

Panduan Analisis Kata untuk Kamus Morfologi *MALINDO Morph*

Hiroki Nomoto, Livie Olivia, Thathit Puspaning Gegana,
Lukman Hakim, Friska Gania Putri, Shabrina Hazimi Putri

Februari 2018

semakan 1: September 2018

Hak cipta terpelihara © 2018 Universiti Bahasa Asing Tokyo (東京外国語大学)

Kandungan

1.	Pendahuluan	2
2.	Kriteria penyemakan	2
2.1.	Bentuk penggandaan yang ditulis dengan angka “2”	2
2.2.	Kata dasar merupakan kata majmuk	3
2.3.	Analisis kata dasar yang tidak tepat	3
2.4.	Analisis imbuhan yang tidak tepat	3
2.5.	Kata sendi nama (<i>di, ke, dsb.</i>) yang dieja dengan kata nama	4
2.6.	Kesalahan ejaan pada kata yang muncul di dalam korpus	4
2.7.	Kegagalan menganalisis akibat huruf yang berubah menjadi angka	4
2.8.	Bahasa alay	5
2.9.	Kata-kata yang jarang digunakan	5
2.10.	Perkataan-perkataan yang tidak dapat dikenalpasti	6
2.11.	Bentuk terikat (<i>antar-, anti-, multi-, non-, dsb.</i>)	6
2.12.	Partikel <i>pun</i>	6
2.13.	Perkataan tertulis di dalam ejaan yang tidak baku atau menggunakan ejaan lama	7
2.14.	Perkataan berawalan <i>N-</i>	7
2.15.	Kata-kata yang disingkatkan penulisannya	7
2.16.	Penentuan kata dasar pada KD dan KBBI yang berbeza	8
2.17.	Penentuan kata dasar untuk perkataan-perkataan yang tidak selesai diucapkan	8

1. Pendahuluan

Panduan ini menjelaskan tentang kriteria-kriteria yang menjadi dasar pembetulan untuk mengesahkan, memilih dan menyemak hasil analisis yang dilakukan secara automatik dengan penganalisis morfologi *MALINDO Morph*. Perkara yang termasuk di dalam tugas pembetulan adalah

- mengesahkan analisis yang betul;
- menghilangkan analisis yang salah; dan
- membetulkan analisis yang hampir betul supaya menjadi betul.

Notasi:

^ TAB

Bahan yang dirujuk dalam tugas pembetulan:

1. Analisis morfologi *Kamus Dewan (KD)*
Pusat Rujukan Persuratan Melayu (<http://prpm.dbp.gov.my/>)
2. Analisis morfologi *Kamus Besar Bahasa Indonesia (KBBI)*
KBBI Daring (<https://kbbi.kemdikbud.go.id/Cari/Index>)
3. Senarai imbuhan bahasa Melayu
Bahasa Melayu TUFs 文法
(http://www.tufs.ac.jp/common/fs/ase/mal/tatabahasa_web/affix_list.html)
[Selain imbuhan yang termasuk di dalam senarai ini, bahasa Indonesia juga mempunyai awalan *N-* dan akhiran *-in*.]

2. Kriteria penyemakan

2.1. Bentuk penggandaan yang ditulis dengan angka “2”

Sebelum pembetulan

couple2^ ber'couple2^ ber-^ 0^ 0^0

↓ (Membuang angka “2” pada kata dasar dan menggantikan dengan angka “0” pada kolum penggandaan menjadi “R-penuh”.)

Selepas pembetulan

couple^ ber'couple2^ ber-^ 0^ 0^R-penuh

2.2. Kata dasar merupakan kata majmuk

Apabila perkataan yang ditemui di korpus adalah kata majmuk atau kata yang terdiri daripada dua kata

Sebelum pembetulan

```
anaktir^      meng'anaktiri^  meN-^-i^0^ 0
anaktiri^     meng'anaktiri^  meN-^0^ 0^ 0
```

↓ (Memansuhkan baris pertama dan membetulkan kata dasar pada baris kedua dengan meletakkan jarak di antara kedua-dua perkataan.)

Selepas pembetulan

```
anak tiri^     meng'anaktiri^  meN-^0^ 0^ 0
```

2.3. Analisis kata dasar yang tidak tepat

Apabila sebahagian daripada kata dasar yang tidak tepat dianalisis sebagai imbuhan

Sebelum pembetulan

```
cur^  di'curi^  di-^  -i^0^ 0
```

↓ (Pembetulan pada kata dasar dan akhiran.)

Selepas pembetulan

```
curi^  di'curi^  di-^  0^ 0^ 0
```

Sebelum pembetulan

```
a^      seakanku^ se-^  -kan+-ku^ 0^0
```

↓ (Setelah mencari contoh penggunaannya di dalam internet, “ku” dalam kata di atas merupakan subjek untuk perkataan berikutnya.)

Selepas pembetulan

```
akan^ seakanku^ se-^  -ku^      0^0
```

2.4. Analisis imbuhan yang tidak tepat

Apabila sebahagian daripada imbuhan dianalisis secara tidak tepat sebagai kata dasar

Sebelum pembetulan

```
penghun^ ber'penghuni^  ber-^  -i^0^ 0
penghuni^ ber'penghuni^  ber-^  0^ 0^ 0
```

↓ (Pembetulan pada kata dasar dan awalan. Apabila terdapat beberapa imbuhan, digabungkan dengan tanda “+” sesuai dengan urutan abjad.)

Selepas pembetulan

```
huni^      ber'penghuni^  ber-+peN-^  0^ 0^ 0
```

2.5. Kata sendi nama (*di*, *ke*, dsb.) yang dieja dengan kata nama

Apabila kata sendi nama seperti *di*, *ke*, dan sebagainya dieja secara tidak tepat bersama kata nama, maka pada kolom awalan diletakkan kata sendi nama tanpa tanda penghubung “-”

Sebelum pembetulan

a^ kedia^ di-+ke-^ 0^ 0^ 0

↓ (Pembetulan pada kata dasar dan awalan.)

Selepas pembetulan

dia^ kedia^ ke^ 0^ 0^ 0

2.6. Kesalahan ejaan pada kata yang muncul di dalam korpus

Apabila ejaan pada perkataan yang ditemui di dalam korpus tidak tepat, perkataan tersebut akan tetap dibentangkan secara sedia ada sehingga pembetulan dilakukan pada kolom kata dasar dan kolom yang lain

Sebelum pembetulan

imara^ meng'imarakan^ meN-^ -kan^ 0^ 0

imarak^ meng'imarakan^ meN-^ -an^ 0^ 0

↓ (Kemungkinan maksud daripada bentuk kata di atas adalah *mengiramakan*, sehingga pada kolom kata dasar diubah menjadi “irama”.)

Selepas pembetulan

irama^ meng'imarakan^ meN-^ -kan^ 0^ 0

Sebelum pembetulan

kureng^ ke'kurengan^ 0^ 0^ ke--an^ 0

kureng^ ke'kurengan^ ke-^ -an^ 0^ 0

kurengan^ ke'kurengan^ ke-^ 0^ 0^ 0

↓ (Kemungkinan maksud daripada bentuk kata di atas adalah *kekurangan*, sehingga pada kolom kata dasar diubah menjadi “kurang”.)

Sebelum pembetulan

kurang^ ke'kurengan^ 0^ 0^ ke--an^ 0

2.7. Kegagalan menganalisis akibat huruf yang berubah menjadi angka

Apabila terjumpa kegagalan menganalisis akibat huruf “o/O” yang menjadi angka “0” atau pada huruf “l/L” yang menjadi angka “1”

Sebelum pembetulan

0rb^ PENGORBANAN^ 0^ -an^ peN--an^ 0

0rbanan^ PENGORBANAN^ peN-^ 0^ 0^ 0

↓ (Menganalisis sebagai *pengorbanan* dan melakukan pembetulan pada kata dasar dan akhiran.)

Selepas pembetulan

korban^ PENGORBANAN^ 0^ 0^ peN--an^ 0

Sebelum pembetulan

1um[^] sebelum[^] ber-+se-[^] 0[^] 0[^] 0

↓ (Menganalisis sebagai *sebelum* dan melakukan pembetulan pada kata dasar dan imbuhan.)

Selepas pembetulan

belum[^] sebelum[^] se-[^] 0[^] 0[^] 0

2.8. Bahasa alay

Pembetulan juga ditujukan terhadap bahasa alay dengan membetulkan kata dasar ke dalam bentuk yang bukan merupakan bentuk bahasa alay

Sebelum pembetulan

4was[^] men4waskan[^] meN-[^] -kan[^] 0[^] 0

4waskan[^] men4waskan[^] meN-[^] 0[^] 0[^] 0

↓ (Mengubah “4was” menjadi “awas”.)

Selepas pembetulan

awas[^] men4waskan[^] meN-[^] -kan[^] 0[^] 0

2.9. Kata-kata yang jarang digunakan

Selepas melakukan kajian di internet atau kamus terdapat bentuk perkataan yang ditemui di dalam korpus yang tidak dapat diketahui dengan pasti

Sebelum pembetulan

eh[^] Se'ehnya[^] 0[^] 0[^] se--nya[^] 0

eh[^] Se'ehnya[^] se-[^] -nya[^] 0[^] 0

ehnya[^] Se'ehnya[^] se-[^] 0[^] 0[^] 0

↓ (Setelah melakukan kajian di internet, diketahui bahawa “se'eh” merupakan perkataan di dalam dialek Pahang, Malaysia.)

Selepas pembetulan

se'eh[^] Se'ehnya[^] 0[^] -nya[^] 0[^] 0

2.10. Perkataan-perkataan yang tidak dapat dikenalpasti

Apabila terdapat perkataan yang meskipun telah dilakukan pencarian lebih mendalam tetapi tetap tidak diketahui maknanya, maka bentuk perkataan yang ditemui di dalam korpus itu dianggap sebagai kata dasar

Sebelum pembedahan

31^ di31an^ di-^ -an^ 0^0
31an^ di31an^ di-^ 0^ 0^0

↓ (Menganalisis perkataan “di31an” sebagai kata dasar dan seluruh kolum diisi dengan angka “0” pada bahagian kanannya.)

Selepas pembedahan

di31an^ di31an^ 0^ 0^ 0^0

Sebelum pembedahan

aaa^ pengaaaaan^ peN-^ -an^ 0^ 0
aaa^ pengaaaaan^ 0^ 0^ peN--an^ 0

↓ (Menganalisis perkataan “pengaaaaan” sebagai kata dasar dan seluruh kolum diisi dengan angka “0” pada bahagian kanannya.)

Selepas pembedahan

pengaaaaan^ pengaaaaan^ 0^ 0^ 0^ 0

2.11. Bentuk terikat (*antar-*, *anti-*, *multi-*, *non-*, dsb.)

Bentuk terikat seperti *antar-*, *anti-*, *multi-*, *non-*, dsb. dianalisis hanya apabila ia muncul bersamaan dengan awalan atau klitik

Sebelum pembedahan

antarpema^ antarpemain^ 0^ -in^ 0^ 0

↓ (Memisahkan bentuk terikat daripada perkataan dasar dan memasukkannya ke dalam kolum awalan.)

Selepas pembedahan

main^ antarpemain^ antar--pe-^ 0^ 0^ 0

2.12. Partikel *pun*

Sebelum pembedahan

sutawijayapun^ sutawijayapun^ 0^0^ 0^ 0

↓ (Meletakkan jarak di antara partikel *pun* dengan perkataan di depannya pada kolum kata dasar dan mengisikan “pun” tanpa tanda penghubung “-” pada kolum akhiran.)

Selepas pembedahan

sutawijaya pun^ sutawijayapun^ 0^pun^ 0^ 0

Sebelum pembedulan

pertempuranpun^ pertempuranpun^ 0^ 0^ 0^ 0
tempuranpun^ pertempuranpun^ per-^ 0^ 0^ 0

↓ (Memisahkan partikel *pun* dan imbuhan daripada kata dasar.)

Selepas pembedulan

tempur^ pertempuranpun^ 0^ pun^ per--an^ 0

2.13. Perkataan tertulis di dalam ejaan yang tidak baku atau menggunakan ejaan lama

Sebelum pembedulan

azaz^ Azazi^ 0^ -i^ 0^ 0
azazi^ Azazi^ 0^ 0^ 0^ 0

↓ (Membetulkan kata dasar kepada ejaan baku.)

Selepas pembedulan

asasi^ Azazi^ 0^ 0^ 0^ 0

Sebelum pembedulan

sadja^ Sadja^ 0^ 0^ 0^ 0

↓ (Membetulkan kata dasar kepada Ejaan Bahasa Indonesia (EBI).)

Selepas pembedulan

saja^ Sadja^ 0^ 0^ 0^ 0

2.14. Perkataan berawalan N-

Sebelum pembedulan

fikir^ Mikir^ N-^ 0^ 0^ 0
pikir^ Mikir^ N-^ 0^ 0^ 0

↓ (Memansuhkan baris “fikir” yang merupakan kata dasar dalam bahasa Melayu dan mengekalkan baris “pikir” yang merupakan bahasa Indonesia. Hal ini kerana penggunaan awalan N- yang umumnya ditemukan di dalam bahasa Indonesia.)

Selepas pembedulan

pikir^ Mikir^ N-^ 0^ 0^ 0

2.15. Kata-kata yang disingkatkan penulisannya

Sebelum pembedulan

diaorg^ diaorg^ 0^ 0^ 0^ 0

↓ (Membetulkan perkataan di kolom kata dasar dengan menukarkan ke dalam bentuk baku.)

Selepas pembedulan

dia orang^ diaorg^ 0^ 0^ 0^ 0

Sebelum pembedulan

byk2^ Byk2^ 0^ 0^ 0^ 0

↓ (Membetulkan kata-kata di kolom kata dasar dengan menukarkan ke dalam bentuk baku.)

Selepas pembedulan

banyak^ Byk2^ 0^ 0^ 0^ R-penuh

2.16. Penentuan kata dasar pada KD dan KBBI yang berbeza

Dalam situasi yang demikian, analisis dipilih melalui sudut pandang linguistik

Di dalam kes yang kata dasar menurut KD adalah *peluang* dan KBBI adalah *luang*.

Sebelum pembedulan

peluang^ peluang-peluang^0^ 0^0^ R-penuh

↓

Selepas pembedulan

luang^ peluang-peluang^pe-^ 0^0^ R-penuh

Kes yang sama juga ditemui pada perkataan lain seperti:

- **Sehingga**

Kata dasar menurut *KD* adalah *hingga*, sedangkan menurut *KBBI* adalah *sehingga*. Dalam kes ini, tidak dilakukan pembedulan terhadap analisis di bawah kerana *hingga* dinilai sebagai kata dasar yang sudah tepat.

hingga^ sehing-ga^ se-^ 0^ 0^0

- **Laki-laki**

Kata dasar menurut *KBBI* adalah *laki*, sedangkan menurut *KD* adalah *laki* atau *laki-laki*. Dalam hal ini, *laki* dinilai sebagai kata dasar yang tepat.

Sebelum pembedulan

laki2^ laki2^ 0^ 0^ 0^ 0

↓

Selepas pembedulan

laki^ laki2^ 0^ 0^ 0^ R-penuh

2.17. Penentuan kata dasar untuk perkataan-perkataan yang tidak selesai diucapkan

Kata dasar ditulis berdasarkan bentuk utuhnya.

kutip^ mengu^ meN-^ 0^ 0^ 0

Ayat asal dalam korpus:

Pekejaan Mat Jenin adalah, **mengu** /// memanjat pokok kelape dan **mengutip** kelape untuk dijual di pasa-pasa. [KL201712, *Korpus Variasi Bahasa Melayu: Standard Lisan*]