

Magyar hangsúly-adatbázis az interneten kutatáshoz és oktatáshoz

Abari Kálmán¹, Olaszy Gábor²

¹ DE Pszichológiai Intézet, Szociál- és Munkapszichológiai Tanszék
abari.kalman@arts.unideb.hu

² BME Távközlési és Médiainformatikai Tanszék
olaszy@tmit.bme.hu

Kivonat: Hangsúlyadatbázis eddig nem készült magyar nyelvre. Beszédtechnológiai kutatásokban és az oktatásban is nagy igény lenne egy referenciaként használható, helyes hangsúlycímkéket tartalmazó mondatgyűjteményre. Fejlesztésünkkel ezt a hiányt kívántuk pótolni. A webes lekérdezőfelülettel rendelkező hangsúlyadatbázisunk 1869 kijelentő mondatot tartalmaz, amelyekben minden szó hangsúlypozícióját megjelöltük. A honlapon számos szempont alapján kereshetünk és a találati lista is több formában (szöveges, kép és hang) jeleníthető meg. A vizsgált magyar mondatok hangsúlymintáztatának gyakorisága is lekérdezhető. A honlap a <http://magyarbeszed.tmit.bme.hu/hangsuly> címen érhető el.

1. Bevezetés

Az itt bemutatott adatbázis a magyar hangsúlyozás szöveges tartalomra való előrejelzéséhez nyújt interaktív támogatást. Az adatbázis hangsúlycímkékkel ellátott magyar mondatokat tartalmaz szöveges formátumban, melyekhez képi megjelenítések és a meghangosított mondatok hangállományai társulnak. Az adatbázis olyan mondatkorpuszra támaszkodik, amelynek két kiindulópontja van. Az alapot egy korábbi gépi beszédfelismerési kutatáshoz alakították ki [5] úgy, hogy fonetikailag kiegyensúlyozott mondatokhoz hoztak létre irodalmi művek szövegeiből. Az adatbázis másik forrása a BME TMIT-en erre alapozott és elkészített párhuzamos, precíziós beszédadatbázis 12 beszélővel [2]. Mindezekből logikusan adódott az a gondolat, hogy erre a mondatmátrixra alapozva elkészítsük a mondatok hangsúlyozási jelekkel kibővített szöveges változatát is, ami egyfajta támaszt adhat későbbi hangsúlykutatásokhoz, valamint felhasználható az oktatásban is. A munka 3 évig tartott. Az adatbázis 1869 kijelentő mondatot tartalmaz. Mivel a hangsúlyozással kapcsolatos nyelvészeti irodalom szerteágazó, el kellett döntenünk, hogy milyen formában közelítünk a témához. A hangsúly jelölésének a legegyszerűbb változatát választottuk, bináris felépítésben gondolkodtunk, vagyis azt jelöltük, hogy van hangsúly (W) vagy nincs hangsúly (N). A másik egyszerűsítésünk, hogy csak szóhangsúlyokat jelölünk a szövegben, ezt is következetesen, azaz minden szó kap egy W vagy N címkét. (Itt megjegyezzük, hogy ezeket a címkéket csak a kutatható adatbázisban láthatja a felhasználó, a mondatlistá-

ban a W címkét kiemeléssel helyettesítjük, ahogy ezt a formát alkalmazzuk e tanulmány sok példájában is a könnyebb olvashatóság kedvéért. A hangsúly mintázatok bemutatásánál pedig a W-t H jelöléssel helyettesítjük, az N-t pedig a - karakterrel.) A harmadik egyszerűsítésünk, hogy a hangsúlyok fizikai kivitelezésénél csak az alaphangfrekvenciát használtuk a hangsúly élmény megvalósításához, úgy, hogy az első szótagi magánhangzókön erőteljes alaphangfrekvencia (F0) emelést hajtottunk végre, majd a második magánhangzóig visszavezéreltük az alaphangfrekvenciát a mondat dallamvonulatát meghatározó alapra. A negyedik egyszerűsítés a mondatok szöveges ábrázolását érinti. A feldolgozás menete során csak kisbetűket alkalmazunk, így a mondatlista mondataiban minden karakter csak kisbetűvel szerepel, és mondatvégi írásjel sincs a mondatok végén.

A hangsúlyadatbázis referencia vizsgálatát külön kutatásban végeztük el [3], az interneten közreadott változat már tehát egyfajta referenciának tekinthető, amely közvetlenül tanulmányozható három formában. Ezek közül a legmegfoghatóbb a szövegbe elhelyezett hangsúlycímkék állománya. A címkék és a mondat tartalma lehetőséget ad keresésekre és csoportosításokra is. A hangformátumot beszédszintetizátorral állítottuk elő az adott mondat hangsúlyjelei szerint. Ez adja az akusztikai megjelenítést. Az F0 változást képből is megjelenítjük, tehát az összevethető a hangzó és az írott, címkézett formával. Az adatbázis hangsúly jelölése nagy pontosságúnak tekinthető, amit úgy kell érteni, hogy nincs benne címkézési hiba, vagyis ahol hangsúlyt jelöltünk az adott szóra, ott a hangsúlyos ejtés nem okoz megértési zavart, furcsa ejtést és fordítva. Vannak olyan mondatok, amelyek többféle hangsúlykiosztással is ejthetők az értelmezés, illetve a közlési szándék szerint. Ezeknél a mondatoknál az egyik helyes formát adják a jelölések.

Az adatbázis érdekessége, hogy kontrasztokat is bemutat hangban, tehát az érdeklődő tanulmányozhatja a jó hangsúlyozással megvalósított mondatot, valamint ugyanannak a mondatnak két másik változatát is. Az egyikben a hangsúlyozást létrehozó alaphangfrekvencia csúcsokat megszüntettük, ez egyfajta neutrális szerkezetet eredményez, ami érzeti szempontból nem biztos, hogy élesen érzékelhető a hangzásban. A másik kontrasztban a rossz hangsúlyozást próbáltuk megvalósítani, többnyire megfordítottuk a hangsúly kiosztás címkéit, azaz azokra a szavakra tettünk hangsúlyt (F0 csúcsot), amelyek az eredeti mondatban hangsúlytalan (N) címkével voltak ellátva. Mindhárom forma tanulmányozható az adatbázisban. Példákat az 1. táblázat tartalmaz.

1. táblázat: Példák az adatbázis elemeire

A) jó hangsúlyozás	[:N]a [:W]híradások[:W]annak[:N]idején [:W]röviden [:N]számoltak [:N]be [:N]az [:N]ügyről
B) neutrális forma	[:N]a [:N]híradások[:N]annak[:N]idején [:N]röviden [:N]számoltak [:N]be [:N]az [:N]ügyről
C) rossz hangsúlyozás	[:N]a [:N]híradások[:N]annak[:W]idején [:N]röviden [:W]számoltak [:W]be [:N]az [:W]ügyről

Fontos tudni, hogy a fenti három változatban a hangzó mondat szegmentális szerkezete (hangidőtartamok, szünetek) ugyanaz, csak az alaphérfvencia csúcsok meglétében/hiányában, illetve a helyében (melyik szón van) különböznek a mondatok fizikai megvalósításai. Az ilyen célzott F0 változtatásokat a Profivox beszédszintetizátor speciális alkalmazási lehetősége biztosította [4]. Az összes mondat mindhárom változatát kézi feldolgozással készítettük el.

Az adatbázis webes lekérdező felülettel is rendelkezik. A honlap (<http://magyarbeszed.tmit.bme.hu/hangsuly>) minden funkciója 2014 januárjától érhető el.

2. A hangsúlyadatbázis szerkezete

A hangsúlyadatbázis fő komponense egy MySQL adatbázis, amely az 1869 mondat három különböző hangsúlyozással címkézett szöveges formáját tartalmazza. Az SQL adatbázist WAV és PNG állományok egészítik ki, amelyek a mondatok meghallgatását és a képek megjelenítését teszik lehetővé. A hangsúlyadatbázis utolsó komponense a PHP/HTML forráskódú állományok gyűjteménye, amely a honlap oldalainak megjelenítéséért és a keresés megvalósításáért felelős.

3. A honlap felépítése

Az Első magyar hangsúlyadatbázis az interneten bárki számára hozzáférhető, használatához egy böngésző szükséges. Az adatbázisból a keresőgépekhez hasonló, könnyen kezelhető felületen keresztül kapjuk meg a helyesen hangsúlyozott mondatok listáját, de a honlap sok egyéb funkciót is tartalmaz.

A honlap funkcionálisan 4 részből áll: 1) keresés az adatbázisban, 2) mondatok listája, 3) leírás a kutatható adatbázisról és 4) segítség a honlap használatához. Az adatbázis keresési lehetőségeiről a következő fejezetben részletesen beszámolunk. A honlap második, mondatokat listázó része az adatbázis 1869 mondatáról ad teljes áttekintést: az összes mondat (egyfajta) helyesen hangsúlyozott listáját mutatja meg. A honlap harmadik részében a kutatásra ingyenesen elérhető adatbázisról kapunk tájékoztatást. A honlap használatáról – hangsúlyosan a keresőfelület működéséről – az utolsó, 4-es pontban találunk információt. A következőkben csak a keresési lehetőségeket mutatjuk be.

3.1. Keresés az adatbázisban

A hangsúlyadatbázis webes keresőfelületét az 1. ábra mutatja be. A hét részre tagolt felhasználói felület első négy pontja az 1869 mondat szűrésére, azaz a találati lista több szempontú szűkítésére használható, míg az 5. pontban a találati lista megjelenési módjaiból választhatunk. A 6. pontban a találati lista rendezettségét állíthatjuk be, a 7.

pontban pedig a keresést indíthatjuk el. A következő hét alfejezetben az 1. ábra hét pontját mutatjuk be részletesen.

3.1.1. Keresés betűsor alapján

Tetszőleges karaktersorozat megadásával az 1869 mondat ortografikus karaktereiben végezhetünk keresést. Rákereshetünk egy korábban vizsgált teljes mondatra pl. az *a világosság felé fordult, és belebámult az üveg papírnyomóba* keresőkérdésre egyetlen mondatot fog tartalmazni a találati lista (*a világosság felé fordult, és **belebámult** az üveg papírnyomóba*). A világ keresőkérdésre egy 27 elemű találati lista a válasz, amely a fenti mondatot éppúgy tartalmazza, mint pl. az *a mai világban nem sikk betegeskedni* mondatot is. A keresés során a keresési mezőbe gépelt karaktersorozatokat tehát úgy értelmezzük, hogy azt tetszőleges karaktersorozat előzheti meg vagy követheti.

1. **Betűsor alapján:**

2. **Szó alapján:** Hangsúlyozása: Nem állítom be
 Hangsúlyos
 Nem hangsúlyos

3. **A mondat hossza (szószám):**
 szószám: -től -ig
 konkrét szószám:

4. **Hangsúlyok száma és helye a mondatban:**
 Első szó:
 Hangsúlyos belső szavak száma: -től -ig
 Utolsó szó:

5. **Megjelenítés** Mondatok listájának megjelenítése Összesítés hangsúlymintázatokra
 A neutrális (B) és egyfajta rossz (C) hangsúlyozású mondatok megjelenítése.
 Grafikus megjelenítés.
 A teszt eredményének megjelenítése (CMOS értékek).
 Mondatok meghallgatása.

6. **Rendezés**
 növekvő mondat betűsora szerint
 csökkenő szószám szerint
 hangsúlyszám szerint
 CMOS érték szerint (AB és AC mondatok is)
 CMOS érték szerint (csak AB-s mondatok)
 CMOS érték szerint (csak AC-s mondatok)
 válaszsám szerint (AB és AC mondatok is)
 válaszsám szerint (csak AB-s mondatok)
 válaszsám szerint (csak AC-s mondatok)

7. **Keresés indítása**

1. ábra. Az Első magyar hangsúlyadatbázis honlapjának keresőfelülete

3.1.2. Keresés szó alapján

Ebben a pontban tetszőleges szó előfordulására kereshetünk rá, miközben a szó hangsúlyhelyzetét is beállíthatjuk. Választhatunk hangsúlyos és hangsúlytalan előfordulások között, valamint dönthetünk úgy, hogy nem vesszük be a szűrőfeltételbe ezt az opciót. Például a *több* szó hangsúlytalan pozícióban 3 mondatban fordul elő (az egyik mondat: *ez pedig **nem több egy közepes nyugati** egyetem költségvetésénél*), míg hangsúlyos helyzetben 10 találat jelenik meg (az egyik ezek közül: ***öt nap alatt több mint kilencven** órát dolgozott*). A keresés a keresőmezőbe írt karaktersorozat pontos előfordulásán alapul.

3.1.3. Keresés a mondat hossza alapján

Az adatbázisban előforduló szavak száma 2 és 14 között változik. Ebben a pontban a szavak száma alapján szűkíthetjük a találati listát. Erre két módunk van. A tartomány alapú mondathossz-beállítás a legkisebb és legnagyobb szószám megadását követeli meg. A másik lehetőség konkrét szószám megadása. Ebben az esetben csak az itt beállított szószámmal rendelkező mondatok jelennek meg a találati listában.

3.1.4. Keresés a hangsúlyok száma és helye alapján

A hangsúlyok számának és helyének beállítása az előző pontban specifikált szószám megadásától függ. Amennyiben ez tartomány alapú, akkor a hangsúlyok helyét a mondat három pozíciójában, a mondat első és utolsó szavában, illetve a mondat belsejében állíthatjuk be. Külön dönthetünk tehát az első és utolsó szó hangsúlyos vagy hangsúlytalan pozíciójáról, illetve a mondatbelseji hangsúlyos szavak számáról. Ez utóbbi egy intervallum megadásával lehetséges. Az 1. ábra 3. pontja ezt a hangsúlymegadási formát mutatja. Tegyük fel, hogy a 3. pontban a szószám intervalluma 2–5, a 4. pontban pedig az első és az utolsó szó is hangsúlyos és a hangsúlyos belső szavak tartománya 1–1. A 8 elemű találati lista ekkor tartalmazza a *szeretném, ha néhány percre elfordulna* és az *üresre facsarunk, aztán megtöltünk önmagunkkal* mondatokat is.

Ha a 3.1.3. pontban konkrét szószámot állítunk be, akkor a hangsúlyok pontos, szavankénti megadására is lehetőségünk van. A 2. ábra a keresési felület 4. pontját emeli ki hét szavas szószám megadás esetén. A hangsúlyok számát a teljes mondatra specifikálhatjuk egy intervallum megadásával. Mivel az adatbázis összes mondatára vonatkozóan a hangsúlyszámok 1 és 8 között változnak, a 2. ábrán látható beállítás nem jelent szűkítést. Az újdonság az ábra további részében figyelhető meg. A mondat mind a hét szavára beállíthatjuk a hangsúlyos vagy hangsúlytalan pozíciót, illetve eltekinthetünk az opció beállításától. A 2. ábra alapján a találati listában csak azok a mondatok jelennek meg, amelyekben a 2., 4. és 7. szó hangsúlyos, a 3. pedig nem hangsúlyos. A többi szó hangsúlypozíciója tetszőleges lehet. A találati lista 7 mondatot tartalmaz (két példa a listából: *a szegénység és betegség együtt járása közsímet,* illetve *ehhez tegnap a tőzsdetanács hozzá is járult*).

4. **Hangsúlyok száma és helye a mondatban:**

Hangsúlyok száma: -től -ig

Szó sorszáma 1. 2. 3. 4. 5. 6. 7.

Nem állítom be

Hangsúlyos

Nem hangsúlyos

2. ábra. A keresési felület 4. pontja konkrét szószám megadása esetén

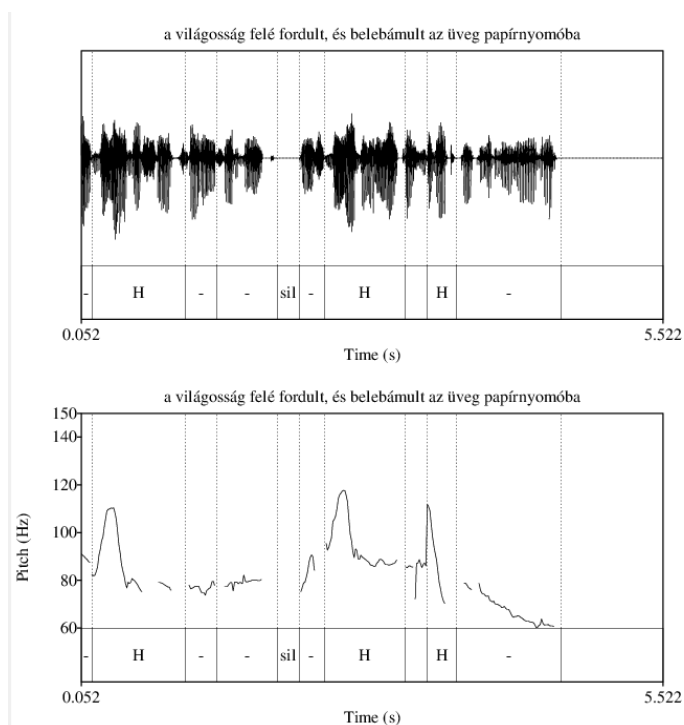
3.1.5. A megjelenítés beállításai

A találati lista az előző négy pont keresőfeltételei alapján áll össze. Alapesetben a találati lista elemei mondatok, melyek megjelenését ebben a pontban határozhatjuk meg (l. 1. ábra 5. pont). Ez az alapértelmezett *Mondatok* listájának megjelenítése opció választásával kezdeményezhető. A másik lehetőség a találatok megjelenítésére, hogy a hangsúlymintázatokra összesítve kérjük a mondatokat (*Összesítés hangsúlymintázatokra* opció). Ez utóbbi lehetőséget az alfejezet végén tárgyaljuk.

A mondatok listája alapértelmezetten a keresés során kiválogatott helyes hangsúlyozású mondatokat tartalmazza (a hangsúlyadatbázisból). Azonban kezdeményezhetjük ugyanazon mondatra a neutrális (B) és egyfajta rossz hangsúlyozású (C) változat megjelenítését is, melyek a helyes hangsúlyozású mondat alatt a B és C pontok után jelennek meg a találati listában (4. ábra).

A grafikus megjelenítési opció segítségével a mondat szerkezetének és a hangsúlykiosztásnak az összefüggéseit lehet tanulmányozni a hangsúlyozás fizikai megvalósulásának tükrében. A találati lista minden eleme ekkor tartalmazza rezgésképet és az alapfrekvencia görbét szinkron megjelenítésben, szóhatárokkal és a hangsúlycímkékkel kiegészítve (3. ábra). A H címke a hangsúlyos, a – címke a hangsúlytalan szavakat jelöli. A mondatbelseji szüneteket a *sil* karaktersorozattal jelöltük.

A mondatok mindhárom típusa meghallgathatóvá tehető a megfelelő jelölőnégyzet bekapcsolásával (*Mondatok meghallgatása*). A 4. ábrán a *világ* keresőszóra adott kételemű találati lista látható. A megjelenítési opciókból az A neutrális (B) és egyfajta rossz (C) hangsúlyozású mondatok megjelenítése és a *Mondatok meghallgatása* jelölőnégyzeteket kapcsoltuk be. A lejátszás gombra kattintva hallgathatjuk meg a megfelelő mondatokat.



3. ábra. Grafikus forma a találati lista egy elemére

1. a világ csupán **erdélyi** kapcsolatban foglalkozik a témával
 - B: a világ csupán erdélyi kapcsolatban foglalkozik a témával
 - C: a világ **csupán** erdélyi kapcsolatban foglalkozik a **témával**
2. **bejárta** a világ **több jelentős** vadászmezejét
 - B: bejárta a világ több jelentős vadászmezejét
 - C: bejárta a **világ** több jelentős **vadászmezejét**

4. ábra. A „világ” szókeresésre adott találati lista mindhárom mondat típus és a meghallgatási opció beállítása után

A hangsúlyadatbázis nyilvános tesztelésének eredményeit [3] itt is közreadjuk (több száz tesztelő hallgatott meg 40-40 mondatpárt). Az eredményeket az A teszt eredményének megjelenítése (CMOS értékek) opció kiválasztásával jeleníthetjük meg. A teszt során az A-típusú (jó hangsúlyozásúnak tartott) mondatokat kellett összevetniük a tesztelőknek vagy a B (neutrális), vagy a C (rossz hangsúlyozású) ugyanazon mondattal egy-egy mondatpárt meghallgatva. Például az *olasz* keresőszóra adott találati lista egyetlen mondatot jelenít meg (*olasz klub csak elvéve igazolt akkoriban magyar labdarúgót*), amely a fenti opció beállítása után táblázatos formában tartalmazza a teszt eredményeit is erre a mondatra. A megjelenő CMOS értékeket [1] kiemeltük az 1. táblázatba.

1. táblázat: Az adatbázis egy mondatának átlagos CMOS értékei

CMOS	CMOS AB	CMOS AC
0,86 (N=7)	0,75 (N=4)	1 (N=3)

Hogyan kell értelmezni a CMOS adatokat? A vizsgált mondatra adott CMOS pontszámok 1, 0 vagy -1 értéket vehetnek fel, és a tesztalany döntésén alapulnak. Az A-típusú mondatra vonatkozó kedvező ítélet esetén 1 értéket kap a mondat. A (B) vagy a (C) mondatra vonatkozó döntés esetén pedig -1 értéket rögzítünk. Más esetekben (ha mindkettőt egyformának tartja, tehát nem dönt egyik mellett sem), akkor 0 értéket adunk. Az 1. táblázat több tesztelőre vonatkozó, átlagos CMOS értékeket tartalmaz. Az AB mondatok meghallgatása során az A-típusra vonatkozó preferenciát a CMOS AB pontszám tartalmazza (0,75). A zárójelben lévő 4-es érték a minta elemszámát jelenti, vagyis összesen 4 tesztalany találkozott (mindegyik egyszer) a fenti mondattal. Három tesztalany az A-típusú mondatot részesítette előnyben, egy pedig egyformának ítélte a B-típusú mondattal. Azaz $0,75 = (1+1+1+0)/4$. Az ugyanezen AC mondatok esetén mindhárom tesztalany az A-típusú mondatokat preferálta, mivel a CMOS AC érték 1. Az 1. táblázat CMOS oszlopában az összesített, az AB és AC mondatokra egyaránt vonatkozó ítéletek átlaga szerepel (0,86). A fenti mondattal összesen 7 tesztalany találkozott.

A találatok eddigi formájától jelentősen eltérő megjelenítést kapunk, ha az Összesítés hangsúlymintázatokra opciót választjuk. Hangsúlymintázatnak nevezzük a mondat szavaira vonatkozó hangsúlyjelek sorozatát balról jobbra értelmezve. Ha a mondatot a hangsúlymintázatával jellemezzük, akkor annyi jel van a hangsúlymintázatban, ahány szó van a mondatban (a névelők is szónak számítanak). Egy két szavas mondat hangsúly mintázata például a H- képpel fejezhető ki, de lehet -H is, vagy HH is. Ezen opciónál a megjelenítés egyéb beállításait és a következő pontban szereplő rendezési szempontokat is figyelmen kívül hagyjuk. Az 1-4. pontokban szereplő keresési feltételeknek megfelelő mondatokat a szavak száma és a hangsúlymintázatok szerint csoportosítjuk. A találati lista így ezen csoportosítás mellett a csoportok elemszámát és kérésre, a csoportba tartozó mondatokat tartalmazza (5. ábra).

Szavak száma	Hangsúlymintázat	Előfordulásszám
<input type="checkbox"/> 2	H-	1
		feltétlenül elgőzösítik
<input type="checkbox"/> 3	HH-	3
		beszélgetőpartnerem testőrökkel érkezik izzadságkészlete kimeríthetetlennek látszott nemzetközi versenyhelyzetben vagyunk
<input type="checkbox"/>	H--	3
<input type="checkbox"/>	-HH	1
<input type="checkbox"/>	H-H	1
<input type="checkbox"/> 4	H---	13
		egyetlenegy esetre tudok visszaemlékezni értelmetlenség közvetítő nyelvekre támaszkodni ezt tedd gondolataid kiindulópontjává időben beadta a jelentkezését jugoszlávia szétzúzásaiban is segédkeztek legjobb belátásuk szerint cselekedhetnek mindenkinek lesz valamilyen szerződése nagy a tömegtájékoztatás felelőssége nagyszerű híreink vannak számotokra nem váltak bensőjükben érzéketlenné oroszszággal összekötő vasútvonalat szeretnének öntisztulási folyamatot kell kezdeményezni semmit sem lehetett bizonyítani
<input type="checkbox"/>	HH--	8

5. ábra. Egy találati lista része, amikor összesítést kértünk a hangsúlymintázatokra

3.1.6. A találati lista rendezése

A találati lista 9 szempont szerint rendezhető. A rendezés iránya minden esetben lehet növekvő és csökkenő is. Rendezhetünk a mondatok jellemzői alapján: ábécé szerint, szószám vagy hangsúlyszám szerint. A nyilvános teszt eredményeit is felhasználhatjuk, így a CMOS értékek és a válaszok száma alapján is rendezhetünk. Ez utóbbi két szempont esetén választhatunk az összesített vagy külön az AB és AC mondatpárokra vonatkozó CMOS értékek és teszt válaszsámok között.

3.1.7. Keresés indítása

A **Keresés** gomb megnyomásával kezdeményezhetjük a találati lista megjelenítését, amely a keresési felület alatt jelenik meg. A találati elemeinek összetevőit a 3.1.5. pontban részleteztük.

4. Összefoglalás

Jelen cikkben az első magyar hangsúlyadatbázisra alapozott webalapú felhasználói felület felépítését és használatát mutattuk be. A hangsúlycímkékkel ellátott mondatok sok szempont alapján lekérdezhetők, a találatként megjelenő mondatokhoz pedig többféle ábrázolást választhatunk. A lekérdezése széles tárháza jó adatbányászati, elemzési alapot nyújt a kutatóknak. Az elkészült hangsúlyadatbázis alkalmas külön-

böző hangsúlyjelölő algoritmusok tesztelésére, a beszédszintézis továbbfejlesztésére. Segítheti a gépi beszédfelismerést, használható az oktatásban és általában is új lendületet adhat a hangsúlykutatásokhoz, a hangsúly és a mondat szerkezet viszonyának vizsgálatához.

Támogatás: Az adatbázis létrehozását a Paelfe (Grant No. AAI-08-01-2011-0001) és az EITKIC_12-1-2012-001 projektek támogatták

Hivatkozások

1. ITU-T: P.800 Methods for subjective determination of transmission quality (1996)
2. Olasz G.: Precíziós, párhuzamos, magyar beszédatadabázis fejlesztése és szolgáltatásai. In: Gósy M. (szerk.): Beszédkutatás. MTA Nyelvtudományi Intézet, Budapest (2013) 261–270
3. Olasz G., Abari K., Bartalis M. In: Gósy M. (szerk.): Beszédkutatás 2014. MTA Nyelvtudományi Intézet, Budapest (2014) Megjelenés alatt.
4. Olasz G., Németh, G., Kiss, G.: Hungarian audiovisual prosody composer and TTS development tool. In: Puppel S., Demenko, G. (szerk.): Prosody 2000. Poznan, Poland (2001) 167-178
5. Vicsi K., Víg A.: Az első magyar nyelvű beszédatadabázis. In: Gósy, M. (szerk.): Beszédkutatás 1998 MTA Nyelvtudományi Intézet, Budapest (1998) 163–177