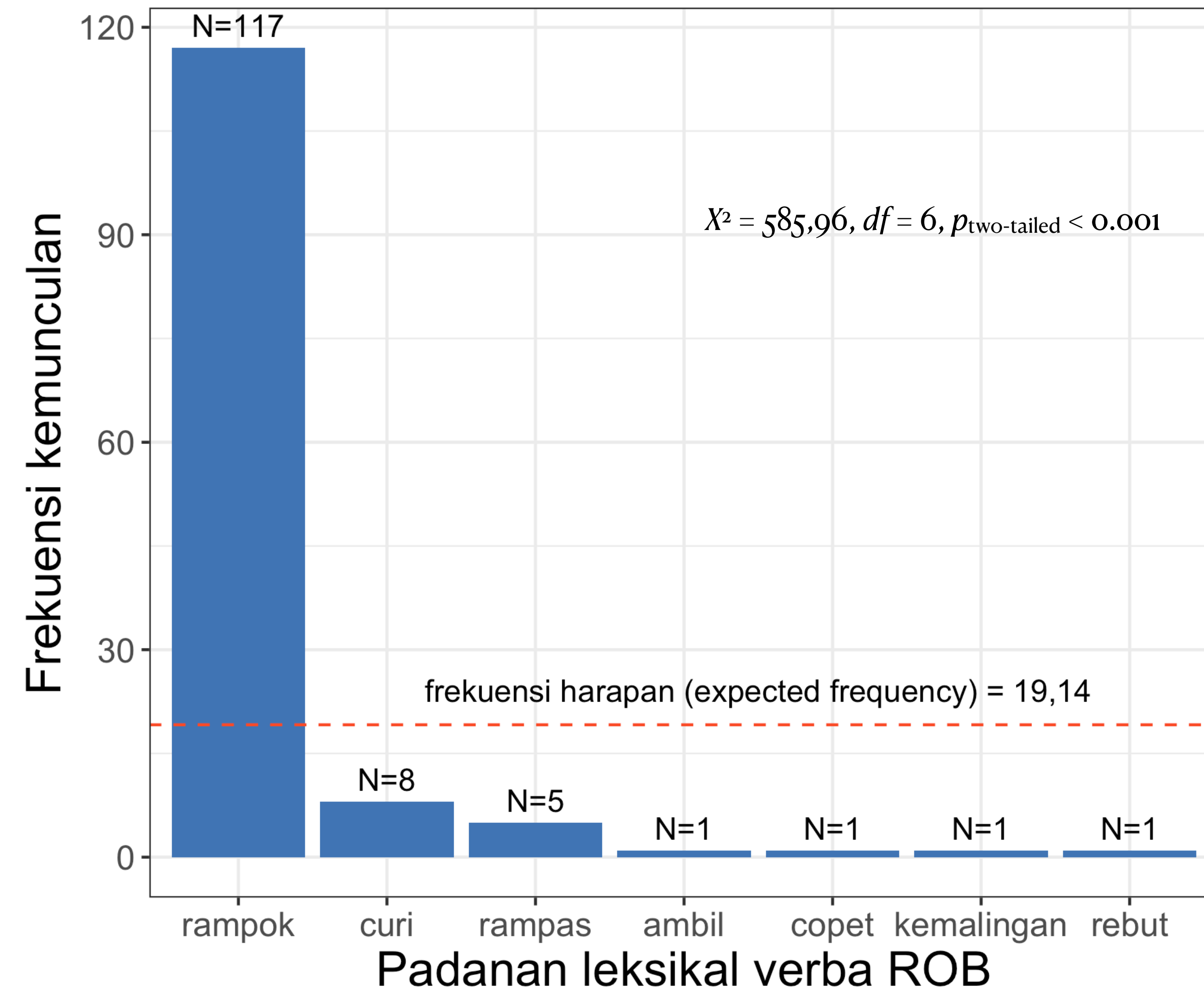
- **Derajat kesepadanan yang sangat tinggi pada tataran kelas kata untuk R&S**
- **Probabilitas R&S.v diterjemahkan juga sebagai verba dalam BI melebihi 90% ($p_{\text{Binomial}} < 0.001$)**



Bentuk leksikal terjemahan verba ROB

- Beragam variasi leksikal: REBUT, CURI, RAMPAS, RAMPOK, AMBIL, COPET, KEMALINGAN

Bentuk leksikal terjemahan verba ROB

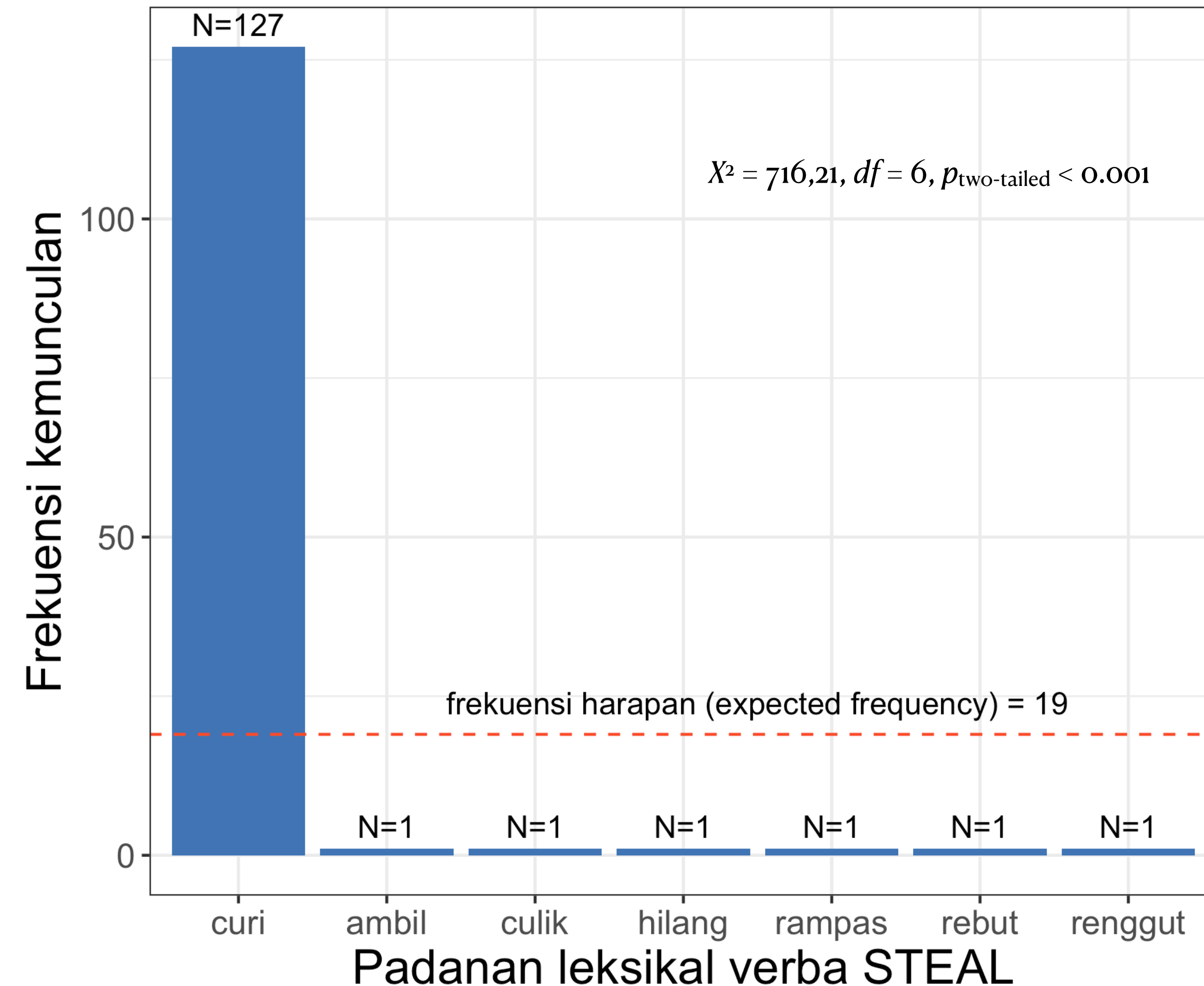


- Beragam variasi leksikal: REBUT, CURI, RAMPAS, RAMPOK, AMBIL, COPET, KEMALINGAN
- Norma: **Lema RAMPOK.v** (*merampok* [85], *dirampok* [28], *rampok* [2], *ngerampok* [1], & *kerampokan* [1])
- Probabilitas RAMPOK (vs. lainnya): **0.87 (87%)**

Bentuk leksikal terjemahan verba STEAL

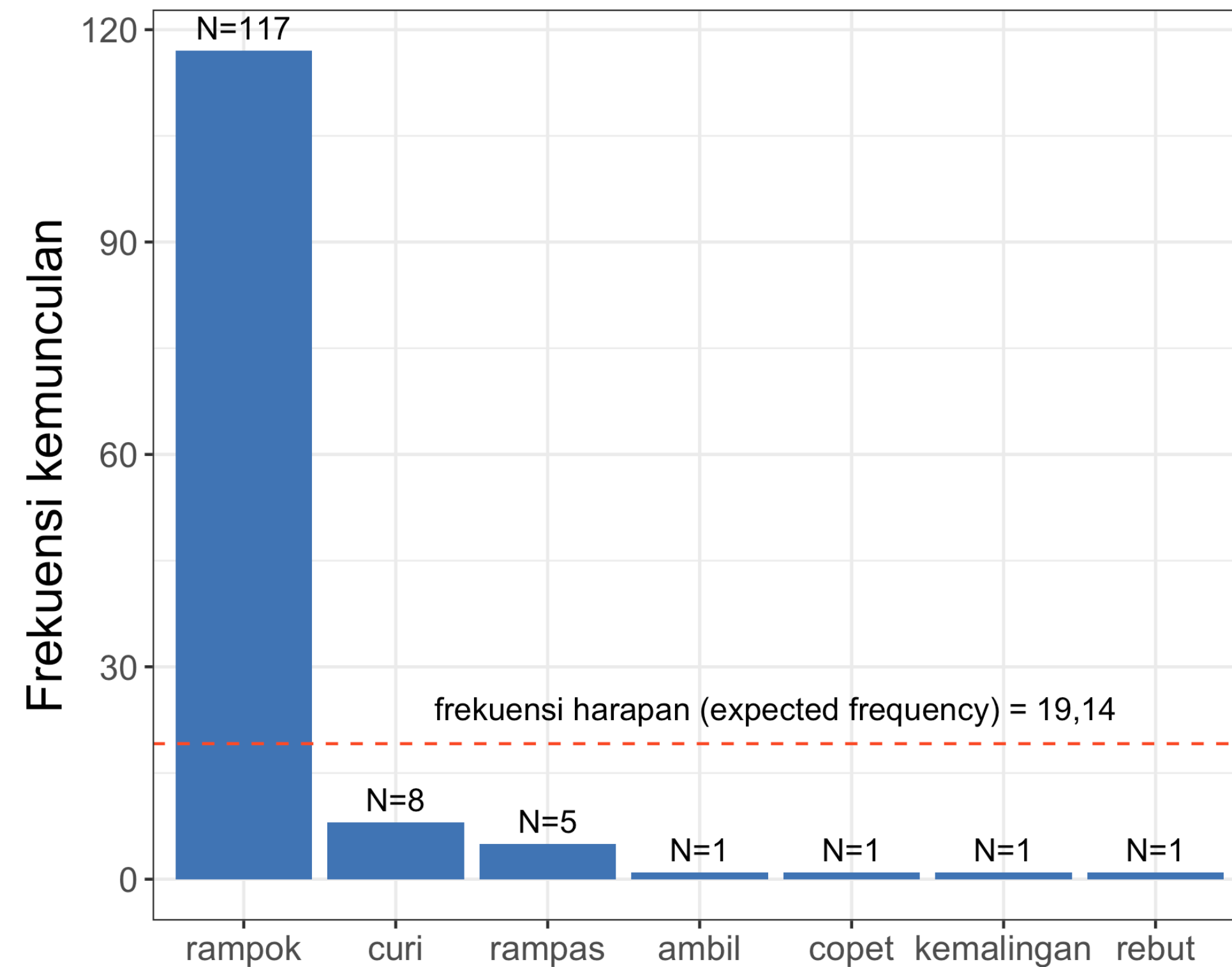
- Beragam variasi leksikal: **AMBIL, REBUT, CULIK, RENGGUT, HILANG, CURI, RAMPAS**

Bentuk leksikal terjemahan verba STEAL

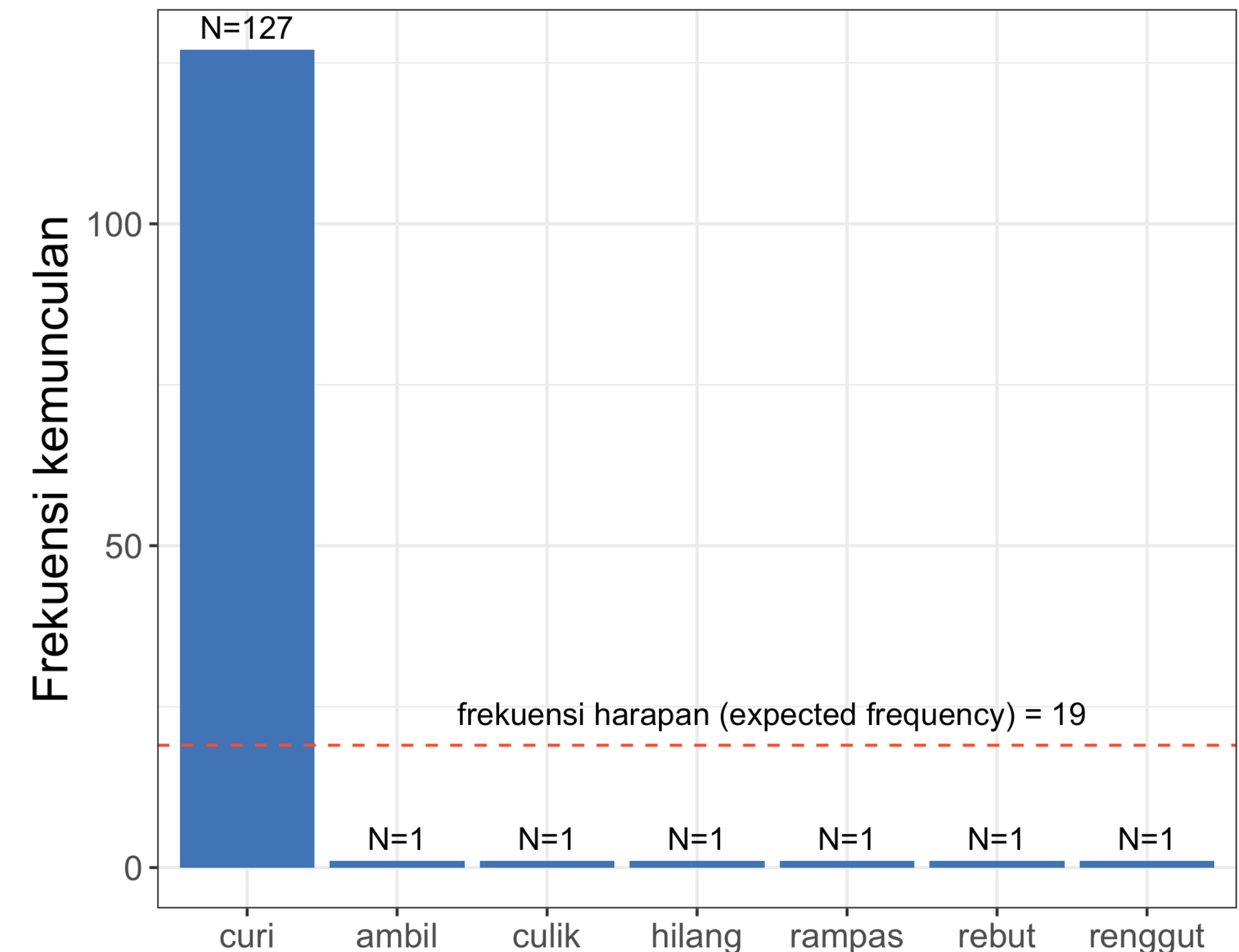


- Beragam variasi leksikal: AMBIL, REBUT, CULIK, RENGGUT, HILANG, CURI, RAMPAS
- Norma: **Lema CURI.v** (*mencuri* [102], *dicuri* [16], *curi* [9])
- Probabilitas CURI (vs. lainnya): **0.95** (**95%**)

Bentuk leksikal terjemahan verba ROB & STEAL



Padanan leksikal verba ROB



Padanan leksikal verba STEAL

Pengamatan lain:

CURI (norma padanan untuk STEAL) masih dapat merujuk pada ROB (prob. 5%).
RAMPOK (norma padanan untuk ROB), sebaliknya, tidak ditemukan merujuk pada STEAL.

Penutup

- ROB & STEAL memiliki beragam tipe bentuk leksikal terjemahan BI
- Data korpus kuantitatif: pengukuran tingkat kelaziman (norma) dari bentuk leksikal tersebut
 - Norma ROB -> RAMPOK (prob = .87; 87%)
 - Norma STEAL -> CURI (prob = .95; 95%)
- CURI bisa sebagai terjemahan ROB, namun tidak sebaliknya untuk RAMPOK sebagai terjemahan STEAL (STEAL lebih spesifik secara semantis ???)
 - Apakah struktur gramatikal CURI sebagai terjemahan dari ROB sama dengan RAMPOK sebagai terjemahan dari ROB?

Referensi

- Catford, J. C. 1965. *A linguistic theory of translation*. London: Oxford University Press.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.
- Kenny, Dorothy. 2001. Equivalence. In Mona Baker (ed.), *Routledge Encyclopedia of Translation Studies*, 77–80. London; New York: Routledge.
- Lison, Pierre & Jörg Tiedemann. 2016. Opensubtitles2016: Extracting large parallel corpora from movie and tv subtitles. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 923–929. Portorož, Slovenia: European Language Resources Association (ELRA).
- Rajeg, Gede Primahadi Wijaya, I Made Rajeg, Putu Dea Indah Kartini & I Gede Semara Dharma Putra. 2021. Material pendukung untuk *MODEL KAJIAN TERJEMAHAN BERBASIS BANK DATA TERJEMAHAN DIGITAL INGGRIS-INDONESIA DAN IMPLIKASI PEDAGOGISNYA*. Open Science Framework (OSF). <https://doi.org/10.17605/OSF.IO/Y6ESA>. <https://osf.io/y6esa/>.

TERIMA KASIH

Pemanfaatan Bank-data Digital Dwibahasa dalam Kajian Terjemahan

Studi kasus padanan bahasa Indonesia untuk verba sinonim bahasa Inggris ROB & STEAL

Gede Primahadi Wijaya Rajeg  <https://orcid.org/0000-0002-2047-8621>

I Made Rajeg  <https://orcid.org/0000-0001-8989-0203>

Putu Dea Indah Kartini

I Gede Semara Dharma Putra

Prodi Sastra Inggris, Fakultas Ilmu Budaya, Universitas Udayana



Metodologi

Pre-processing dengan kode pemrograman R

The image shows three overlapping RStudio windows. The top window displays an XML file 'en-id-OpenSubtitles.tmx' with lines 12396 to 12399. Line 12397 contains an English subtitle: '<tuv xml:lang="en"><seg>I'm going to tell her to watch out, I'm going to steal the guy.</seg></tuv>'. Line 12398 contains the corresponding Indonesian subtitle: '<tuv xml:lang="id"><seg>Aku akan menyuruhnya untuk berhati-hati, aku akan mencuri pria itu.</seg></tuv>'. The middle window shows 'eng_OpenSub.txt' with a list of subtitles. Line 3097 is highlighted: 'I'm going to tell her to watch out, I'm going to steal the guy.'. The bottom window shows 'id_OpenSub.txt' with the Indonesian translation of line 3097 highlighted: 'Aku akan menyuruhnya untuk berhati-hati, aku akan mencuri pria itu.'. A search bar at the bottom left shows the search term '\bsteal' and a 'Replace' button.