

HIDOL (Hang-idő konverter)

A HIDOL program a beszédhangok időtartamának és a szóhosszak adatainak feldolgozására szolgál (átlagok, eloszlások, összehasonlítások tendenciák meghatározása). A beszédhang-időtartamok mérésénél meghatározhatjuk a mérendő beszédhang környezetét is. Így, előre megtervezett célzott mérések is elvégezhetők. A mérendő beszédhang helyzetét a szón belül is meghatározhatjuk.

A HIDOL program a mérési eredményeket txt kiterjesztésű, tabulátorral tagolt fájlokban adja meg.

Az eredményfájlok minden mérés után felül íródnak. Ezért azokat közvetlenül a mérés elvégzése után célszerű más néven eltárolni.

A program 2004-ben készült, fejlesztői: Olasz Gábor és Abari Kálmán.

A program a hang- és szóidőtartamokat kézzel címkézett speciális formátumú beszédadatbázisból veszi, ezért a kapott mérési adatok pontosnak tekinthetők.

A beszédadatbázismondatai (wav kiterjesztésű fájlok) a beszédadatbázisból meghallgathatók.

A program csak a saját beszédadatbázisával működik!

A program a BME TMIT hangszimbólumait használja, amelyek a betűképnek felelnek meg, kivéve a következőket (betű=hangszimbólum):

á=A.; é=E.; ö=O; ü=U; gy=G; ty=T; ny=N; zs=Z; s=S; sz=s; cs=C.

A dz; dzs; hangok nem szerepelnek a hangtárban. Ezeket a hangok a d;z, illetve a d;Z hangkapcsolatokkal értelmezi a program. A kettőspont fonológiailag hosszú hangot jelöl.

A letöltött állományok jelentése.

1. HIDOL_beszédadatbázis_meresre könyvtár. Ebben a könyvtárban van a teljes beszédadatbázis, valamint a címkézési információkat tartalmazó fájlok. Ezt a könyvtárt kell beolvasni, miután elindítottuk a HIDOL programot. **Ez az első lépés!**
2. A specifikációk elnevezésű könyvtárban példákat adunk meg mérési specifikációkra txt fájlokban (lásd később).
3. A HIDOL .exe fájl a feldolgozó program. Ezt kell elindítani.
4. A hangcímkék.txt tartalmazza az alapvető feldolgozásokhoz meghatározott hangcímkéket és hozzájuk tartozó hangkörnyezeti csoportokat. Ezt a fájlt kell beolvasni a *Hangok címkéi* fültre lépés után. **Ez a második lépés!** Ezt a fájlt csak a HIDOL programmal célszerű tovább bővíteni, azaz más hangcsoportokat definiálni és ellátni címkével (felül kell írni minden egyes bővítés után).
5. A hangkapcsolatok.txt fájl tartalmazza a hangkörnyezeti (eddig definiált) információkat. Ennek betöltése a harmadik lépés! Ezt a fájlt csak a programmal célszerű tovább bővíteni, azaz más hangkörnyezeti változatokat definiálni (felül kell írni minden egyes bővítés után).
6. Az output.txt fájl a mindenkor mérés eredményét tartalmazza. Ha meg akarjuk őrizni, akkor más néven el kell menteni!
7. A hangokkal, szavakkal, mondatokkal kapcsolatos txt fájlok a beszédadatbázis részletezései.

A program használata

Elindítjuk a HIDOL.exe programot, ekkor megjelenik az alapképernyő. Sorban megyünk végig a sorszámozott füleken (1-5).

1. *Bemeneti fájlok.*

A *Könyvtár* gombra kattintva betöltjük a HIDOL_beszédadatbázis_meresre elnevezésű könyvtár tartalmát a programba (több száz mondat). Ezekben az állományokban fogja keresni a program hangokat. A betöltés eredményeképpen a képernyőn megjelenik a betöltött mondatalmaz adatállománya. Ha nincs betöltve beszédadatbázis, a program nem mér semmit.

2. Hangok címkéi.

Erre a fülre kattintva feltűnnek az előre meghatározott hangcímkék és a hozzájuk rendelt hangok, hangcsoportok. Ezeket külön fájlban tároljuk (hangcimkek.txt). Ezt a fájlt célszerű betölteni a *Megnyitás fájlból* nyomógombra kattintva. Töltsük be a hangcimkek.txt fájlt, amely már többféle hangcímkét és a hozzá tartozó hangcsoportot tartalmaz.

A hangcímké elnevezése tetszőleges lehet. A címkével kell hivatkozni a hozzá tartozó hangcsoportra. Például a *PTK* címkéhez definiáltuk a *ptk* hangokat, a *Vr* címkéhez a rövid magánhangzókat stb. Ha hozzáadtunk ehhez a fához új hangcímkét és megadtuk a hozzá tartozó hangcsoportot is, akkor célszerű a *Mentés fájlba* gombbal elmenteni és felülírni a hangcimkek.txt fájlt, hogy a munkánk ne vesszen el.

3. Hangkapcsolatok.

Ezzel a füllel a mérendő hang hangkörnyezetét lehet megadni (már sokféle környezet van eleve definiálva). Három hangnyi területet lehet ezekkel a hangmeghatározásokkal lefedni a mérendő hang előtt is és utána is (összesen 7 hangnyi terület). Így finom hangkörnyezeti hatások is vizsgálhatók. A megjelenő *Hang1*, *Hang2*, *Hang3* oszlopokban egy-egy sor jelenti a hármas hangkörnyezeti definíciót.

Hozzá is adhatunk a listához új elemhármast a képernyő jobb oldali mezőjében található *Hozzáad* gombbal. Ez előtt azonban meg kell adni a *Hang1*, *Hang2*, *Hang3* adatokat. Figyelem! Csak olyan hangcímkék használhatók itt, amelyeket a *Hangok címkéi* pontban lévő hangcimkek.txt fájlban már szerepelnek. Erre a szinkronításra mindig figyelni kell! Az ilyen hozzáadásra csak igen speciális mérések elvégzésekor kerül sor, mivel a program alapértelmezésben 53 féle kombinációt ajánl fel a hangcimkek.txt fájlban.

Példa: a *C_BÁRMI_BÁRMI* hanghármast első eleme (C) azt jelenti, hogy minden mássalhangzót figyelembe vesz a program, a mássalhangzó után azonban bármilyen hang szerepelhet. A *BÁRMI_BÁRMI_C* hanghármast utolsó eleme jelent mássalhangzót. Ezt a két hanghármast fogjuk felhasználni például CVC típusú hangkörnyezeti mérésekhez a specifikáció pontban.

4. Specifikációk.

Itt specifikáljuk a tényleges mérést. A specifikáció típusa lehet hangra vonatkozó (H), illetve szóra vonatkozó (SZ). Példaképpen megadunk számos specifikációt txt fájlok formájában a specifikációk elnevezésű alkönyvtárban. Ezek közül bármelyik betölthető a *Megnyitás fájlból* gombra kattintással. A megadott specifikáció fájlok jelentése a következő.

sp_BBB_AaouUiEOe_BBB.txt= a megadott magánhangzók hosszának átlagát és eloszlását számolja ki bármilyen hang áll előttük és mögöttük.

sp_BBB_bpdtgkgyty_BBB.txt= a megadott mássalhangzók hosszának átlagát és eloszlását számolja ki bármilyen hang áll előttük és mögöttük.

sp_BBB_C_Cr_Ch_BBB.txt= az összes mássalhangzó, valamint az összes rövid és hosszú mássalhangzó hosszának átlagát és eloszlását számolja ki bármilyen hang áll előttük és mögöttük.

sp_BBB_mnny_BBB.txt= az m, n, ny, hangok hosszának átlagát és eloszlását számolja ki bármilyen hang áll előttük és mögöttük.

sp_BBB_V_BBB.txt= az összes magánhangzó hosszának átlagát és eloszlását számolja ki bármilyen hang áll előttük és mögöttük.

sp_Cr+C_Cr+V_egyenkent.txt= a rövid mássalhangzó hosszának átlagát és eloszlását számolja, ha mássalhangzó követi, illetve ha V követi.

sp_szo_1_5szg_első_belső.txt= a szavak hosszának átlagát adja meg 1,2,3,4 és 5 szótagú, mondatkezdő és mondatbelseji szavakra.

sp_szo_2sztg_belső.txt= két szótagú szavak hosszának átlagát és eloszlását számolja ki mondatbelseji szavakra.

sp_V+lrC_V+CV_hasonlitasa.txt= a V hosszának átlagát adja meg hangonként, ha l+C, illetve r+C követi, valamint a V átlagát, ha csak C követi.

Gyakorlásképpen az *Új specifikáció* gombra kattintva hozzunk létre egy specifikációt például a V-k hosszának megmérése CVC környezetben. Ehhez az *Előtte*, *Hang*, *Utána* oszlopokba kell adatot definiálnunk (választanunk).

Előtte: legörgetve kikeressük a *BÁRMI_BÁRMI_C* hanghármast, ezt választjuk ki (ez lesz a V előtti hanghármast).

Hang: legörgetve kikeressük a V-t, ezt választjuk ki.

Utána: legörgetve kikeressük a *C_BÁRMI_BÁRMI* hanghármast, ezt választjuk ki (ez lesz a V-t követő hanghármast).

Pozíció a szóban: a mérendő hang helyzetét a szón belül is meghatározhatjuk. Alapértelmezés: *BÁRMI*, ezt válasszuk ki.

Ezzel specifikáltuk a mérés terét. Minden CVC helyzetű magánhangzó hosszát ki fogja gyűjteni a program az adatbázisból. Mindezekkel definiáltuk a mérendő hangot és a hangkörnyezetét. Ezt a specifikációt elmenthetjük tetszőleges névvel.

Meg kell adni a specifikáció nevét (ezzel a névvel elmenthetjük egy txt fájl formájában, hogy a későbbiek során már készen behívassuk).

Legyen ennek a specifikációnak a neve: *C_V_C*. Az eredményt ezzel a névvel fogja megadni a program az *output.txt* fájlban az alábbiak szerint.

```
SPEC.NEV: C_V_C
ÖSSZ (db) : 4569
ÁTL. (ms) : 85
MIN. (ms) : 23
MAX. (ms) : 250
```

Ezeken szerint 4569 db CVC helyzetű magánhangzó van a beszédatadbázisban, ezek hosszának átlaga 85 ms.

5. Kimeneti fájlok.

Ezzel a füllel végezzük el a mérést, amely három lépésből áll (mindegyiket külön el kell indítani).

1. *A konverzió indítása*. IGEN válasz után elvégzi a beszédatadbázis összes adatának rendezését. **Ezt csak egyszer kell elvégezni egy futtatáson belül.**
2. *A lekérdezés indítása*. IGEN válasz után itt végzi el a program a specifikáció szerinti méréseket és összegyűjti az adatokat. **Ezt minden új specifikáció esetén le kell futtatni.**
3. *Az eredmény megtekintése*. Itt tudjuk megnézni a végeredményt, amelynek három részletezettségi foka van (átlag, eloszlás, részletezett adatok). Az előbbi mérés átlagát a fent megmutatott 5 sor képviseli.

Az eloszlást is bejelölve, a „megmutat” utasításra a következő sorokat kapjuk (10 ms-os sávokba összesítve).

```
SPEC.NEV: C_V_C
ÖSSZ (db) : 4569
ÁTL. (ms) : 85
MIN. (ms) : 23
MAX. (ms) : 250
ELOSZLÁSOK (db) :
0- 10: 0
10- 20: 0
20- 30: 8
30- 40: 46
40- 50: 198
```

50- 60: 474
60- 70: 750
70- 80: 814
80- 90: 674
90- 100: 504
100- 110: 348
110- 120: 255
120- 130: 203
130- 140: 120
140- 150: 84
150- 160: 39
160- 170: 30
170- 180: 11
180- 190: 5
190- 200: 2
200- 210: 2
210- 220: 0
220- 230: 0
230- 240: 0
240- 250: 2

Látható, hogy a 70-80 ms-os sávban találta a legtöbb magánhangzót (814 db) a program. Ha az eloszlásokat is bejelöljük a jobb alsó négyzetben, akkor a mért adatok 10ms-os sávokban való eloszlása is megjelenik.

A részletezett adatok listáját is megkaphatjuk (bal alsó jelölő négyzet). Ilyenkor a mérések pontos helyét (melyik mondat melyik szavában melyik hang) és az egyedi adatokat adja meg a program. Így bármely mért adatnak pontosan utána lehet nézni.

Szóhosszak mérése

A szóhosszak mérését a *Specifikációk* fülön állíthatjuk be. A szavakat csoportosíthatjuk szótagszám szerint, és hogy van-e bennük hosszú magánhangzó, illetve, hogy a hosszú magánhangzókból hány van a szóban.

A *Szövegszó* jelölő négyzetnek két jelentése van. Ha bejelöljük a pipát, akkor a határozott névelőket nem veszi be a mérésbe. Ha nincs bejelölve a pipa, akkor csak a határozott névelőket válogatja ki a program (egy szótagú szavak mérésénél ad választási lehetőséget).

Példaként megadjuk a két szótagú szavak átlagos hosszának mérési beállításait (lehet bennük bármilyen hosszú magánhangzó is). Szűkítést jelent, hogy csak mondat belseji szavakat keresünk. Ennek a mérésnek a specifikációja: *szótagszám=2; Pozíció a mondatban: belső; szövegszó=* bejelölve, továbbá a hosszú magánhangzók ablakaiban a *BÁRMENNYI* jelölés látszik.

Az eredmény:

SPEC.NEV: 2sztg_szo_belso
ÖSSZ (db) : 509
ÁTL. (ms) : 394
MIN. (ms) : 172
MAX. (ms) : 767