

HIDOL - Speech Sound Duration Measuring Program (SSDMP)

The HIDOL program serves for measuring and giving statistics on speech sound-durations. When measuring speech sound duration, the sound environment of the measured sound can be adjusted as well. The position of the measured speech sound can also be determined within the word. Word length can be measured as well. Output statistics are: averages, distributions, comparisons, trends). Results of the measure are given in a txt file.

The program measures the speech sound- and word durations from a manually tagged Hungarian speech database.

The program was created in 2004 in Budapest. Developers are: Gábor Olaszky and Kálmán Abari. This was the first computer-aided speech sound duration analysis program and data environment for Hungarian speech.

The speech database contains read sentences in the pronunciation of a native Hungarian male speaker. The sentences in the speech database (files with a wav extension) can be listened directly from the speech database.

The program only works with its own speech database!

FOR USING THIS PROGRAM IT IS GOOD IF YOU ARE FAMILIAR WITH HUNGARIAN LANGUAGE!

The program uses the BME TMIT speech sound symbols that correspond to the Hungarian font image, except the following ones (Hungarian letter(s) = speech sound symbol):

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

The dz, dzs letters are not in the sound library. The corresponding speech sounds are labelled with the combinations ao d+z and d+Z Hungarian letters. The colon indicates a phonologically long sound (for example b:, or O:, or U:).

The downloaded files.

1. HIDOL_beszedatabazis_meresre directory. This directory contains the complete speech database as well as files containing tagging information. This directory must be read after starting the HIDOL program. This is the first step when HIDOL runs already!
2. In the directory named *specifikaciok*, we give examples of measurement specifications (see later).
3. The HIDOL.exe file is the processing program. **This needs to be started.**
4. The *hangcimkek.txt* file contains the speech sounds defined for processing and their associated sound environment groups. This file must be read after going to the *Hangok címkéi* window. **This is the second step!** It is possible to expand this file, but only with the HIDOL program.
5. The *hangkapcsolatok.txt* file contains the sound environment information (defined so far more than 50 variants). Loading this is the third step! Advice: if you want to expand this file, do it only with the HIDOL program.
6. The *output.txt* file contains the result of the current measurement. To preserve your results, save this file under a different name!
7. Other txt files related to sounds, words, sentences are components of the speech database. Do not change them!

Using the HIDOL program

Start HIDOL.exe and the basic screen will appear. We line up along the numbered upper tabs (1-5).

1. *Bemeneti fájlok. (Input files).*

Clicking the *Könyvtár (Library)* button the contents of the directory named *HIDOL_beszedadatbázis_meresre* (*HIDOL_speech database directory*) will be loaded (hundreds of sentences with hand made labells). The program will search for the measuring items in these files. As a result of the loading, the data files of the loaded sentences is displayed on the screen. If no speech database is loaded, the program will not measure anything.

2. Hangok címkéi. (Sound tags).

Clicking on this tab will bring up the predefined speech sound tags and the sounds associated with them. These are stored in the separate *hangcimkek.txt* file (*voice tags.txt*).

Click on window *Megnyitás fájlból (Open from file)* and load the *hangcimkek.txt* (*voice tags*) file, which already contains several voice tags and the corresponding voice groups. The name of the voice tag can be arbitrary. The label must refer to the corresponding voice group. For example, we defined *ptk* speech sound group, let we define the *PTK* tag. For short vowels let it be the *Vr* tag, and so on. If we added a new voice tag to this tree and also specified the corresponding voice group, then it is advisable to save and overwrite the *hangcimkek.txt* (*voice tags*) file with the *Mentés fájlba (Save to file)* button so that our work don't get lost.

3. Hangkapcsolatok (speech sound's connections).

With this window you can specify the sound environment of the speech sound to be measured (there are already many environments defined). Three sounds can be defined before the measured speech sound and also three ones after it (totally a group of 7 speech sounds). Thus, subtle sound-environmental effects can also be investigated. In the *Hang1 (Sound1)*, *Hang2 (Sound2)*, *Hang3 (Sound3)* columns one row represents the triple sound environment definition.

You can also add a new triplet to the list by clicking the *Hozzáad (Add)* button in the field on the right side of the screen. Before this, however, the data for *Hang1 (Sound1)*, *Hang2 (Sound2)*, *Hang3 (Sound3)* must be entered. Attention! Only voice tags that are already in the *hangcimkek.txt* (*Voice tags*) file can be used here. You always have to pay attention to this synchronicity! Such extra additions are advised only when performing very special measurements, as the program offers 53 combinations in the *hangcimkek.txt* (*voice tags*) file by default.

Example: the first element (C) of the *C_BÁRMI_BÁRMI (C_ANY_ANY)* string means that all consonants will be measured independently of the speech sounds after it.

The last element of *BÁRMI_BÁRMI_C (ANY_ANY_C)* soundtrack means any consonant (C) that can be preceded by any sound. These two triads will be used, for example, for CVC-type sound environment measurements in the *Specifikációk (specification)* section.

4. Specifikációk. (Specifications).

Here we specify the actual measurement. The specification type can be speech sound based (H) or word (SZ) based. Speech sound is the default. A number of specifications are given in the *specifikaciok (specifications)* directory in txt format.

Any of these can be loaded from using the *Megnyitás fájlból (Open file)* button. The meaning of the recommended specification files is as follows.

sp_BBB_AaouUIEOe_BBB.txt = average and distribution of Hungarian short vowel lengths, independently whatever sound is in front of and behind them.

sp_BBB_bpdtgkgyty_BBB.txt = the average and distribution of the length of the corresponding consonants independently whatever sound is in front of and behind them.

sp_BBB_C_Cr_Ch_BBB.txt = average and distribution of Hungarian consonants and all short and long ones lengths, independently whatever sound is in front of and behind them.

sp_BBB_mnny_BBB.txt = calculates the average and distribution of the lengths of m, n, ny speech sounds independently whatever sound is in front of and behind them.

sp_BBB_V_BBB.txt = calculates the average and distribution of the length of all vowels independently whatever sound is in front of and behind them.

sp_Cr + C_Cr + V_egyenként.txt = calculates the mean and distribution of the length of the short consonant, if consonant follows it or V follows it.

sp_szo_1_5szg_első_belso.txt = gives the average length of words in inside position in the sentence and having 1,2,3,4 and 5 syllables,

sp_szo_2sztg_belso.txt = calculates the average and distribution of the length of two syllable words inside the sentence.

sp_V + IrC_V + CV_hasonlitas.txt = calculates the average length of V if I + C sounds follow it, or r + C follow the V, or only any consonant follows the V.

In practice, click the *Új spec (New Specification)* button to create a specification. Let we measure the length of the vowels in general in CVC environment. To do this, we need to define (select) data in the *Előtte, Hang, Utána (Before, Sound, After)* columns.

Előtte (Before): scroll down to find *BÁRMI_BÁRMI_C (ANY_ANY_C)* soundtrack and select this (this will be the sound environment before the V).

Hang (Sound): scroll down to find V, select this.

Utána (After): scroll down to find *C_ BÁRMI_BÁRMI (C_ ANY_ ANY)* triad, select it (this will be V following triad).

Pozíció a szóban (Position in the word): the position of the sound to be measured can also be determined within the word. Default: *BÁRMI, (ANY)*, let's choose this.

The measurement area is specified. The length of each vowel in CVC position will be extracted from the database. With all this, we defined the sound to be measured and the sound environment. You can save this specification with any name typing in the *Specifikáció neve (Name of Specification)* window.

You must enter a name for the specification (with this name you can save it as a .txt file. Let this specification be named *C_V_C*. The result with this name will be given in the *output.txt* file as follows.

```
SPEC.NEV: C_V_C
ÖSSZ (db) : 4569          (the number of measure)
ÁTL. (ms) : 85           (average)
MIN. (ms) : 23           (min.)
MAX. (ms) : 250          (max.)
```

According to them, there are 4569 CVC vowels in the speech database, the average length of which is 85 ms.

5. *Kimeneti fájlok Output files.*

Click on *Kimeneti fájlok* tab to perform the measurement, which consists of three steps (each must be started separately).

1. *A konverzió indítása (Start the conversion)*. After answering *Igen (YES)*, the program sorts all the data in the speech database. You only need to do this once per run.
2. *A lekérdezés indítása (Start measuring)*. If *Igen (YES)*, the program performs the measurements according to the specification here and collects the data. This must be run for each new specification.
3. *Az eredmény megtekintése (View the result)*. Here we can see the end result in three forms (mean, distribution, detailed data). The average of the former measurement is the default screen. To ask for the distribution check the *Eloszlás megjelenítése* box and click on *Megmutat (show)* window. The total distribution will be shown in 10 ms bands.

```
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

It can be seen that the program found the most vowels (814 pieces) in the 70-80 ms band.

You can also get a list of detailed data (lower left check box: *Részletek megjelenítése*). In this case, the exact location of every measurement (in which sentence in which word is the sound) and the individual data are given by the program. This way, any measured data can be viewed exactly.

Word length measurement

You can set the word length measurement clicking on the *Specifikációk (Specifications)* tab and inside it on the *Új spec. (New spec.)* window. Select the (SZ) point.

Words can be grouped by syllable number *Szótagszám* and whether they contain long vowels. The number of A: or E: or *Egyéb hosszú magánhangzó - any other long vowel* can be defined in the appropriate windows.

The *Szövegszó (Text word)* check box has two meanings. If the check mark is selected, the definite adjectives will not be included in the measurement. (This gives a choice when measuring single-syllable words).

The *Pozíció a mondatban (Position in the sentence)* check box has four position markers concerning the words: *ELSŐ (first)*, *BELSŐ (inside)*, *VESSZŐ ELŐTT (before comma)* *UTOLSÓ (last)*.

As an example, we give the settings for measuring the average length of two - syllable words (any long vowel is accepted in the word, *Egyéb hosszú magánhangzó: BÁRMENNYI*). Let we narrow the search for words inside the sentence (*Pozíció a mondatban: BELSŐ*).

The specification of this measurement is: syllable number = 2; Position in sentence: internal; text word = checked, *BÁRMENNYI (ANY)* is displayed in the long vowel windows.

The result:

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

It can be seen that the program found 509 such two syllable words in the speech database. The average word length is 394 ms. The minimum length is 172 ms, the maximum is 767 ms.