

SKRIPSI

MORPHOLOGICAL PARSER UNTUK TEKS DALAM
BAHASA INDONESIA



Andreas Novian Dwi Triastanto

NPM: 2013730038

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS
UNIVERSITAS KATOLIK PARAHYANGAN
1994

UNDERGRADUATE THESIS

**MORPHOLOGICAL PARSER FOR TEXTS IN BAHASA
INDONESIA**



Andreas Novian Dwi Triastanto

NPM: 2013730038

**DEPARTMENT OF INFORMATICS
FACULTY OF INFORMATION TECHNOLOGY AND SCIENCES
PARAHYANGAN CATHOLIC UNIVERSITY
1994**

LEMBAR PENGESAHAN

MORPHOLOGICAL PARSER UNTUK TEKS DALAM
BAHASA INDONESIA

Andreas Novian Dwi Triastanto

NPM: 2013730038

Bandung, 1 Januari 1994

Menyetujui,

Pembimbing

Dott. Thomas Anung Basuki

Ketua Tim Penguji

Anggota Tim Penguji

Dott. Thomas Anung Basuki

Dott. Thomas Anung Basuki

Mengetahui,

Ketua Program Studi

Mariskha Tri Adithia, P.D.Eng

PERNYATAAN

Dengan ini saya yang bertandatangan di bawah ini menyatakan bahwa skripsi dengan judul:

***MORPHOLOGICAL PARSER* UNTUK TEKS DALAM BAHASA INDONESIA**

adalah benar-benar karya saya sendiri, dan saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung segala risiko dan sanksi yang dijatuhkan kepada saya, apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya, atau jika ada tuntutan formal atau non-formal dari pihak lain berkaitan dengan keaslian karya saya ini.

Dinyatakan di Bandung,
Tanggal 1 Januari 1994

Meterai Rp. 6000

Andreas Novian Dwi Triastanto
NPM: 2013730038

ABSTRAK

abstrak

Kata-kata kunci: kata-kata kunci

ABSTRACT

abstract

Keywords: keywords

kepada saya sendiri

KATA PENGANTAR

kata pengantar

Bandung, Januari 1994

Penulis

DAFTAR ISI

KATA PENGANTAR	xv
DAFTAR ISI	xvii
DAFTAR GAMBAR	xix
DAFTAR TABEL	xxi
1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	1
1.3 Tujuan	1
1.4 Batasan Masalah	2
1.5 Metodologi Penelitian	2
1.6 Sistematika Pembahasan	2
2 DASAR TEORI	5
2.1 Morfologi	5
2.1.1 Morfologi dalam Linguistik	5
2.1.2 Objek Kajian Morfologi	6
2.2 Morfem	7
2.2.1 Identifikasi Morfem	7
2.2.2 Alomorf dan Morf	9
2.2.3 Jenis Morfem	9
2.2.4 Morfem Dasar, Bentuk Dasar, Pangkal (<i>Stem</i>), Akar, dan Leksem	12
2.2.5 Morfem Afiks	13
2.3 Proses Morfologi	13
2.3.1 Bentuk Dasar	14
2.3.2 Pembentuk Kata	14
2.3.3 Hasil Proses Pembentukan	15
2.4 Morfofonemik dalam Pembentukan Kata	15
2.4.1 <i>Prefiksasi ber-</i>	16
2.4.2 <i>Prefiksasi me- (termasuk klofiks me-kan dan me-i)</i>	16
2.4.3 <i>Prefiksasi pe- dan konfiksasi pe-an</i>	17
2.4.4 <i>Prefiksasi per- dan konfiksasi per-an</i>	18
2.4.5 <i>Prefiksasi ter-</i>	19
3 ANALISIS	21
3.1 Leksikon	21
3.2 Proses <i>Morphological Parsing</i>	22
DAFTAR REFERENSI	25

DAFTAR GAMBAR

2.1	Hierarki linguistik	5
2.2	Objek kajian linguistik	6
2.3	Morfem bebas dan terikat	10
2.4	Morfem dasar dan afiks	11
3.1	Struktur data trie	21

DAFTAR TABEL

2.1	Bentuk alomorf dari morfem {ber-}[1]	9
2.2	Bentuk alomorf dari morfem {me-}[1]	9

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Pemrosesan bahasa alami atau dalam bahasa Inggris disebut dengan *natural language processing* (NLP) adalah cabang ilmu komputer dan linguistik yang mengkaji interaksi antara komputer dan manusia menggunakan bahasa alami. NLP sering dianggap sebagai cabang dari kecerdasan buatan dan bidang kajiannya bersinggungan dengan linguistik komputasional. Kajian NLP antara lain mencakup segmentasi tuturan (*speech segmentation*), segmentasi teks (*text segmentation*), penandaan kelas kata (*part-of-speech tagging*), serta pengawataksaan makna (*word sense disambiguation*). Salah satu alat yang digunakan oleh komputer dalam proses mengenali bahasa alami manusia adalah *morphological parser*.

Morphological parser berfungsi untuk membagi sebuah kata menjadi komponen-komponen penyusunnya. Proses ini dapat mengenali komponen kata seperti awalan, kata dasar, sisipan, dan akhiran serta dapat mengenali jika kata tersebut merupakan kata ulang maupun kata majemuk. Proses di mana *morphological parser* melakukan tugasnya dalam menguraikan kata menjadi komponen-komponen penyusunnya disebut dengan *morphological parsing*. Proses ini dapat membantu mengurangi ambiguitas selama proses mengetahui makna suatu kalimat. Sebagai contoh, kata "mengurus" bisa mempunyai makna menjadi kurus maupun mengerjakan sebuah urusan, bergantung pada apa kata dasar dari kata tersebut. Jika kita bisa membagi kata tersebut menjadi komponen penyusunnya, kita bisa lebih yakin mengenai makna dari kata tersebut dalam kalimat. *Morphological parsing* merupakan salah satu proses penting dalam NLP.

Morphological parser sudah banyak dibuat untuk beberapa bahasa yang ada di dunia[2]. Saat ini, belum ada yang membuat *morphological parser* untuk kalimat dalam bahasa Indonesia dengan benar. Padahal, aturan morfologi pada bahasa Indonesia relatif lebih sederhana dibandingkan aturan pada bahasa lain.

1.2 Rumusan Masalah

Sehubungan dengan latar belakang yang telah diuraikan di atas, maka dibuat rumusan masalah sebagai berikut ini.

- Bagaimana aturan morfologi bahasa Indonesia?
- Bagaimana struktur data dari *lexicon* yang digunakan pada perangkat lunak?
- Bagaimana cara mengimplementasikan aturan morfologi bahasa Indonesia ke dalam perangkat lunak?
- Bagaimana performansi dari perangkat lunak yang dihasilkan?

1.3 Tujuan

Tujuan dari penelitian ini adalah sebagai berikut:

- Mengetahui aturan morfologi bahasa Indonesia
- Mengetahui struktur data dari *lexicon* yang digunakan pada perangkat lunak
- Mengimplementasikan aturan morfologi bahasa Indonesia ke dalam perangkat lunak
- Mengetahui performansi dari perangkat lunak yang dihasilkan

1.4 Batasan Masalah

Terdapat beberapa batasan masalah untuk penelitian ini:

- Kalimat yang dapat diproses adalah kalimat dalam bahasa Indonesia yang ditulis sesuai ejaan yang disempurnakan (EYD)
- Kata-kata serapan dari bahasa asing yang belum ada dalam Kamus Besar Bahasa Indonesia (KBBI) tidak dapat diproses
- Kata yang dapat diproses adalah kata yang dibentuk dari proses morfologi berupa afiksasi, komposisi, dan reduplikasi

1.5 Metodologi Penelitian

Tahap-tahap yang akan dilakukan dalam penelitian ini adalah sebagai berikut:

1. Melakukan studi literatur tentang morfologi bahasa Indonesia dan perangkat lunak *morphological parser* yang sudah ada
2. Melakukan analisis pada *morphological parser* bahasa Indonesia dan *lexicon* yang digunakan serta merancang struktur data dari *lexicon*
3. Merancang dan mengimplementasikan *lexicon* dan *morphological parser* ke dalam perangkat lunak
4. Mengumpulkan contoh kalimat dalam bahasa Indonesia sebagai bahan pengujian
5. Melakukan pengujian terhadap perangkat lunak

1.6 Sistematika Pembahasan

Keseluruhan bab yang disusun dalam karya tulis ini terbagi ke dalam bab-bab sebagai berikut:

1. BAB 1 - PENDAHULUAN membahas mengenai latar belakang, rumusan masalah, tujuan, batasan masalah, metodologi penelitian, dan sistematika pembahasan.
2. BAB 2 - DASAR TEORI membahas mengenai morfem, proses morfologi bahasa Indonesia, *lexicon* bahasa Indonesia dan struktur data dari *lexicon*.
3. BAB 3 - ANALISIS membahas mengenai analisis *morphological parser* bahasa Indonesia, *lexicon* bahasa Indonesia, dan struktur data *lexicon* yang digunakan pada perangkat lunak *Morphological Parser*.
4. BAB 4 - PERANCANGAN membahas mengenai perancangan antarmuka dan struktur data pada perangkat lunak *Morphological Parser*.

-
5. BAB 5 - IMPLEMENTASI DAN PENGUJIAN membahas mengenai implementasi dan pengujian yang dilakukan pada perangkat lunak *Morphological Parser*.
 6. BAB 6 - KESIMPULAN DAN SARAN membahas mengenai kesimpulan dan saran mengenai perangkat lunak *Morphological Parser*.

BAB 2

DASAR TEORI

Pada bab ini dijelaskan mengenai

2.1 Morfologi

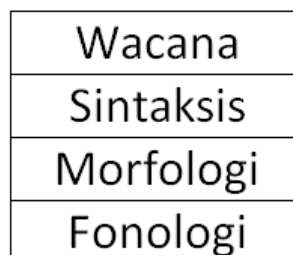
Secara etimologi, kata *morfologi* berasal dari kata *morf* yang berarti 'bentuk' dan kata *logi* yang berarti 'ilmu' [1]. Secara harfiah, kata *morfologi* berarti 'ilmu mengenai bentuk'. Di dalam kajian linguistik, *morfologi* berarti 'ilmu mengenai bentuk-bentuk dan pembentukan kata'; sedangkan di dalam kajian biologi, *morfologi* berarti 'ilmu mengenai bentuk-bentuk sel-sel tumbuhan atau jasad-jasad hidup'. Kesamaan dari dua bidang kajian tersebut adalah keduanya mengkaji tentang bentuk.

Jika morfologi dalam kajian linguistik membicarakan tentang bentuk-bentuk dan pembentukan kata, maka segala bentuk dan jenis morfem yang merupakan satuan bentuk sebelum menjadi kata perlu dibicarakan juga. Pembicaraan mengenai pembentukan kata akan melibatkan pembicaraan mengenai komponen atau unsur pembentukan kata, yaitu morfem, baik morfem dasar maupun morfem afiks, dengan berbagai alat proses pembentukan kata, yaitu afiks dalam proses pembentukan kata melalui proses afiksasi, duplikasi atau pengulangan dalam proses pembentukan kata melalui proses reduplikasi, penggabungan dalam proses pembentukan kata melalui proses komposisi, dan sebagainya.

Ujung dari proses morfologi adalah terbentuknya *kata* dalam bentuk dan makna sesuai dengan keperluan dalam satu tindak tutur. Bila bentuk dan makna yang terbentuk dari satu proses morfologi sesuai dengan yang diperlukan dalam tutur, maka bentuknya dapat dikatakan berterima; tetapi jika tidak sesuai dengan yang diperlukan, maka bentuk itu dikatakan tidak berterima. Keberterimaan atau ketidakberterimaan bentuk itu dapat juga karena alasan sosial.

2.1.1 Morfologi dalam Linguistik

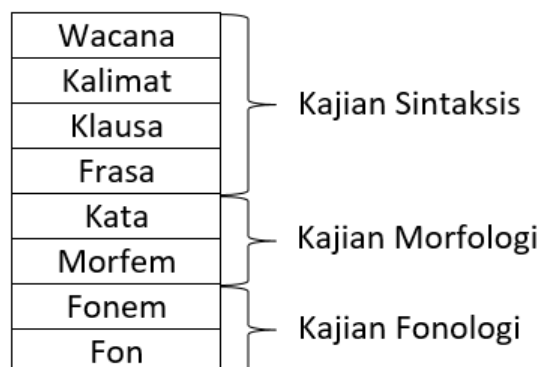
Di dalam hierarki linguistik, kajian morfologi berada di antara kajian fonologi dan sintaksis seperti tampak pada gambar 2.1 berikut:



Gambar 2.1: Hierarki linguistik [1]

Sebagai kajian yang terletak di antara kajian fonologi dan sintaksis, maka kajian morfologi mempunyai kaitan baik dengan fonologi maupun dengan sintaksis. Keterkaitan antara morfologi

dan fonologi tampak dengan adanya kajian yang disebut dengan *morfonologi* atau *morfofonemik* yaitu ilmu yang mengkaji terjadinya perubahan fonem akibat adanya proses morfologi. Keterkaitan antara morfologi dan sintaksis tampak dengan adanya kajian yang disebut *morfosintaksis* (gabungan dari kata *morfologi* dan *sintaksis*). Keterkaitan ini muncul karena adanya masalah morfologi yang perlu dibicarakan bersama dengan masalah sintaksis. Misalnya, satuan bahasa yang disebut *kata* merupakan satuan terbesar dalam kajian morfologi, sedangkan dalam kajian sintaksis merupakan satuan terkecil dalam pembentukan kalimat atau satuan sintaksis lainnya. Jadi, satuan bahasa yang disebut *kata* itu menjadi objek dalam kajian morfologi dan kajian sintaksis. Dalam gambar 2.2 dapat dilihat kedudukan *kata* dalam keseluruhan objek kajian linguistik.



Gambar 2.2: Objek kajian linguistik[1]

Keterangan singkat

Wacana adalah satuan bahasa terbesar atau tertinggi, yang berisi satu satuan ujaran yang lengkap dan utuh; dan dibangun oleh kalimat atau kalimat-kalimat yang dihubungkan secara kohesi dan koherensi.

Kalimat adalah satuan sintaksis yang dibangun oleh konstituen dasar (biasanya berupa klausa), dilengkapi dengan konjungsi (bila diperlukan), disertai dengan intonasi final (deklaratif, interogatif, imperatif, atau interjektif).

Klausa adalah satuan sintaksis yang berinti adanya sebuah predikat dan adanya fungsi lainnya. Maka sering dikatakan klausa adalah konstruksi yang bersifat predikatif.

Frase adalah satuan sintaksis berupa kelompok kata yang posisinya tidak melewati batas fungsi sintaksis (subjek, predikat, objek, atau keterangan).

Kata dalam sintaksis merupakan satuan terkecil yang biasa dan dapat menduduki salah satu fungsi sintaksis (subjek, predikat, objek, atau keterangan); dalam morfologi merupakan satuan terbesar, dibentuk melalui salah satu proses morfologi (afiksasi, reduplikasi, komposisi, akronimisasi, dan konversi).

Morfem adalah satuan gramatikal terkecil yang bermakna (secara inheren).

Fonem adalah satuan bunyi terkecil (dalam kajian fonologi) yang dapat membedakan makna kata.

Fon adalah satuan bunyi bahasa yang dilihat tanpa memperhatikan statusnya sebagai pembeda makna kata (dalam kajian fonetik).

2.1.2 Objek Kajian Morfologi

Objek kajian morfologi adalah satuan-satuan morfologi, proses-proses morfologi, dan alat-alat dalam proses morfologi itu[1]. Satuan morfologi adalah:

1. Morfem (akar atau afiks).
2. Kata.

Lalu, proses morfologi melibatkan komponen:

1. Dasar (bentuk dasar).
2. Alat pembentuk (afiks, duplikasi, komposisi, akronimisasi, dan konversi).
3. Makna gramatikal.

Morfem adalah satuan gramatikal terkecil yang bermakna. Morfem dapat berupa akar (dasar) dan dapat pula berupa afiks. Perbedaannya, morfem berupa akar dapat menjadi dasar dalam pembentukan kata, sedangkan morfem berupa afiks hanya "menjadi" penyebab terjadinya makna gramatikal. Kemudian, *kata* adalah satuan gramatikal yang terjadi sebagai hasil dari proses morfologis. Jika berdiri sendiri, setiap kata memiliki makna leksikal dan dalam kedudukannya dalam satuan ujaran memiliki makna gramatikal.

Dalam proses morfologi, dasar atau bentuk dasar merupakan bentuk yang mengalami proses morfologi. Dasar ini dapat berupa sebuah kata dasar maupun bentuk polimorfemis (bentuk berimbuhan, bentuk ulang, atau bentuk gabungan). Alat pembentuk kata dapat berupa afiks dalam proses afiksasi, pengulangan dalam proses reduplikasi, dan penggabungan dalam proses komposisi.

Makna gramatikal adalah makna yang "muncul" dalam proses gramatika. Makna gramatikal ini biasa didikotomikan dengan makna leksikal, yakni makna yang secara inheren dimiliki oleh sebuah leksem, yang merupakan satuan dari leksikon. Makna gramatikal ini mempunyai hubungan dengan komponen makna leksikal setiap dasar (akar).

2.2 Morfem

Morfem adalah satuan gramatikal terkecil yang memiliki makna[1]. Dengan kata terkecil berarti "satuan" itu tidak dapat dianalisis menjadi lebih kecil lagi tanpa merusak maknanya. Sebagai contoh, bentuk *membeli* dapat dianalisis menjadi dua bentuk terkecil yaitu {me-} dan {beli}. Bentuk {me-} adalah sebuah morfem, yakni morfem afiks yang secara gramatikal memiliki sebuah makna; dan bentuk {beli} juga sebuah morfem, yakni morfem dasar yang secara leksikal memiliki makna. Kalau bentuk *beli* dianalisis menjadi lebih kecil lagi menjadi *be-* dan *li*, keduanya tidak memiliki makna apapun. Jadi, keduanya bukan morfem. Contoh lain, bentuk *berpakaian* dapat dianalisis ke dalam satuan-satuan terkecil menjadi {ber-}, {pakai}, dan {-an}. Ketiganya adalah morfem, di mana {ber-} adalah morfem prefiks, {pakai} adalah morfem dasar, dan {-an} adalah morfem sufiks. Ketiganya memiliki makna. Morfem {ber-} dan morfem {-an} memiliki makna gramatikal, sedangkan morfem {pakai} memiliki makna leksikal. Perlu dicatat dalam konvensi linguistik sebuah bentuk dinyatakan sebagai morfem ditulis dalam kurung kurawal ({...}).

2.2.1 Identifikasi Morfem

Satuan bahasa merupakan komposit antara bentuk dan makna[1]. Oleh karena itu, untuk menetapkan sebuah bentuk adalah morfem atau bukan didasarkan pada kriteria bentuk dan makna itu. Hal-hal berikut dapat menjadi pedoman untuk menentukan morfem dan bukan morfem itu.

1. Dua bentuk yang sama atau lebih memiliki makna yang sama merupakan sebuah morfem. Umpamanya kata *bulan* pada ketiga kalimat berikut adalah sebuah morfem yang sama.
 - *Bulan* depan dia akan menikah.
 - Sudah tiga *bulan* dia belum bayar uang SPP.
 - *Bulan* November lamanya 30 hari.
2. Dua bentuk yang sama atau lebih bila memiliki makna yang berbeda merupakan dua morfem yang berbeda. Misalnya kata *bunga* pada kedua kalimat berikut adalah dua buah morfem yang berbeda.

- Bank Indonesia memberi *bunga* 5 persen per tahun.
 - Dia datang membawa seikat *bunga*.
3. Dua buah bentuk yang berbeda, tetapi memiliki makna yang sama, merupakan dua morfem yang berbeda. Umpamanya, kata *ayah* dan kata *bapak* pada kedua kalimat berikut adalah dua morfem yang berbeda.
- *Ayah* pergi ke Medan.
 - *Bapak* baru pulang dari Medan.
4. Bentuk-bentuk yang mirip (berbeda sedikit) tetapi maknanya sama adalah sebuah morfem yang sama, asal perbedaan bentuk itu dapat dijelaskan secara fonologis. Umpamanya, bentuk-bentuk *me-*, *mem-*, *men-*, *meny-*, *meng-*, dan *menge-* pada kata-kata berikut adalah sebuah morfem yang sama.
- *melihat*
 - *membina*
 - *mendengar*
 - *menyusul*
 - *mengambil*
 - *mengecat*
5. Bentuk yang hanya muncul dengan pasangan satu-satunya adalah sebuah morfem juga. Umpamanya bentuk *renta* pada konstruksi *tua renta*, dan bentuk *kuyup* pada konstruksi *basah kuyup* adalah juga morfem. Contoh lain, bentuk *bugar* pada *segar bugar*, dan bentuk *mersik* pada *kering mersik*.
6. Bentuk yang muncul berulang-ulang pada satuan yang lebih besar apabila memiliki makna yang sama adalah juga merupakan morfem yang sama. Misalnya bentuk *baca* pada kata-kata berikut adalah sebuah morfem yang sama.
- *membaca*
 - *pembaca*
 - *pembacaan*
 - *bacaan*
 - *terbaca*
 - *keterbacaan*
7. Bentuk yang muncul berulang-ulang pada satuan bahasa yang lebih besar, apabila mempunyai bentuk bahasa yang sama namun maknanya berbeda (polisemi) merupakan morfem yang sama. Umpamanya, kata *kepala* pada kalimat-kalimat berikut memiliki makna yang berbeda, tetapi tetap merupakan morfem yang sama.
- Ibunya menjadi *kepala* sekolah di sana.
 - Nomor teleponnya tertera pada *kepala* surat itu.
 - *Kepala* jarum itu terbuat dari plastik.
 - Setiap *kepala* mendapat bantuan sepuluh ribu rupiah.
 - Tubuhnya memang besar tetapi sayang *kepalanya* kosong.

2.2.2 Alomorf dan Morf

Morfem sebenarnya merupakan barang abstrak karena ada dalam konsep. Sedangkan yang konkret, yang ada dalam pertuturan adalah alomorf, yang tidak lain adalah realisasi dari morfem itu[1]. Jadi, sebagai realisasi dari morfem itu, alomorf ini bersifat nyata/ada. Umpamanya morfem {kuda} direalisasikan dalam bentuk unsur leksikal *kuda*, dan morfem {-kan} direalisasikan dalam bentuk sufiks *-kan* seperti terdapat pada *meluruskan* atau *membacakan*.

Pada umumnya sebuah morfem hanya memiliki sebuah alomorf. Namun, ada juga morfem yang direalisasikan dalam beberapa bentuk alomorf. Misalnya, morfem {ber-} memiliki tiga bentuk alomorf, yaitu *ber-*, *be-*, dan *bel-*, seperti terdapat pada gambar 2.1 berikut.

Morfem	Alomorf	Contoh (pada kata)
<i>ber-</i>	<i>ber-</i> <i>be-</i> <i>bel-</i>	bertemu, berdoa beternak, bekerja belajar.

Tabel 2.1: Bentuk alomorf dari morfem {ber-}[1]

Malah morfem {me-} memiliki enam buah alomorf seperti tampak pada gambar 2.2.

Morfem	Alomorf	Contoh (pada kata)
<i>me-</i>	<i>me-</i> <i>mem-</i> <i>men-</i> <i>meny-</i> <i>meng-</i> <i>menge-</i>	melihat, merawat. membaca, membawa, menduga, mendengar, menyisir, menyusul, menggali, mengebor, mengecat, mengetik

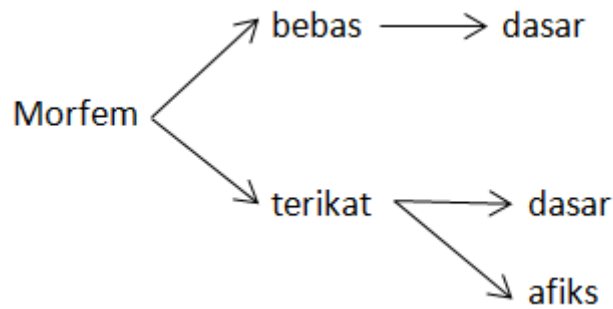
Tabel 2.2: Bentuk alomorf dari morfem {me-}[1]

Di samping istilah *morfem* dan *alomorf* ada pula istilah *morf*. Dalam kajian morfologi, morf berarti bentuk yang belum diketahui statusnya, apakah sebagai morfem atau sebagai alomorf. Jadi, sebenarnya wujud fisik morf adalah sama dengan wujud fisik alomorf. Sedangkan morfem merupakan "abstraksi" dari alomorf atau alomorf-alomorf yang ada.

2.2.3 Jenis Morfem

Dalam kajian morfologi biasanya dibedakan adanya beberapa morfem berdasarkan kriteria tertentu, seperti kriteria kebebasan, keutuhan, makna, dan sebagainya. Berikut adalah jenis-jenis morfem tersebut.

1. Berdasarkan kebebasannya untuk dapat digunakan langsung dalam pertuturan, dibedakan adanya morfem bebas dan morfem terikat. Morfem bebas adalah morfem yang tanpa keterkaitannya dengan morfem lain dapat langsung digunakan dalam pertuturan. Misalnya, morfem {pulang}, {merah}, dan {pergi}. Morfem bebas ini tentunya berupa morfem dasar. Sedangkan morfem terikat adalah morfem yang harus terlebih dahulu bergabung dengan morfem lain untuk dapat digunakan dalam pertuturan. Dalam hal ini, semua afiks dalam bahasa Indonesia termasuk morfem terikat. Di samping itu, banyak juga morfem terikat yang berupa morfem dasar, seperti {henti}, {juang}, dan {geletak}. Untuk dapat digunakan, ketiga morfem ini harus terlebih dahulu diberi afiks atau digabung dengan morfem lain. Misalnya {juang} menjadi *berjuang*, *pejuang*, dan *daya juang*; *henti* harus digabung dulu dengan afiks tertentu seperti menjadi *berhenti*, *perhentian*, dan *menghentikan*; dan *geletak* harus diberi imbuhan dulu, misalnya menjadi *tergeletak*, dan *mengeletak*. Adanya morfem bebas dan terikat dapat digambarkan sebagai berikut.



Gambar 2.3: Morfem bebas dan terikat[1]

Berkenaan dengan bentuk dasar terikat, perlu dikemukakan catatan sebagai berikut:

Pertama, bentuk dasar terikat seperti *gaul*, *juang*, dan *henti* lazim juga disebut sebagai *prakategorial* karena bentuk-bentuk tersebut belum memiliki kategori sehingga tidak dapat digunakan dalam pertuturan.

Kedua, Verhaar (1978) juga memasukkan bentuk-bentuk seperti *beli*, *baca*, dan *tulis* ke dalam kelompok *prakategorial*, karena untuk digunakan di dalam kalimat harus terlebih dahulu diberi prefiks *me-*, prefiks *di-*, atau prefiks *ter-*. Dalam kalimat imperatif memang tanpa imbuhan bentuk-bentuk tersebut dapat digunakan. Namun, kalimat imperatif adalah hasil transformasi dari kalimat aktif transitif (yang memerlukan imbuhan).

Ketiga, bentuk-bentuk seperti *renta* (yang hanya muncul dalam *tua renta*), *kerontang* (yang hanya muncul dalam *kering kerontang*), dan *kuyup* (yang hanya muncul dalam *basah kuyup*) adalah juga termasuk morfem terikat. Lalu, oleh karena hanya muncul dalam pasangan tertentu, maka disebut *morfem unik*.

Keempat, bentuk-bentuk yang disebut klitika merupakan morfem yang agak sukar ditentukan statusnya, apakah morfem bebas atau morfem terikat. Kemunculannya dalam pertuturan selalu terikat dengan bentuk lain, tetapi dapat dipisahkan. Umpamanya klitika *-ku* dalam konstruksi *bukuku* dapat dipisahkan sehingga menjadi *buku baruku*. Dilihat dari posisi tempatnya dibedakan adanya proklitika, yaitu klitika yang berposisi di muka kata yang diikuti seperti klitika *ku-* dalam bentuk *kubawa* dan *kauambil*. Sedangkan yang disebut enklitika adalah klitika yang berposisi di belakang kata yang dilekati, seperti klitika *-mu* dan *-nya* pada bentuk *nasibmu* dan *duduknya*.

Kelima, bentuk-bentuk yang termasuk preposisi dan konjungsi seperti *dan*, *oleh*, *di*, dan *karena* secara morfologis termasuk morfem bebas, tetapi secara sintaksis merupakan bentuk terikat (dalam satuan sintaksisnya).

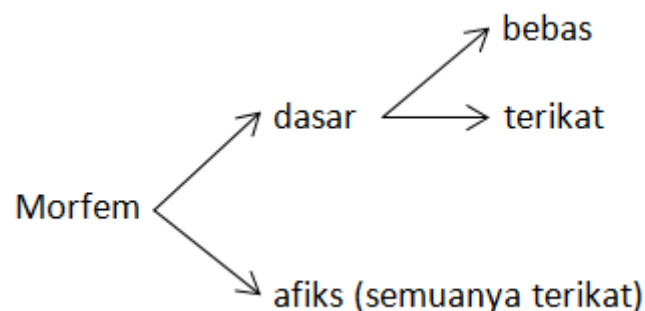
Keenam, bentuk-bentuk yang oleh Kridalaksana (1989) disebut proleksem, seperti *a* (pada *asusila*), *dwi* (pada *dwibahasa*), dan *ko* (pada *kopilot*) juga termasuk morfem terikat.

2. Berdasarkan keutuhan bentuknya dibedakan adanya morfem utuh dan morfem terbagi. Morfem utuh secara fisik merupakan satu-kesatuan yang utuh. Semua morfem dasar, baik bebas maupun terikat, serta prefiks, infiks, dan sufiks termasuk morfem utuh. Sedangkan yang dimaksud morfem terbagi adalah morfem yang fisiknya terbagi atau disisipi morfem lain. Karenanya semua konfiks (seperti *pe-an*, *ke-an*, dan *per-an*) adalah termasuk morfem terbagi. Namun, mengenai morfem terbagi ini ada dua catatan yang perlu diperhatikan.

Pertama, semua konfiks adalah morfem terbagi; tetapi pada bentuk *ber-an* ada yang berupa konfiks dan ada yang bukan konfiks. Jika kata dalam bentuk *ber-an* tidak memiliki arti ketika hanya ditambahkan prefiks *ber-* atau sufiks *-an* saja, maka bentuk *ber-an* tersebut adalah berupa konfiks. Namun, jika kata tersebut memiliki arti ketika hanya ditambahkan prefiks *ber-* atau sufiks *-an* saja, maka bentuk *ber-an* tersebut adalah berupa *klofiks* (akronim dari kelompok afiks). Contoh, kata *bermunculan* adalah dasar *muncul* ditambahkan konfiks *ber-an* sementara kata *berpakaian* adalah prefiks *ber-* yang ditambahkan pada bentuk *pakaian*.

Kedua, dalam bahasa Indonesia ada afiks yang disebut infiks, yaitu afiks yang ditempatkan di tengah (di dalam kata). Umpamanya infiks *-el-* pada dasar *tunjuk* menjadi kata *telunjuk*. Di sini infiks itu memecah morfem *tunjuk* menjadi dua bagian, yaitu *t-el-unjuk*. Dengan demikian morfem *t-unjuk* menjadi morfem terbagi, bukan morfem utuh.

3. Berdasarkan kemungkinan menjadi dasar dalam pembentukan kata, dibedakan *morfem dasar* dan *morfem afiks*. Morfem dasar adalah morfem yang dapat menjadi dasar dalam suatu proses morfologi. Misalnya, morfem {beli}, {makan}, dan {merah}. Namun, perlu dicatat bentuk dasar yang termasuk dalam kategori preposisi dan konjungsi tidak pernah mengalami proses afiksasi. Sedangkan, yang tidak dapat menjadi dasar, melainkan hanya sebagai pembentuk disebut morfem afiks, seperti morfem {me-}, {-kan}, dan {pe-an}. Berdasarkan pembagian ini, maka dapat dibuat gambar 2.4 berikut.



Gambar 2.4: Morfem dasar dan afiks[1]

4. Berdasarkan jenis fonem yang membentuknya dibedakan adanya *morfem segmental* dan *morfem suprasegmental* atau *morfem nonsegmental*. Morfem segmental adalah morfem yang dibentuk oleh fonem-fonem segmental, yakni morfem yang berupa bunyi dan dapat disegmentasikan. Misalnya morfem {lihat}, {ter-}, {sikat}, dan {-lah}. Sedangkan morfem suprasegmental adalah morfem yang terbentuk dari nada, tekanan, durasi, dan intonasi. Dalam bahasa Indonesia tidak ditemukan morfem suprasegmental ini; tetapi dalam bahasa Cina, Thai, dan Burma morfem tersebut kita dapat.
5. Berdasarkan kehadirannya secara konkret dibedakan adanya *morfem wujud* dan *morfem tanwujud*. Yang dimaksud dengan morfem wujud adalah morfem yang secara nyata ada; tetapi yang tanwujud kehadirannya tidak nyata. Morfem tanwujud ini tidak ada dalam bahasa Indonesia, tetapi ada dalam bahasa Inggris.
6. Berdasarkan ciri semantik dibedakan adanya morfem bermakna leksikal dan morfem tak bermakna leksikal. Sebuah morfem disebut bermakna leksikal karena di dalam dirinya, secara inheren, telah memiliki makna. Semua morfem dasar bebas, seperti {makan}, {pulang}, dan {pergi} termasuk morfem bermakna leksikal. Sebaliknya, morfem afiks seperti {ber-}, {ke-}, dan {ter-} termasuk morfem tak bermakna leksikal. Morfem bermakna leksikal dapat langsung menjadi unsur dalam pertuturan, sementara morfem tidak bermakna leksikal tidak dapat.

Dikotomi morfem bermakna leksikal dan tidak bermakna leksikal ini, untuk bahasa Indonesia timbul masalah. Morfem-morfem seperti {juang}, {henti}, dan {gaul} memiliki makna leksikal atau tidak. Kalau dikatakan memiliki makna leksikal, pada kenyataannya morfem-morfem itu belum dapat digunakan dalam pertuturan sebelum mengalami proses morfologi. Kalau dikatakan tidak bermakna leksikal, pada kenyataannya morfem-morfem tersebut bukan afiks.

2.2.4 Morfem Dasar, Bentuk Dasar, Pangkal (*Stem*), Akar, dan Leksem

Morfem dasar, bentuk dasar (lebih lazim dasar (*base*) saja), pangkal (*stem*), akar, dan leksem adalah lima istilah yang lazim digunakan dalam kajian morfologi. Namun, seringkali digunakan secara kurang cermat, malah seringkali berbeda. Oleh karena itu, ada baiknya istilah-istilah tersebut dibicarakan dulu sebelum pembicaraan mengenai proses-proses morfologi.

Istilah *morfem dasar* biasanya digunakan sebagai dikotomi dengan morfem afiks. Jadi, bentuk-bentuk seperti {beli}, {juang}, dan {kucing} adalah morfem dasar. Morfem dasar ini ada yang termasuk morfem bebas seperti {beli}, {kucing}, dan {pulang}; tetapi ada pula yang termasuk morfem terikat, seperti {juang}, {henti}, dan {tempur}. Sedangkan morfem afiks seperti {ber-}, {di-}, dan {-an} jelas semuanya termasuk morfem terikat seperti dijelaskan pada gambar 2.4 di atas.

Sebuah morfem dasar dapat menjadi bentuk dasar atau dasar (*base*) dalam suatu proses morfologi. Artinya, dapat diberi afiks tertentu dalam proses afiksasi, dapat diulang dalam proses reduplikasi, atau dapat digabung dengan morfem yang lain dalam suatu proses komposisi atau pemajemukan.

Istilah *bentuk dasar* atau *dasar (base)* biasanya digunakan untuk menyebut sebuah bentuk yang menjadi dasar dalam suatu proses morfologi. Bentuk dasar ini dapat berupa morfem tunggal, tetapi dapat juga berupa gabungan morfem. Umpamanya pada kata *berbicara* yang terdiri dari morfem {ber-} dan morfem {bicara}; maka morfem {bicara} adalah menjadi bentuk dasar dari kata *berbicara* itu, yang kebetulan juga berupa morfem dasar. Pada kata *dimengerti* bentuk dasarnya adalah *mengerti*, dan pada kata *keanekaragaman* bentuk dasarnya adalah *aneka ragam*. Pada bentuk reduplikasi *rumah-rumah* bentuk dasarnya adalah *rumah*, pada bentuk reduplikasi *berlari-lari* bentuk dasarnya *berlari*, dan pada bentuk reduplikasi *kemerah-merahan* bentuk dasarnya adalah *kemerahan*. Lalu, pada komposisi *sate ayam* bentuk dasarnya adalah *sate*, pada komposisi *ayam betina* bentuk dasarnya adalah *ayam*, dan pada komposisi *pasar induk* bentuk dasarnya adalah *pasar*. Jadi, bentuk dasar adalah bentuk yang langsung menjadi dasar dalam suatu proses morfologi. Wujudnya dapat berupa morfem tunggal, dapat juga berupa bentuk polimorfemis.

Istilah *pangkal* atau *stem* digunakan untuk menyebut bentuk dasar dalam proses pembentukan kata inflektif, atau pembubuhan afiks inflektif. Hal ini terutama terjadi pada bahasa-bahasa fleksi, seperti bahasa Arab, bahasa Itali, bahasa Jerman, dan bahasa Prancis. Dalam bahasa Indonesia proses pembentukan kata inflektif hanya terjadi pada proses pembentukan verba transitif, yakni verba yang berprefiks *me-* (yang dapat diganti dengan *di-*, prefiks *ter-*, dan prefiks *zero*). Misalnya, pada kata *membeli* pangkalnya adalah *beli*, pada kata *mendaratkan* pangkalnya adalah *daratkan*, dan pada kata *menangisi* pangkalnya adalah bentuk *tangisi*.

Istilah *akar (root)* digunakan untuk menyebut bentuk yang tidak dapat dianalisis lebih jauh lagi. Artinya, akar adalah bentuk yang tersisa setelah semua afiksnya ditinggalkan. Misalkan pada kata *memberlakukan* setelah semua afiksnya ditanggalkan (yaitu prefiks *me-*, prefiks *ber-*, dan sufiks *-kan*) dengan cara tertentu, maka yang tersisa adalah akar *laku*. Akar *laku* ini tidak dapat dianalisis lebih jauh lagi tanpa merusak makna akar tersebut. Contoh lain, kata *keberterimaan* kalau semua afiksnya ditanggalkan akan tersisa akarnya yaitu bentuk *terima*. Bentuk *terima* ini pun tidak dapat dianalisis lebih jauh lagi.

Istilah *leksem* ada digunakan dalam dua bidang kajian linguistik, yaitu bidang *morfologi* dan bidang *semantik*. Dalam kajian morfologi, leksem digunakan untuk mewadahi konsep "bentuk yang akan menjadi kata" melalui proses morfologi. Umpamanya bentuk PUKUL (dalam konvensi 'morfologi' leksem ditulis dengan huruf kapital semua) adalah sebuah leksem yang akan menurunkan kata-kata yang seperti *memukul*, *dipukul*, *terpukul*, *pukul*, *pukulan*, *pemukul*, dan *pemukulan*. Sedangkan dalam kajian semantik leksem adalah satuan bahasa yang memiliki sebuah makna. Jadi, bentuk-bentuk seperti *kucing*, *membaca*, *matahari*, *membanting tulang*, dan *sumpah serapah* adalah leksem.

Dari bentuk *leksem* ada bentuk-bentuk turunannya, yaitu *leksikon*, *leksikal*, *leksikologi*, dan *leksikografi*. Istilah leksikon dalam arti 'kumpulan leksem' dapat dipadankan dengan istilah *kosakata* atau *perbendaharaan kata*.

2.2.5 Morfem Afiks

Sudah disebutkan di atas bahwa morfem afiks adalah morfem yang tidak dapat menjadi dasar dalam pembentukan kata, tetapi hanya menjadi unsur pembentuk dalam proses afiksasi. Dalam bahasa Indonesia dibedakan adanya morfem afiks yang disebut:

1. *Prefiks*, yaitu afiks yang dibubuhkan di kiri bentuk dasar, yaitu prefiks *ber-*, prefiks *me-*, prefiks *per-*, prefiks *di-*, prefiks *ter-*, prefiks *se-*, dan prefiks *ke-*.
2. *Infiks*, yaitu afiks yang dibubuhkan di tengah kata, biasanya pada suku awal kata, yaitu infiks *-el-*, infiks *-em-*, dan infiks *-er-*.
3. *Sufiks*, adalah afiks yang dibubuhkan di kanan bentuk dasar, yaitu sufiks *-kan*, sufiks *-i*, dan sufiks *-an*.
4. *Konfiks*, yaitu afiks yang dibubuhkan di kiri dan di kanan bentuk dasar secara bersamaan karena konfiks ini merupakan satu kesatuan afiks. Konfiks yang ada dalam bahasa Indonesia adalah konfiks *ke-an*, konfiks *ber-an*, konfiks *pe-an*, konfiks *per-an*, dan konfiks *se-nya*.
5. *Klitika*¹, adalah imbuhan yang dalam ucapan tidak mempunyai tekanan sendiri dan tidak merupakan kata karena tidak dapat berdiri sendiri. Jadi, klitika merupakan bentuk yang selalu terikat pada bentuk (kata) lain. Dilihat dari posisi tempatnya, dibedakan adanya proklitika, yaitu klitika yang berposisi di sebelah kiri kata yang diikuti seperti klitika *ku-* dan *kau-* dalam bentuk *kubawa* dan *kauambil*. Sedangkan yang disebut enklitika adalah klitika yang berposisi di belakang kata yang dilekati, seperti klitika *-ku*, *-mu*, *-nya*, dan *-lah* pada bentuk *bukuku*, *nasibmu*, *duduknya*, dan *pergilah*. Ada juga bentuk klitika yang ditulis terpisah dari kata yang diimbuhan, yaitu klitika *pun* pada bentuk *kami pun*.
6. Dalam bahasa Indonesia ada bentuk kata yang *berklofiks*, yaitu kata yang dibubuhi afiks pada kiri dan kanannya; tetapi pembubuhannya itu tidak sekaligus, melainkan bertahap. Kata-kata berklofiks dalam bahasa Indonesia adalah yang berbentuk *me-kan*, *me-i*, *memper-*, *memper-kan*, *memper-i*, *ber-kan*, *di-kan*, *di-i*, *diper-*, *diper-kan*, *diper-i*, *ter-kan*, dan *ter-i*.
7. Dalam ragam nonbaku ada afiks nasal yang direalisasikan dengan nasal *m-*, *n-*, *ny-*, *ng-*, dan *nge-*. Kridalaksana (1989) menyebut afiks nasal ini dengan istilah *simulfiks*. Contoh: *nulis*, *nyisir*, *ngambil*, dan *ngecat*.

2.3 Proses Morfologi

Proses morfologi pada dasarnya adalah proses pembentukan kata dari sebuah bentuk dasar melalui pembubuhan afiks (dalam proses afiksasi), pengulangan (dalam proses reduplikasi), penggabungan (dalam proses komposisi), pemendekan (dalam proses akronimisasi), dan pengubahan status (dalam proses konversi)[1]. Prosedur ini berbeda dengan analisis morfologi yang menceraikan kata (sebagai satuan sintaksis) menjadi bagian-bagian atau satuan-satuan yang lebih kecil. Jadi, kalau dalam analisis morfologi; seperti menggunakan teknik *Immediate Constituent Analysis* (IC Analysis), terhadap kata *berpakaian*, misalnya, mula-mula kata *berpakaian* dianalisis menjadi bentuk *ber-* dan *pakaian*; lalu bentuk *pakaian* dianalisis lagi menjadi bentuk *pakai* dan *-an*. Maka dalam proses morfologi prosedurnya dibalik: mula-mula dasar *pakai* diberi sufiks *-an* menjadi *pakaian*. Kemudian kata *pakaian* itu diberi prefiks *ber-* menjadi *berpakaian*. Jadi, kalau analisis morfologi menceraikan data kebahasaan yang ada, sedangkan proses morfologi mencoba menyusun dari komponen-komponen kecil menjadi sebuah bentuk yang lebih besar yang berupa kata kompleks atau kata yang polimorfemis.

¹id.wikibooks.org/wiki/Bahasa_Indonesia/Klitika

Proses morfologi melibatkan komponen (1) bentuk dasar, (2) alat pembentuk (afiksasi, reduplikasi, komposisi, akronimisasi, dan konversi), (3) makna gramatikal, dan (4) hasil proses pembentukan.

2.3.1 Bentuk Dasar

Pada subbab 2.2.4 telah disinggung bahwa *bentuk dasar* adalah bentuk yang kepadanya dilakukan proses morfologi itu. Bentuk dasar itu dapat berupa akar seperti *baca*, *pahat*, dan *juang* pada kata *membaca*, *memahat*, dan *berjuang*. Dapat berupa bentuk polimorfemis seperti bentuk *bermakna*, *berlari*, dan *jual beli* pada kata *kebermaknaan*, *berlari-lari*, dan *berjual beli*.

Dalam proses reduplikasi bentuk dasar dapat berupa akar, seperti akar *rumah* pada kata *rumah-rumah*, akar *tinggi* seperti pada kata *tinggi-tinggi*, dan akar *marah* pada kata *marah-marah*. Dapat juga berupa kata berimbuhan seperti *menembak* pada kata *menembak-nembak*, kata berimbuhan *bangunan* pada kata *bangunan-bangunan*, dan kata berimbuhan *kemerahan* pada kata *kemerah-merahan*. Dapat juga berupa kata gabung seperti *rumah sakit* pada kata *rumah-rumah sakit*, dan *anak nakal* pada kata *anak-anak nakal*.

Dalam proses komposisi dapat berupa akar *sate* pada kata *sate ayam*, *sate padang*, dan *sate lontong*; dapat berupa dua buah akar seperti akar *kampung* dan akar *halaman* pada kata *kampung halaman*, atau akar *tua* dan akar *muda* pada kata *tua muda*.

Ada perbedaan bentuk antara *pelajar* dan *pengajar*. Menurut kajian tradisional dan struktural bentuk dasar dari kedua kata itu adalah sama, yaitu akar *ajar*. Dalam kajian proses di sini bentuk dasar kedua kata itu tidaklah sama. Bentuk dasar kata *pelajar* adalah *belajar* sedangkan bentuk dasar kata *pengajar* adalah *mengajar*. Ini dikarenakan makna gramatikal kata *pelajar* adalah 'orang yang belajar' sedangkan makna gramatikal kata *pengajar* adalah 'orang yang mengajar'. Contoh lain, bentuk dasar kata *penyatuan* adalah *menyatukan* karena makna *penyatuan* adalah 'hal/proses menyatukan'. Sedangkan bentuk dasar kata *persatuan* adalah *bersatu* atau *mempersatukan* karena makna gramatikalnya adalah 'hal bersatu' atau 'hal mempersatukan'. Namun, secara teoretis dapat juga dikatakan bentuk dasar kata *pelajar* dan *pengajar* adalah sama yaitu *ajar*; tetapi bentuk *pelajar* dibentuk dari dasar *ajar* melalui verba *belajar*, sedangkan *pengajar* dibentuk dari dasar *ajar* melalui verba *mengajar*. Demikian juga kata *penyatuan* dibentuk dari dasar *satu* melalui verba *menyatukan*, sedangkan kata *persatuan* dibentuk dari dasar *satu* melalui verba *bersatu* atau *mempersatukan*.

Dari uraian di atas, jelas bahwa konsep *bentuk dasar* tidak sama dengan pengertian *morfem dasar* atau *kata dasar*. Ini dikarenakan bentuk dasar dapat juga berupa bentuk-bentuk polimorfemis.

2.3.2 Pembentuk Kata

Komponen kedua dalam proses morfologi adalah alat pembentuk kata. Sejauh ini alat pembentuk kata dalam proses morfologi adalah (a) afiks dalam proses afiksasi, (b) pengulangan dalam proses reduplikasi, (c) penggabungan dalam proses komposisi, (d) pemendekan atau penyingkatan dalam proses akronimisasi, dan (e) pengubahan status dalam proses konversi.

Dalam proses afiksasi sebuah afiks diimbuhkan pada bentuk dasar sehingga hasilnya menjadi sebuah kata. Umpamanya pada dasar *baca* diimbuhkan afiks *me-* sehingga menghasilkan kata *membaca* yaitu sebuah verba transitif aktif; pada dasar *juang* diimbuhkan afiks *ber-* sehingga menghasilkan verba intransitif *berjuang*.

Berkenaan dengan jenis afiksnya, biasanya proses afiksasi itu dibedakan atas *prefiksasi*, yaitu proses pembubuhan prefiks, *konfiksasi* yakni proses pembubuhan konfiks, *sufiksasi* yaitu proses pembubuhan sufiks dan *infiksasi* yakni proses pembubuhan infiks. Perlu dicatat dalam bahasa Indonesia proses infiksasi sudah tidak produktif lagi. Dalam hal ini perlu juga diperhatikan adanya *klofiksasi*, yaitu kelompok afiks yang proses afiksasinya dilakukan bertahap. Misalnya pembentukan kata *menangisi*, mula-mula pada dasar *tangis* diimbuhkan sufiks *-i*; setelah itu baru dibubuhkan prefiks *me-*.

Proses prefiksasi dilakukan oleh prefiks *ber-*, *me-*, *di-*, *ter-*, *ke-*, dan *se-*; infiksasi dilakukan oleh infiks *-el-*, *-em-*, dan *-er-*; sufiksasi dilakukan sufiks *-an*, *-kan*, dan *-i*; sedangkan konfiksasi dilakukan oleh konfiks *pe-an*, *per-an*, *ke-an*, *se-nya*, dan *ber-an* (ada yang bukan konfiks). Namun, perlu dicatat ada afiks yang sangat produktif yaitu prefiks *ber-* dan prefiks *me-*; ada yang cukup produktif, yaitu prefiks *ter-*, sufiks *-kan*, sufiks *-i*, dan sufiks *-an*; dan juga ada yang tidak produktif lagi, yaitu infiks *-el-*, *-em-*, dan *-er-*.

Alat pembentuk kedua adalah pengulangan bentuk dasar yang digunakan dalam proses reduplikasi. Hasil dari proses reduplikasi ini lazim disebut dengan istilah *kata ulang*. Secara umum dikenal adanya tiga macam pengulangan, yaitu pengulangan secara utuh, pengulangan dengan perubahan bunyi vokal maupun konsonan, dan pengulangan sebagian.

Alat pembentuk ketiga adalah penggabungan sebuah bentuk pada bentuk dasar yang ada dalam proses komposisi. Penggabungan ini juga merupakan alat yang banyak digunakan dalam pembentukan kata karena banyaknya konsep yang belum ada wadahnya dalam bentuk sebuah kata. Misalnya, bahasa Indonesia hanya punya sebuah kata untuk berbagai macam warna merah. Oleh karena itulah dibentuk gabungan kata seperti *merah jambu*, *merah darah*, dan *merah bata*.

Alat pembentuk keempat adalah abreviasi khusus yang digunakan dalam proses akronimisasi. Disebut abreviasi khusus karena semua abreviasi menghasilkan akronim. Abreviasi dari bentuk *Sekolah Menengah Atas* menjadi SMA adalah bukan akronim; tetapi hasil abreviasi dari *Jakarta Bogor Ciawi* menjadi *Jagorawi* adalah akronim.

Alat kelima dalam pembentukan kata adalah perubahan status dalam proses yang disebut konversi. Misalnya, bentuk *gunting* yang berstatus nomina dalam kalimat "gunting ini terbuat dari baja", dapat diubah statusnya menjadi bentuk *gunting* yang berstatus verba, seperti dalam kalimat "gunting dulu baik-baik, nanti baru dilem".

2.3.3 Hasil Proses Pembentukan

Proses morfologi atau proses pembentukan kata mempunyai dua hasil yaitu *bentuk* dan *makna gramatikal*. Bentuk dan makna gramatikal merupakan dua hal yang berkaitan erat; bentuk merupakan wujud fisiknya dan makna gramatikal merupakan isi dari wujud fisik atau bentuk itu.

Wujud fisik dari hasil proses afiksasi adalah kata berafiks, disebut juga kata berimbuhan, kata turunan, atau kata terbitan. Wujud fisik dari proses reduplikasi adalah kata ulang, atau disebut juga bentuk ulang. Wujud fisik dari hasil proses komposisi adalah kata gabung, disebut juga gabungan kata, kelompok kata, atau kata majemuk (tentang istilah kata majemuk banyak menimbulkan persoalan. Nanti, akan dibahas tersendiri pada subbab lainnya).

2.4 Morfofonemik dalam Pembentukan Kata

Morfofonemik (disebut juga morfonologi atau morfofonologi) adalah kajian mengenai terjadinya perubahan bunyi atau perubahan fonem sebagai akibat dari adanya proses morfologi, baik proses afiksasi, proses reduplikasi, maupun proses komposisi[1]. Umpamanya, dalam proses pengimbuhan sufiks *-an* pada dasar *hari* akan muncul bunyi [y], yang dalam ortografi tidak dituliskan, tetapi dalam ucapan dituliskan.

$$Hari + an \rightarrow [hariyan]$$

Contoh lain, dalam proses pengimbuhan sufiks *-an* pada dasar *jawab* akan terjadi pergeseran letak bunyi [b] kebelakang, membentuk suku kata baru.

$$Ja.wab + an \rightarrow [ja.wa.ban]$$

Morfofonemik dalam pembentukan kata bahasa Indonesia terutama terjadi dalam proses afiksasi. Dalam proses reduplikasi dan komposisi hampir tidak ada. Dalam proses afiksasi pun terutama, hanya dalam prefiksasi *ber-*, prefiksasi *me-*, prefiksasi *pe-*, prefiksasi *per-*, prefiksasi *ter-*, konfiksasi *pe-an*, konfiksasi *per-an*.

2.4.1 Prefiksasi *ber-*

Morfofonemik dalam proses pengimbuhan prefiks *ber-* berupa: (a) pelesapan fonem /r/ pada prefiks *ber-* itu; (b) perubahan fonem /r/ pada prefiks *ber-* itu menjadi fonem /l/; dan (c) pengeklalan fonem /r/ yang terdapat prefiks *ber-* itu.

1. Pelesapan fonem /r/ pada prefiks *ber-* itu terjadi apabila bentuk dasar yang diimbuhi mulai dengan fonem /r/, atau suku pertama bentuk dasarnya berbunyi [er]. Misalnya:
ber + renang → berenang
ber + ragam → beragam
ber + racun → beracun
ber + kerja → bekerja
ber + ternak → beternak
ber + cermin → becemin
2. Perubahan fonem /r/ pada prefiks *ber-* menjadi fonem /l/ terjadi bila bentuk dasarnya akar *ajar*; tidak ada contoh lain.
ber + ajar → belajar
3. Pengeklalan fonem /r/ pada prefiks *ber-* tetap /r/ terjadi apabila bentuk dasarnya bukan yang ada pada poin 1 dan 2 di atas.
ber + obat → berobat
ber + korban → berkorban
ber + getah → bergetah
ber + lari → berlari
ber + tamu → bertamu

2.4.2 Prefiksasi *me-* (termasuk klotiks *me-kan* dan *me-i*)

Morfofonemik dalam proses pengimbuhan dengan prefiks *me-* dapat berupa: (a) pengeklalan fonem; (b) penambahan fonem; dan (c) peluluhan fonem.

1. Pengeklalan fonem di sini artinya tidak ada fonem yang berubah, tidak ada yang dilepas dan tidak ada yang ditambahkan. Hal ini terjadi apabila bentuk dasarnya diawali dengan konsonan /r, l, w, y, m, n, ng, dan ny/. Contoh:
me + rawat → merawat
me + lirik → melirik
me + wasiat → wasiat
me + yakin → meyakinkan
me + makan → memakan
me + nanti → menanti
me + nganga → nganga
me + nyanyi → nyanyi
2. Penambahan fonem, yakni penambahan fonem nasal /m, n, ng, dan nge/. Penambahan fonem nasal /m/ terjadi apabila bentuk dasarnya dimulai dengan konsonan /b/ dan /f/. Umpamanya:
me + baca → membaca
me + buru → memburu
me + fitnah → memfitnah
me + fokus → memfokus(kan)

Penambahan fonem nasal /n/ terjadi apabila bentuk dasarnya dimulai dengan konsonan /d/. Umpamanya:

me + dengar → mendengar

me + duga → menduga

me + dapat → mendapat

Penambahan fonem nasal /ng/ terjadi apabila bentuk dasarnya dimulai dengan konsonan /g, h, dan kh/ dan huruf vokal /a, i, u, e, dan o/. Contoh:

me + goda → menggoda

me + hina → menghina

me + khayal → mengkhayal

me + ambil → mengambil

me + iris → mengiris

me + ukur → mengukur

me + elak → mengelak

me + obral → mengobral

Penambahan fonem nasal /nge/ terjadi apabila bentuk dasarnya hanya terdiri dari satu suku kata. Misalnya:

me + bom → mengebom

me + cat → mengecat

me + lap → mengelap

3. Peluluhan fonem terjadi apabila prefiks *me-* diimbuhkan pada bentuk dasar yang dimulai dengan konsonan bersuara /s, k, p, dan t/. Dalam hal ini konsonan /s/ diluluhkan dengan nasal /ny/, konsonan /k/ diluluhkan dengan nasal /ng/, konsonan /p/ diluluhkan dengan nasal /m/, dan konsonan /t/ diluluhkan dengan nasal /n/. Contoh:

me + sikat → menyikat

me + kirim → mengirim

me + pilih → memilih

me + tolong → menolong

2.4.3 Prefiksasi *pe-* dan konfiksasi *pe-an*

Morf fonemik dalam proses pengimbuhan dengan prefiks *pe-* dan konfiks *pe-an* sama dengan morf fonemik yang terjadi dalam proses pengimbuhan dengan prefiks *me-*, yaitu (a) pengeklalan fonem; (b) penambahan fonem; dan (c) peluluhan fonem.

1. Pengeklalan fonem, artinya tidak ada perubahan fonem, dapat terjadi apabila bentuk dasarnya diawali dengan konsonan /r, l, y, w, m, n, ng, dan ny/. Contoh:

pe + rawat → perawat

pe + latih → pelatih

pe + yakin → peyakin

pe + waris → pewaris

pe - an + manfaat → pemanfaatan

pe - an + nanti → penantian

pe + nganga → penganga

pe + nyanyi → penyanyi

2. Penambahan fonem, yakni penambahan fonem nasal /m, n, ng, dan nge/ antara prefiks dan bentuk dasar. Penambahan fonem nasal /m/ terjadi apabila bentuk dasarnya diawali oleh konsonan /b/. Contoh:

pe + baca → pembaca

pe + bina → pembina

pe + buru → pemburu

Penambahan fonem nasal /n/ terjadi apabila bentuk dasarnya diawali oleh konsonan /d/.
Contoh:

pe + dengar → *pendengar*

pe + duga → *penduga*

pe + didik → *pendidik*

Penambahan fonem nasal /ng/ terjadi apabila bentuk dasarnya diawali dengan konsonan /g, h, dan kh/ dan vokal /a, i, u, e, o/. Contoh:

pe + gali → *penggali*

pe + hambat → *penghambat*

pe + khianat → *pengkhianat*

pe + angkat → *pengangkat*

pe + inap → *penginap*

pe + usir → *pengusir*

pe + elak → *pengelak*

pe + obral → *pengobral*

Penambahan fonem nasal /nge/ terjadi apabila bentuk dasarnya berupa bentuk dasar satu suku. Contoh:

pe + bom → *pengebom*

pe + cat → *pengecat*

pe + lap → *pengelap*

3. Peluluhan fonem, apabila prefiks *pe-* (atau *pe-an*) diimbuhkan pada bentuk dasar yang diawali dengan konsonan tak bersuara /s, k, p, dan t/. Dalam hal ini konsonan /s/ diluluhkan dengan nasal /ny/, konsonan /k/ diluluhkan dengan nasal /ng/, konsonan /p/ diluluhkan dengan nasal /m/, dan konsonan /t/ diluluhkan dengan nasal /n/. Contoh:

pe + saring → *penyaring*

pe + kumpul → *pengumpul*

pe + pilih → *pemilih*

pe + tulis → *penulis*

2.4.4 Prefiksasi *per-* dan konfiksasi *per-an*

Morfofonemik dalam pengimbuhan prefiks *per-* dan konfiks *per-an* dapat berupa: (a) pelepasan fonem /r/ pada prefiks *per-* itu; (b) perubahan fonem /r/ dari prefiks *per-* itu menjadi fonem /l/; dan (c) pengekaln fonem /r/ tetap /r/.

1. Pelepasan fonem /r/ terjadi apabila bentuk dasarnya dimulai dengan fonem /r/, atau suku kata pertamanya /er/. Contoh:
per + ringan → *peringan*
per + rendah → *perendah*
per + ternak → *peternak*
per + kerja → *pekerja*
2. Perubahan fonem /r/ menjadi /l/ terjadi apabila bentuk dasarnya berupa kata *ajar*.
per + ajar → *pelajar*
3. Pengekaln fonem /r/ terjadi apabila bentuk dasarnya bukan yang disebutkan pada poin 1 dan 2 di atas. Contoh:
per + kaya → *perkaya*
per + kecil → *perkecil*
per + lambat → *perlambat*
per + tegas → *pertegas*

2.4.5 *Prefiksasi ter-*

Morfofonemik dalam proses pengimbuhan dengan prefiks *ter-* dapat berupa: (a) pelesapan fonem /r/ dari prefiks *ter-* itu; (b) perubahan fonem /r/ dari prefiks *ter-* itu menjadi fonem /l/; dan (c) pengeklalan fonem /r/ itu.

1. Pelesapan fonem dapat terjadi apabila prefiks *ter-* diimbuhkan pada bentuk dasar yang dimulai dengan konsonan /r/. Misalnya:
ter + rasa → *terasa*
ter + rangkum → *terangkum*
ter + rebut → *terebut*
2. Perubahan fonem /r/ pada prefiks *ter-* menjadi fonem /l/ terjadi apabila prefiks *ter-* itu diimbuhkan pada bentuk dasar *anjur*.
ter + anjur → *telanjur*
3. Pengeklalan fonem /r/ pada prefiks *ter-* tetap menjadi /r/ apabila prefiks *ter-* itu diimbuhkan pada bentuk dasar yang bukan disebutkan pada poin 1 dan 2 di atas. Contoh:
ter + dengar → *terdengar*
ter + jauh → *terjauh*
ter + lempar → *terlempar*
ter + baik → *terbaik*

BAB 3

ANALISIS

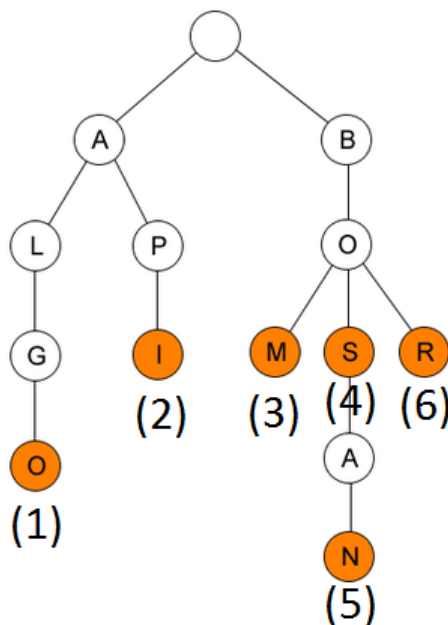
Pada bab ini dijelaskan mengenai

3.1 Leksikon

Leksikon, seperti dijelaskan pada subbab 2.2.4, dapat dipadankan dengan istilah *kosakata* atau *perbendaharaan kata*. Leksikon dibutuhkan pada proses *morphological parsing* untuk mengetahui apakah sebuah kata yang sedang diproses adalah sebuah bentuk dasar yang valid atau tidak dalam bahasa Indonesia. Leksikon menyimpan kumpulan bentuk dasar dan turunannya untuk nantinya diakses ketika proses *morphological parsing* dilakukan.

Leksikon dalam proses *morphological parsing* harus bisa diakses dengan cepat dan efektif. Hal ini dikarenakan leksikon akan diakses sangat sering dalam proses ini. Leksikon akan diakses sekitar 3-5 kali untuk setiap kata yang sedang diproses. Oleh karena itu, leksikon perlu disimpan pada struktur data yang memungkinkan waktu akses yang cepat supaya keseluruhan proses dapat dijalankan dalam waktu yang masuk akal.

Struktur data yang saat ini terkenal paling cepat untuk diakses adalah struktur data *trie*. Trie adalah struktur data berbentuk pohon yang menyimpan himpunan string yang jika ditelusuri setiap node mulai dari akar hingga daun akan membentuk suatu string yang merupakan kunci yang kita cari. Setiap string yang dihasilkan dari node awal yang sama akan mempunyai awalan (prefiks) yang sama, karena itulah trie disebut juga pohon prefiks.



Gambar 3.1: Struktur data trie

Struktur data trie yang digambarkan pada bagan 3.1 menyimpan enam string kunci dari dua buah awalan, yaitu string "A" dan "B". Jika kita telusuri dari node akar "A" sampai node daun "O", kita akan mendapat string "ALGO" yang ditandai dengan nomor (1). String lain yang disimpan pada contoh tersebut adalah string "API" pada nomor (2), string "BOM" pada nomor (3), string "BOS" pada nomor (4), string "BOSAN" pada nomor (5), dan string "BOR" pada nomor (6).

Perlu diperhatikan bahwa sebuah string kunci tidak harus disimpan dengan node terakhir ada pada posisi daun, seperti pada string "BOS" pada nomor 4. Node terakhir pada string tersebut merupakan node internal. Penyimpanan seperti ini bisa dilakukan dengan menandai setiap node yang merupakan akhir dari sebuah string yang membentuk kata.

Kata yang disimpan dalam leksikon terdiri dari dua jenis kata, yaitu kata dasar dan kata turunan. Contoh kata dasar adalah kata 'makan', 'sapu', dan 'kerja' sementara contoh kata turunan adalah kata 'makan-makan', 'menyapu', dan 'kerja bakti'. Kata-kata turunan disimpan sebagai bagian dari kata dasar dan dapat diakses ketika dibutuhkan. Dalam Kamus Besar Bahasa Indonesia Dalam Jaringan (KBBI daring)¹, kata dasar dan kata turunan disimpan secara terpisah namun keduanya dapat diakses melalui cara yang sama, yaitu dengan menuliskannya pada kolom pencarian. Sementara pada Kamus Besar Bahasa Indonesia Luar Jaringan (KBBI luring)², hanya kata dasar saja yang bisa diakses dengan menuliskannya pada kolom pencarian. Pada penelitian kali ini akan digunakan struktur penyimpanan seperti pada KBBI luring.

Struktur penyimpanan seperti pada KBBI luring memungkinkan untuk mengenali perbedaan antara kata dasar dan kata yang telah melalui proses morfologi seperti afiksasi, reduplikasi, dan komposisi. Perangkat lunak yang dirancang pada penelitian ini harus dapat menentukan apakah sebuah kata merupakan kata dasar yang valid dalam bahasa Indonesia.

3.2 Proses *Morphological Parsing*

Pada subbab 2.3 telah dibahas mengenai proses morfologi, yang pada dasarnya adalah proses pembentukan kata melalui beberapa proses, yaitu pembubuhan afiks (afiksasi), pengulangan (reduplikasi), penggabungan (komposisi), pemendekan (akronimisasi), dan pengubahan status (konversi). Proses *morphological parsing* merupakan kebalikan dari proses morfologi. Masukan bagi proses *morphological parsing* adalah kata atau kalimat yang telah melalui proses morfologi dan keluarannya adalah komponen-komponen penyusunnya.

Proses *morphological parsing* untuk setiap kata dalam masukan dapat dituliskan sebagai berikut:

1. Periksa leksikon, jika kata tersebut ada dalam leksikon, masukkan sebagai salah satu kemungkinan keluaran
2. Periksa adanya simbol penghubung (-), yang menandakan hasil proses reduplikasi, lalu lakukan pemisahan kata dan lakukan proses parsing pada kedua kata tersebut
3. Jika ada kata yang mengikuti, periksa kemungkinan kata yang sedang diproses dan kata yang mengikuti adalah dua kata hasil komposisi, lalu lakukan proses parsing pada kedua kata tersebut
4. Periksa adanya kemungkinan afiks, baik itu prefiks, sufiks, infiks, maupun konfiks. Pisahkan afiks yang ditemukan dengan komponen kata yang lain dan lakukan pengecekan leksikon pada komponen kata tersebut
5. Jika sudah dilakukan pemisahan terhadap kemungkinan afiks namun kata yang sedang diproses tidak ditemukan dalam leksikon, kemungkinan kata tersebut bukan kata dalam bahasa Indonesia

¹<https://kbbi.kemdikbud.go.id/>

²<http://ebsoft.web.id/kbbi-kamus-besar-bahasa-indonesia-offline-gratis/>

Sebagai contoh, jika dilakukan proses *morphological parsing* pada kata 'kemerah-merahan', maka prosesnya adalah sebagai berikut:

- Periksa leksikon, kata tersebut tidak ditemukan dalam leksikon
- Ditemukan simbol penghubung (-) sehingga diketahui kata tersebut adalah hasil proses reduplikasi. Pisahkan kata sehingga didapat kata 'kemerah' dan 'merahan'
- Periksa leksikon kembali untuk kedua kata tersebut, kedua kata tersebut tidak ditemukan dalam leksikon
- Periksa kemungkinan afiks pada kata 'kemerah' dan 'merahan'
- Didapatkan prefiks {ke-} + bentuk dasar {merah} dan bentuk dasar {merah} + sufiks {-an}, yang setelah ditinjau lebih lanjut didapatkan konfiks {ke-an} + bentuk dasar {merah}
- Hasil akhir proses parsing adalah konfiks {ke-an} + bentuk dasar {merah} + reduplikasi

Walaupun berbentuk mirip dengan kata 'kemerah-merahan', proses parsing pada kata 'berlari-larian' sedikit berbeda. Prosesnya adalah sebagai berikut:

- Periksa leksikon, kata tersebut tidak ditemukan dalam leksikon
- Ditemukan simbol penghubung (-) sehingga diketahui kata tersebut adalah hasil proses reduplikasi. Pisahkan kata sehingga didapat kata 'berlari' dan 'larian'
- Periksa leksikon kembali untuk kedua kata tersebut, kata 'berlari' ditemukan sebagai turunan dari kata dasar {lari} yang ditambahkan prefiks {ber-}
- Periksa kemungkinan afiks pada kata 'larian'
- Didapatkan bentuk dasar {lari} + sufiks {-an}
- Hasil akhir proses parsing adalah prefiks {ber-} + bentuk dasar {lari} + sufiks {-an} + reduplikasi

Pada proses pemeriksaan leksikon yang pertama, pemeriksaan dilakukan hanya pada kata dasar, sementara pada proses pemeriksaan leksikon yang kedua dan seterusnya dilakukan pada kata dasar dan turunannya. Kata 'kemerah-merahan' dan kata 'berlari-larian' ada dalam leksikon sebagai turunan dari kata dasar {merah} dan {lari} sehingga kedua kata tersebut tidak ditemukan dalam proses pemeriksaan leksikon yang pertama. Hal ini dilakukan supaya dapat membedakan antara kata yang dibentuk dari proses konfiksasi dengan kata yang dibentuk dari proses klofiksasi.

Untuk kata dengan kemungkinan hasil parsing lebih dari satu, seperti kata 'beruang', prosesnya adalah sebagai berikut:

- Periksa leksikon, ditemukan bentuk dasar {beruang}, masukkan sebagai salah satu kemungkinan keluaran
- Periksa kemungkinan afiks pada kata 'beruang'
- Didapatkan prefiks {ber-} + bentuk dasar {uang}
- Hasil akhir proses parsing adalah bentuk dasar {beruang} dan prefiks {ber-} + bentuk dasar {uang}

Bentuk-bentuk yang tidak secara khusus ada dalam bahasa Indonesia seperti bentuk angka, nama orang, dan kata dalam bahasa asing ditulis sebagai *bentuk asing* sebagai hasil dari proses parsing.

Beberapa contoh yang sudah dibahas di atas adalah contoh proses parsing yang dilakukan pada sebuah kata dalam bahasa Indonesia. Perangkat lunak *morphological parser* yang dirancang pada penelitian ini akan dapat memproses tidak hanya kata tapi juga kalimat dan paragraf yang ditulis dalam bahasa Indonesia. Proses parsing pada kalimat dan paragraf memerlukan beberapa langkah tambahan yaitu:

1. Hilangkan tanda baca yang tidak diperlukan dalam proses parsing. Tanda baca yang diperlukan dalam proses parsing hanya tanda baca penghubung kata (-) sebagai tanda hasil proses reduplikasi
2. Gantikan tanda baca yang dihilangkan dengan karakter spasi sebagai tanda pemisah kata
3. Pisahkan setiap kata lalu lakukan proses parsing untuk setiap kata tersebut

DAFTAR REFERENSI

- [1] Chaer, A. (2008) *Morfologi Bahasa Indonesia (Pendekatan Proses)*. Rineka Cipta, Jakarta.
- [2] Sak, H., Gungor, T., dan Saraclar, M. (2008) Turkish language resources: Morphological parser, morphological disambiguator and web corpus. *GoTAL '08 Proceedings of the 6th International Conference on Advances in Natural Language Processing*, Gothenburg, Sweden, 25-27 August, pp. 417–427. Springer-Verlag Berlin, Heidelberg.