



# Morfológia

# Bevezetés

- **Morfológia:** szavakat és belső szerkezetüket tanulmányozza
- **Lexéma:** szó egy adott jelentésben
- **Lemma:** szótő (ragozatlan alak)
  - Főnév: alanyeset, egyes szám (*ház*)
  - Ige: jelen idő, kijelentő mód, E/3. (*megy*)
- **Szóalak:** egy lexéma (ragozott) alakja (*ment*, *házak*)



# Hány szó van egy szövegben?

*Az V. kerület mai döntése szerint a Kossuth téren kivágott fák kötelező pótlásaként Csepelen és a Tabánban fognak fákat ültetni.*

- Attól függ, mi számít szónak...





# Morféma és (allo)morfok

- Morféma: legkisebb nyelvi egység, mely önálló jelentéssel bír
- Morf: egy morféma felszíni (megjelenő) alakja
- Allomorf: azonos morfémának környezetfüggő változatai (*ban/ben, sátor/sátr*)



# Szabad és kötött morfémák

- Szabad morféma: önállóan előfordulhat - lemmák, „rendes” szavak
- Kötött morféma: csak másik morfémával (szótőhöz kapcsolódva) fordulhat elő – toldalékok, speciális szavak (*batka*)

# Affixumok

- Szótőhöz illesztett elemek
- Prefixum: szó elején  
*dis+regard*
- Szuffixum: szó végén  
*ember+rel*
- Circumfixum: szó elején és végén  
*leg+kövé+ebb*



# Affixumok

- Infixum: szó közepén  
*sulat – s+um+ulat* („ír” – „írt”, tagalog)
- Interfixum: összetételekben (fúgamorféma)  
*Geburt + s + Tag*
- Transzfixum  
*k-t-b* „írás”  
*katab* „írt”  
*ki'taab* „könyv”  
*mak'taba* „könyvesbolt”
  - Magyar? K-r-t



- Portmanteau morfolok: egynél több jelentést hordoznak elválaszthatatlanul

*fut + unk* (T/1. alanyi ragozás)

- Szuppletív alakok

*went = goed*

*nincs = nem van*





# Toldalékolás

- **Inflexió:** nem változik a szófaj, jelentés alapvetően nem változik, új szóalakok jönnek létre
- **Deriváció:** változhat a szófaj, jelentés változik, új lexémák jönnek létre
- **Magyarban:**
  - Rag
  - Jel
  - Képző

**szótő + képző(k) + jel(ek) + rag**





# Szóképzés

- Képzők a szótőhöz közvetlenül kapcsolódnak (vagy egy másik képző után)
- Összetett szavak
  - Endocentrikus: az egyik tag jelentése / szófaja meghatározó az egész jelentésének szempontjából (*csokinyúl* = *egyfajta nyúl*)
  - Egzocentrikus: egyik tag jelentése /szófaja sem releváns az egész jelentésének szempontjából (*hiszekegy*)



# Képzők a magyarban 1

| Tag            | explanation                            | example  | POS         |
|----------------|----------------------------------------|----------|-------------|
| FREQ           | frequentative                          | gat      | VERB → VERB |
| MEDIAL         | medial                                 | ódik     | VERB → VERB |
| CAUS           | causative                              | tat      | VERB → VERB |
| PART           | adverbial participle                   | va       | VERB → ADV  |
| PERF_PART      | perfect adverbial participle           | ván      | VERB → ADV  |
| IMPERF_PART    | imperfect adjectival participle        | ó        | VERB → ADJ  |
| FUT_PART       | future adjectival participle           | andó     | VERB → ADJ  |
| PERF_PART      | perfect adjectival participle          | ott      | VERB → ADJ  |
| NEG_PERF_PART  | negative perfect adjectival participle | atlan    | VERB → ADJ  |
| GERUND         | gerund                                 | ás       | VERB → NOUN |
| NEG_MODAL_PART | negative modal adjectival participle   | hatatlan | VERB → ADJ  |
| MODAL_PART     | modal adjectival participle            | ható     | VERB → ADJ  |
| REG_ACT        | regular activity                       | kodik    | NOUN → VERB |
| ABSTRACT       | abstract                               | ság      | NOUN → NOUN |
| MRS            | mrs                                    | né       | NOUN → NOUN |
| DIMIN          | diminutive                             | ka       | NOUN → NOUN |
| ATTRIB         | attributive                            | s        | NOUN → ADJ  |
| MET_ATTRIB     | metonymical attributive                | i        | NOUN → ADJ  |
| INAL_ATTRIB    | inalienable attributive                | jú       | NOUN → ADJ  |
| NEG_ATTRIB     | negative attributive                   | talan    | NOUN → ADJ  |
| TYPE1          | type1                                  | szeru    | NOUN → ADJ  |
| TYPE2          | type2                                  | féle     | NOUN → ADJ  |
| TYPE3          | type3                                  | nemu     | NOUN → ADJ  |
| TYPE_RANK      | type rank                              | rangú    | NOUN → ADJ  |
| NEG_ATTRIB2    | negative attributive2                  | mentes   | NOUN → ADJ  |
| TYPE4          | type4                                  | fajta    | NOUN → ADJ  |
| LOC_INE        | locative inessive                      | beli     | NOUN → ADJ  |
| QUANTITY       | quantity                               | nyi      | NOUN → NUM  |
| ESS_FOR        | essivus formalis                       | képpen   | NOUN → ADV  |
| COM            | comitative                             | stul     | NOUN → ADV  |
| PERIOD1        | period1                                | anként   | NOUN → ADV  |
| PERIOD2        | period2                                | onta     | NOUN → ADV  |
| ACT            | activity                               | oz       | NOUN → VERB |
| ACT2           | activity2                              | ol       | NOUN → VERB |



# Képzők a magyarban 2

|                      |                              |               |             |
|----------------------|------------------------------|---------------|-------------|
| COMPAR               | comparative                  | bb            | ADJ → ADJ   |
| SUPERLAT             | superlative                  | leg-bb        | ADJ → ADJ   |
| SUPERSUPERLAT        | supersuperlative             | legesleg-bb   | ADJ → ADJ   |
| COMPAR_DESIGN        | comparative designative      | bbik          | ADJ → ADJ   |
| SUPERLAT_DESIGN      | superlative designative      | leg-bbik      | ADJ → ADJ   |
| SUPERSUPERLAT_DESIGN | supersuperlative designative | legesleg-bbik | ADJ → ADJ   |
| MANNER               | manner                       | lag           | ADJ → ADV   |
| MANNER               | manner                       | an            | ADJ → ADV   |
| INTRANS_RESULT       | intransitive resultative     | odik/ul       | ADJ → VERB  |
| TRANS_RESULT         | transitive resultative       | ít            | ADJ → VERB  |
| MULTIPL-ITER         | multiplicative iterative     | szor          | NUM → ADV   |
| MULTIPL-ITER         | multiplicative iterative     | szoroz        | NUM → VERB  |
| ITER_ATTRIB          | iterative attributive        | szori         | NUM → ADJ   |
| MULTIPL_ATTRIB       | multiplicative attributive   | szoros        | NUM → ADJ   |
| MULTIPL              | multiplicative               | szorta        | NUM → ADV   |
| AGGREG               | aggregative                  | an            | NUM → ADV   |
| FRACT                | fractional                   | ad            | NUM → NUM   |
| ORD                  | ordinal                      | odik          | NUM → NUM   |
| DATE                 | date                         | odika         | NUM → NOUN  |
| ATTRIB               | attributive                  | i             | POSTP → ADJ |

# Inflexió

- **Rag: szó végén helyezkedik el (lezárja a szót), általában a szó mondatbeli szerepét határozza meg (szintaxis!)**
  - Esetragok: kb. 20
- **Jel: szó jelentését módosítja**
  - Birtokjel
  - Birtokos személyjel
  - Többes szám jele



# 1. Főnevek (Noun) – N

| Pozíció | Attribútum        | Lehetséges értékek                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Kód                                                                                                             | Toldalékok (jelek, ragok)                                                                                                                                                                                                                                                     | Példa                                                                                                                                                                                                                                                                                                            |
|---------|-------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1       | Típus             | neutral                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | n                                                                                                               |                                                                                                                                                                                                                                                                               |                                                                                                                                                                                                                                                                                                                  |
| 2       | Nem               | -                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | -                                                                                                               |                                                                                                                                                                                                                                                                               |                                                                                                                                                                                                                                                                                                                  |
| 3       | Szám              | egyes (singular)<br>többes (plural)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | s<br>p                                                                                                          | Ø<br>-k;<br>-i, -ai/-ei, -jai/-jei;<br>-ék                                                                                                                                                                                                                                    | asztal(om)<br>asztalok,<br>asztalaim<br>szomszédék                                                                                                                                                                                                                                                               |
| 4       | Eset              | alany (nominative)<br>tárgy (accusative)<br>birtokos (genitive)<br>részes(határozó) (dative)<br>eszköz(határozó) (instrumental)<br>"belső közelítő" (illative)<br>"belviszony" (inessive)<br>"távolító" (elative)<br>"külső közelítő" (allative)<br>"közelében levés" (adessive)<br>"távolító külviszony" (ablativ)<br>"ráhelyezés" (sublative)<br>"rajtalevés" (superessive)<br>"eltávolítás" (delative)<br>"határ" (terminative)<br>essive<br>(essive-)formal<br>temporalis<br>causalis<br>sociative<br>factive<br>distributive<br>locativus | n<br>a<br>g<br>d<br>i<br>x<br>2<br>e<br>t<br>3<br>b<br>s<br>p<br>h<br>9<br>w<br>f<br>m<br>c<br>q<br>y<br>u<br>l | Ø<br>-t<br>Ø, -nak/-nek<br>-nak/-nek<br>-val/-vel<br>-ba/-be<br>-ban/-ben<br>-ból/-ből<br>-hoz/-hez/-höz<br>-nál/-nél<br>-tól/-től<br>-ra/-re<br>-n/-on/-en/-ön<br>-ról/-ről<br>-ig<br>-ul/-ül<br>-ként, -képp(en)<br>-kor<br>-ért<br>-stul/-stül<br>-vá/-vé<br>-nként<br>-tt | asztal(om)<br>asztal(oma)t<br>asztalnak<br>asztalnak<br>asztallal<br>asztalba<br>asztalban<br>asztalból<br>asztalhoz<br>asztalnál<br>asztaltól<br>asztalra<br>asztalon<br>asztalról<br>asztalig<br>asztalul<br>asztalként<br>órákor<br>asztalért<br>asztalostul<br>asztallá<br>asztalonként<br>Győrött, Péccsett |
| 5       | Határozottság     | -                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | -                                                                                                               |                                                                                                                                                                                                                                                                               |                                                                                                                                                                                                                                                                                                                  |
| 6       | Klitikum          | -                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | y                                                                                                               | -e                                                                                                                                                                                                                                                                            | katona-e                                                                                                                                                                                                                                                                                                         |
| 7       | Élő/élettelen     | -                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | -                                                                                                               |                                                                                                                                                                                                                                                                               |                                                                                                                                                                                                                                                                                                                  |
| 8       | Birtokos száma    | egyes (singular)<br>többes (plural)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | s<br>p                                                                                                          | -m; -d; -a/-e, -ja/-je, Ø<br>-nk, -unk/-ünk;<br>-tok/-tek/-tök;<br>-uk/-ük, -juk/-jük, -k                                                                                                                                                                                     | asztalom, asztalaim<br>asztalunk, asztalaink                                                                                                                                                                                                                                                                     |
| 9       | Birtokos személye | első<br>második<br>harmadik                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 1<br>2<br>3                                                                                                     | -m; -nk, -unk/-ünk;<br>-d; -tok/-tek/-tök;<br>-a/-e, -ja/-je, Ø;<br>-uk/-ük, -juk/-jük, -k                                                                                                                                                                                    |                                                                                                                                                                                                                                                                                                                  |
| 10      | Birtok(olt) száma | egyes (singular)<br>többes (plural)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | s<br>p                                                                                                          | -é<br>-éi                                                                                                                                                                                                                                                                     | asztalé(t), asztaloké(t),<br>asztalomé(t), asztalaimé(t)                                                                                                                                                                                                                                                         |





## 2. Igék (Verb) – V

| Pozíció | Attribútum                              | Lehetséges értékek                                                                                                                                                                          | Kód                                       | Toldalékok (jelek, ragok, képzők)                                                                                                                                                                                                                                                                      | Példa                                                                                                                                                                                                                                                                                                                                                                                              |
|---------|-----------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1       | Típus                                   | fő (main)<br>segéd (auxiliary)<br>ható (modal)<br>gyakorító (frequentative)<br>műveltető (causative)<br>gyakorító+ható<br>műveltető+ható<br>műveltető+gyakorító<br>műveltető+gyakorító+ható | m<br>a<br>o<br>f<br>s<br>1<br>2<br>3<br>4 | -hAt<br>-gAt<br>-(t)At<br>-gAthAt<br>-(t)AthAt<br>-(t)AtgAt<br>-(t)AtgAthAt                                                                                                                                                                                                                            | fogok (menni)<br>mehetek<br>pofozgat<br>etet<br>boncolgathat<br>fektethet<br>etetget<br>futtatgathat                                                                                                                                                                                                                                                                                               |
| 2       | Mód/forma                               | kijelentő (indicative)<br>felszólító (imperative)<br>feltételes (conditional)<br><br>főnévi igenév (infinitive)                                                                             | i<br>m<br>c<br><br>n                      | Ø<br>-j, -jj, -gy, -ggy, -s<br>-(n)na/-(n)ne, -ana/-ene;<br>-(n)ná/-(n)né, -aná/-ené<br>-ni                                                                                                                                                                                                            |                                                                                                                                                                                                                                                                                                                                                                                                    |
| 3       | Idő                                     | jelen (present)<br>múlt (past)                                                                                                                                                              | p<br>s                                    | Ø<br>-t/-tt/-ott/-ett/-ött                                                                                                                                                                                                                                                                             |                                                                                                                                                                                                                                                                                                                                                                                                    |
| 4       | Személy                                 | első<br><br><br>második<br><br><br>harmadik                                                                                                                                                 | 1<br><br><br>2<br><br><br>3               | -ok/-ek/-ök, -om/-em/-öm<br>-ak/-ek, -am/-em<br>-k; -m<br>-m<br>-am/-em<br>-(u)nk/-(ü)nk<br><br>-sz, -asz/-esz, -ol/-el/-öl<br>-ál/-él, Ø; -l<br>-d<br>-ál/-él<br>-(o)tok/-(e)tek/-(ö)tök<br>-atok/-etek<br>Ø, -ik<br>-on/-en/-ön, -ék<br>Ø, -k<br>-a/-e<br>Ø<br>-(a)nak/-(e)nek<br>-uk/-ük<br>-ak/-ek | várok, eszem<br>várjak, egyem<br>várnék, enném<br>várnom<br>vártam (volna)<br>várunk, várjunk, várnánk,<br>várunk, vártunk (volna)<br>vársz, eszel<br>várj(ál); várnál<br>várnod<br>vártál (volna)<br>vártok, várnátok, várnotok<br>várjatok, vártatok (volna)<br>vár, eszik<br>várjon, egyék<br>várna, ennék<br>várnia<br>várt (volna)<br>várnak, várjanak, várnának<br>várniuk<br>vártak (volna) |
| 5       | Szám                                    | egyes (singular)<br>többes (plural)                                                                                                                                                         | s<br>p                                    |                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                    |
| 6       | Nem                                     | -                                                                                                                                                                                           | -                                         |                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                    |
| 7       | Igenem                                  | -                                                                                                                                                                                           | -                                         |                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                    |
| 8       | Tagadó                                  | -                                                                                                                                                                                           | -                                         |                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                    |
| 9       | Határozottság (alanyi/ tárgyas ragozás) | nem (no)<br><br>igen (yes)<br><br><br><br>-lak/-lekes ragozás                                                                                                                               | n<br><br>y<br><br><br>2                   | l. fentebb (4), kivéve a főnévi igenév ragozását a tárgyas személyragos igealakhoz a <i>valamit</i> névmás helyett az <i>azt</i> névmást tehetjük hozzá mindhárom módban, mindkét időben egyes szám 1. személyű igealakok                                                                              | várlak stb.                                                                                                                                                                                                                                                                                                                                                                                        |
| 10      | Klitikum                                |                                                                                                                                                                                             | y                                         | -e                                                                                                                                                                                                                                                                                                     | jött-e                                                                                                                                                                                                                                                                                                                                                                                             |



# Elemzendő szavak

- *Faxolásaitoknak*
- *Elhozathattátok*
- *Áthatolhatatlanba*
- *Ingatlanügynökséghez*
- *Futtatgathattad*





# Morfológiai elemzés

- **Feladata:** minden szövegszóhoz hozzárendelni az összes lehetséges elemzését (szófaj és egyéb morfológiai jegyek) és a hozzájuk tartozó szótöveket (lemmák)
- **magyar vs. angol**
  - Szótári alapalakok
  - Lehetséges szóalakok száma
  - Lehetséges elemzések (kódok) száma (angol 36 vs. magyar ~1000)
  - Eltárolható-e minden szóalak?



# Morfológiai elemző

- **Lexikon: szótövek és toldalékok**
- **Szabályok: a szóalakok felszíni reprezentációja és a nyelvi elemek szótári reprezentációja hogyan függ össze**
- **Csak zárt szóosztályokra működik jól (névmások, kötőszavak...) – ezek kimerítően felsorolhatók a szótárban**

# Lemmatizálás

- Szótövesítés: a morfológiai elemző feladata – a szóalakot töre (lemmára) és toldalékokra bontja
- Képzés?
- Abszolút és relatív szótő:  
*faxolásaitoknak* (*fax* vs. *faxolás*)
- Mikor melyiket érdemes használni?  
*szolgáltatások – szolgáltatás – szolgáltató – szolga*





# Ismeretlen szavak elemzése

- A szókincs állandó bővülése miatt nincs benne minden szó a szótárban
- Lehetséges toldalékok listájára lehet csak hagyatkozni
- Tartalmaz-e a szóalak jobb oldala lehetséges toldaléktömböket?
  - Ha igen, levágjuk, és a maradékot kezeljük szótóként
  - Az affixumok alapján kap morfológiai elemzést
- Előállítható-e két szótárban meglevő lemmából? (*egérpad*)

# Ismeretlen szavak

- **Lehetnek:**
  - összetételek
  - Névelemek / tulajdonnevek
  - képzések
- **fémkapunk**
- **félmillió**
- **csokinyúl**
- **NATO-hoz**
- **Elemzési módszerek(Zsibrita et al. 2010):**
  - Két vagy több elemezhető részre vágjuk szét
  - Szabályok a nem lehetséges kombinációk kiszűrésére (\*V+N)
  - Az utolsó elem elemzése lesz az egész szóé
  - Kötőjeles szavaknál behelyettesítés



# félmillió

|        |     |           |
|--------|-----|-----------|
| fél    | N   | half      |
|        | ADJ | half      |
|        | NUM | half      |
|        | V   | be afraid |
| millió | NUM | million   |

**fél+millió**  
**NUM**

szabályok:

NUM + NUM

\* non-NUM + NUM



# fémkapunk

|      |   |                   |
|------|---|-------------------|
| fém  | N | metal             |
| kap  | V | get               |
| punk | N | punk              |
| kapu | N | gate              |
| unk  | S | 1PI (verb)        |
| nk   | S | 1PIPoss<br>(noun) |

**fém+kap+unk**  
**VERB**

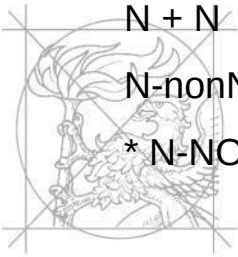
**fém+kapu+nk**  
**NOUN**

szabályok:

N + N

N-nonNOM + V

\* N-NOM + V



# csokinyúl

|        |   |             |
|--------|---|-------------|
| csoki  | N | chocolate   |
| nyúl   | N | rabbit      |
|        | V | stretch     |
| kinyúl | V | stretch out |

szabályok:

N + N

N-nonNOM + V

\* N-NOM + V

**csoki+nyúl**

**VERB**

**NOUN**

**cso+kinyúl (?)**

**VERB**





# NATO-hoz

|      |   |       |
|------|---|-------|
| NATO | ? | NATO  |
| hoz  | V | bring |
|      | S | to    |

szabályok:

N + - + S

N-nonNOM + - + V

\* N-NOM + - + V

V + - + V

helyettesítés:

NATO- -> kalap 'hat'

**NATO-hoz**

**NATO: V**

**VERB**

**NATO-hoz (kalaphoz)**

**NATO: N**

**NOUN**

**Szabályok rendezése:**

**1. helyettesítés**

**2. szegmentálás**