

A kérdés-válasz fordulók dallamszerkezetének fonetikai vizsgálata magyar nyelvű társalgásokban – rádiós anyag felhasználásával –

Beke András

1. Bevezetés

A társalgás az egyik legalapvetőbb és legtöbbet használt beszédmód, ezért fontos, hogy megvizsgáljuk szerkezeti elemeit (a mondat szintjén: kérdő, felszólító, felkiáltó, óhajtó, kijelentő mondatokat) és ezek akusztikai szerkezetének építőelemeit: a hangmagasságot (F0) és a hangintenzitást (I). A társalgások vizsgálata magyar nyelven a legutóbbi időkben indult el a spontán társalgások vizsgálatának egyik szegmenseként (Bolla 1992; Couper-Kuhlen–Selting 1996; Markó 2007). Jelen tanulmányban csupán a kérdés-válasz dallamszerkezetét igyekeztünk leírni, amit a jövőben további altémákra szeretnénk bontani a kérdés fajtái szerint: kiegészítendő kérdés, eldöntendő kérdés, választó kérdés, befejezetlen kérdés, morfémával jelzett kérdés.

A kérdés dallamszerkezetét többen is vizsgálták már különféle szempontból (Csúri 1919; Hegedűs 1930; Deme 1962; Fónagy–Magdics 1963; Varga 1994; Gósy 1993; Olaszy–Koutny 2001; Olaszy 2001). Az egyik legrészletesebb leírást Fónagy–Magdics (1967) szerzőpáros végezte el a kérdő mondatok fajtáinak dallamszerkezetének hallás alapján történő lekottázásával. Olaszy (2002) a leggyakoribb kérdésformák dallammene-
tének és intenzitás szerkezetének fonetikai leírását végezte el. A kijelentő mondatok leírása szintén megtörtént (Deme 1962; Fónagy–Magdics 1967; Olaszy 1995). A mondatfajták prozódiai vizsgálata napjainkban előtérbe került, mivel a gépi beszéddel működtetett ún. dialógus-rendszerekben a párbeszédes forma tartalmazhatja ezeket a kifejezési formákat, így ezek generálásához szabályokat kell kialakítani. A verbális kommunikáció komplex és folyamatos hangzásformája ugyanis megköveteli, hogy az egymás után ugyanazon beszélőtől elhangzó közlések egymáshoz hang-

zásban kapcsolódjanak, prozódiai struktúrájuk egymásból következzen, vagyis a beszéd hangzása folyamatos és kifejező legyen, és érzékeltesse a szituáció nyelvi tartalmát.

Hipotézisünk szerint a magyar nyelvű dialógusokban a kérdés dallamszerkezete összefügg a válasz dallamszerkezetével. Vizsgálatunkban arra keressük a választ, hogy milyen törvényszerűségekkel lehet leírni ezt, s ezen törvényszerűségekkel hogyan lehet a kérdés dallamszerkezetéből előre következtetni a válasz dallamszerkezetére. Tehát konkrét adatokkal jellemezzük a magyar nyelvű dialógusokban lévő kérdés-válasz fordulók dallamszerkezetét. Olyan adatrendszert szeretünk volna felállítani, amelynek felhasználásával meg lehet határozni a kérdés – válasz dallamátmenetet, vagyis előre meg lehet jósolni a válasz dallamszerkezetét a kérdés dallamszerkezetének függvényében.

2. Anyag és módszer

A nyelvi anyagot rádiós műsorokból állítottuk össze (Danubius rádió). Ebből 62 mondatnyi szöveget vágtunk ki. A szöveg párbeszédből épül fel, amelyet két személy valósít meg. Az egyik bemondó, aki kérdez, a második betelefonáló. A vizsgált szöveg betelefonálás műsorból származik, így a szöveg spontán jellegét biztosítottuk. A két műsorvezető közül az egyik középkorú férfi: (MF= összesen 18 mondat), a másik középkorú nő (MN= összesen 13 mondat). A betelefonálók között vannak nők (TN: összesen 11 mondat) és férfiak (TF: összesen 20 mondat) is.

A hangfelvételek meghallgatásából elkészítettük az anyagok szövegének írott változatát is. A prozódiai méréseket és jellemzéseket ennek a szövegnek a mondataira végeztük el.

A vizsgált szöveget megnyilatkozásokként felváltuk. Természetesen nem tekintettük szónak a hezitálásokat (akár nazális mormogás, akár a „svá” különböző időtartamú ejtésével valósultak is meg); továbbá a befejezetlen hangsorokat. Ezeket mind a számolásból, mind az elemzésből kihagytuk; nem hagytuk ki ugyanakkor az ismétléseket, akár kötőszavak,

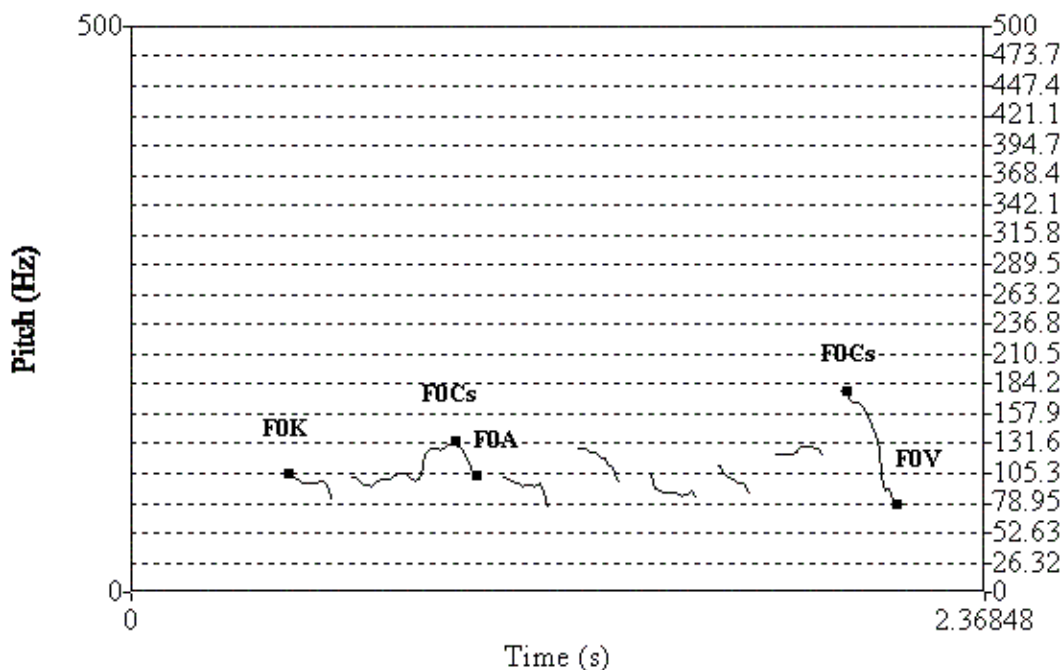
akár egyéb szófajú szavak voltak. Szünetnek tekintettünk minden olyan jelkimaradást, amely legalább 100 ms hosszúságú volt (Laczkó 1993, Horváth 2004).

A hangmintákat szótagokra bontottuk, és ezeken mértük meg a szótagszintű FoCs-értékeket (a szótagban mérhető legmagasabb frekvencia-érték) Praat 4.4 szoftverrel. A statisztikai számításokban az SPSS statisztikai programot használtuk, míg a regressziók megállapításában az Excel programot.

A mért paraméterek a következők:

- a beszélőre jellemző szűkített F_0 -átlag értéke Hz-ben;
- a mondatkezdés alaphangfrekvenciája (**FoK**), amely az első hangsúlyos szón mérhető frekvenciacsúcs Hz-értékét jelenti;
- a mondat befejezésekor mérhető érték (**FoV**) az utolsó szótagban;
- a hangsúlyokban mérhető alaphangfrekvencia-csúcs értéke (**FoCs**);
- a hangsúly csúcsa után az alaphangfrekvencia visszatérési értéke (**FoA**)

1. ábra: Az „Általában a férfiak felkészületlenek?” eldöntendő kérdő mondat dallamgörbéje és a dallamgörbén mért paraméterek



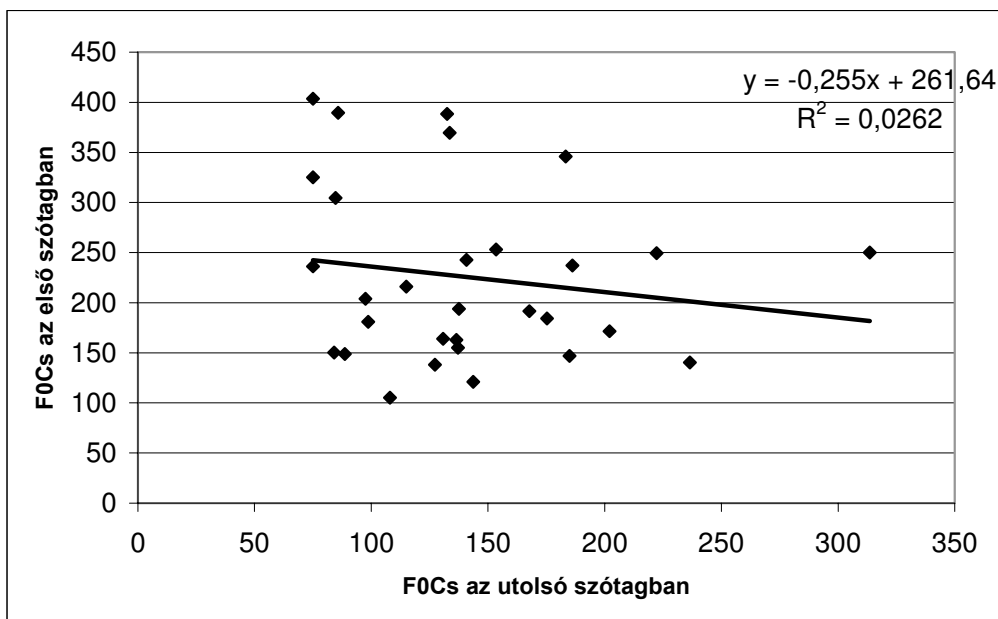
A vizsgálatban négy különböző mérést végeztünk el. Az elsőben megvizsgáltuk, hogy létezik-e összefüggés a kérdés utolsó szótagjában detektált F_0Cs -érték és a válasz első szótagjában mért F_0Cs -érték között. E vizsgálat kétfelé tagolódott. Az elsőben úgy dolgoztunk, hogy a F_0Cs -értékeket %-os formába alakítottuk át úgy, hogy a beszélő átlagos frekvenciáját vettük alapul, vagyis azt vettük 100%-nak. Az első vizsgálaton belül a másodikban a mért adatok relatív számadataival dolgoztunk. A második mérésben egy olyan módszert alkalmaztunk, amelyben a kérdés utolsó szótagjában mért F_0Cs -értéket vettük 100%-nak, és ahhoz képest vizsgáltuk meg a válasz első szótagjában mért F_0Cs -értéket (ekkor százalékos adatokat adtunk meg). A harmadik mérésben megvizsgáltuk, hogy van-e kapcsolat a kérdés utolsó szótagjában mért F_0V -érték (a szótagban az utolsó Hz-érték) és a válasz első szótagjában mért F_0K -érték között. A negyedikben olyan módszert alkalmaztunk, amelyben a kérdés utolsó négy szótagjában mért F_0Cs -értékeknek és a válasz első három szótagjában mért F_0Cs -értékeknek vettük az átlagát, és ezek között néztük meg az összefüggést.

3. A vizsgálat eredményei

3.1. Az 1. vizsgálat eredményei

A kérdés utolsó szótagjában mért F_0Cs -értéket és a válasz első szótagjában mért F_0Cs -értéket vizsgáltuk. (Az adatokat l. a 4. táblázatban.) A 2. ábrán láthatjuk az eszerint felrajzolt adatokat lineáris regressziókkal és a R^2 értékekkel. Ebben a módszerben az $R^2=0,0262$, azaz az összefüggés a két adatsor között nem erős. Tehát a kérdés F_0Cs -értéke nincs összefüggésben a válasz F_0Cs -értékével.

2. ábra: A kérdés utolsó szótagjában mért FoCs-értékek és a válasz első szótagjában mért FoCs-értékek közötti korreláció



Ebben a mérésben a FoCs-értékeket a beszélő átlagfrekvenciájához mértük, így százalékos adatokat kaptunk (az adatokat l. a 2. táblázatban). Így megfigyelhettük, hogy a beszélő az átlagos frekvenciaértékéhez képest hány százalékkal emeli meg vagy süllyeszti le a beszéddallamát. Azonban itt is megállapíthattuk, hogy a két adatsor között nincs magas fokú korreláció (l. 6. ábra). Tehát a kérdés alapfrekvenciájához mért utolsó FoCs-érték nincs korrelációban a válasz első szótagjában mért FoCs %-os értékével. Azonban itt megállapíthatjuk, hogy az esés-emelkedés átlagos értéke +50% és -50% közé esik.

1. táblázat: Az alapfrekvenciához mért F0 mozgás %-os arányban

Személyek	Az alapfrekvencia (F0): 100%	FoCs 1szótag	2	3	4	Emelkedés/süllyedés(%)
Műsorvezető (férfi)	106,32	127,87	87,27	75,27	137,17	
	100,00%	120,27%	82,08%	70,80%	129,02%	29,02%
Betelefonáló (férfi)	121,19	155,03	137,3	162,94		
	100,00%	127,92%	113,29%	134,45%		34,45%

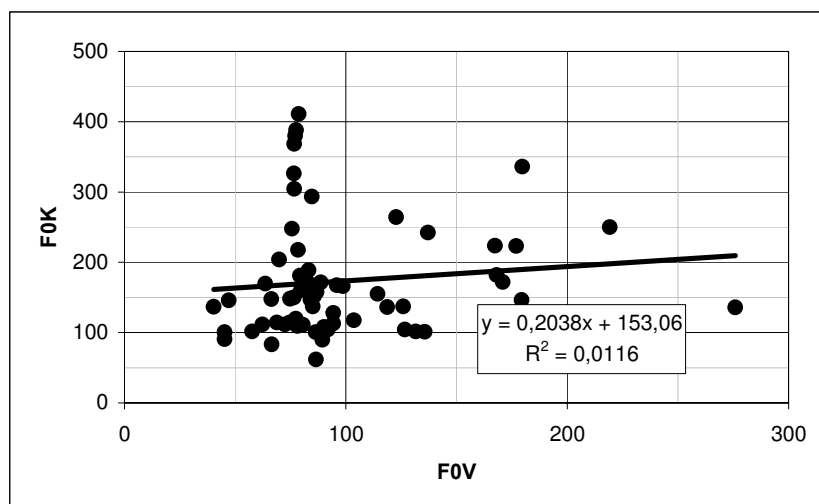
3.2. A 2. vizsgálat eredményei

Ebben a vizsgálatban a kérdés utolsó szótagján mért F_0Cs -értéket vettük 100%-nak, és ehhez viszonyítottuk a válasz első szótagján mért F_0Cs -értéket (l. 8. ábra). Ekkor a $R^2=0,9494$. De a mérések során kiderült, hogy ez a módszer nem megfelelő, hiszen az utolsó szótag F_0Cs -értékét 1-nek vettük, ami a számításban nagy torzításokat okozott. Majd amikor ezt korrigáltuk, egyértelműen kiderült, hogy a két adatsor között a korreláció alacsony fokú, hiszen $R^2=0,0262$. Tehát a két adatsor között nincs erősebb összefüggés.

3.3. A 3. vizsgálat eredményei

Ebben megvizsgáltuk, hogy van-e összefüggés a kérdés utolsó szótagjában mért F_0V -érték és a válasz első szótagjában mért első F_0K -érték között (az adatokat l. a 3. táblázatban). A függvény jól mutatja, hogy ez az összefüggés minimális, hiszen a $R^2=0,0116$, tehát nem függ az F_0V -értékétől a F_0K -értéke (l. 3. ábra).

3. ábra: A kérdés utolsó szótagjában mért F_0V -érték és a válasz első szótagjában mért első F_0K -érték közötti korreláció

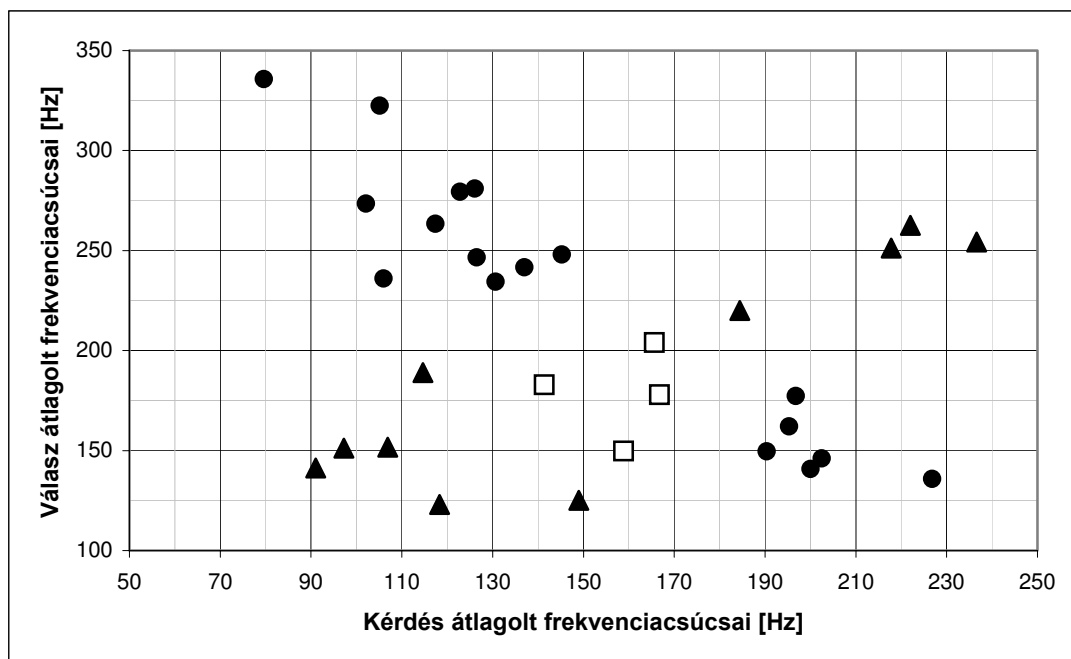


3.4. A 4. vizsgálat eredményei

A negyedik mérési eljárásban a kérdés utolsó négy szótagjában mért F_0 Cs-értékeket és a válasz első három szótagjában mért F_0 Cs-értékeket átlagoltuk, és az átlagok között vizsgáltuk meg az összefüggést. Ez a vizsgálati módszer sikeresnek bizonyult, hiszen jól korreláló adatsorokat tudtunk létrehozni.

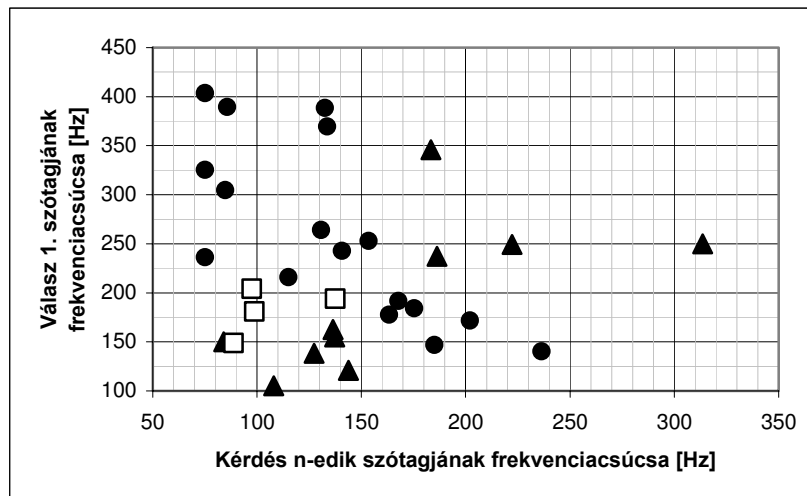
A 4. ábra mutatja a válaszok első három szótagjában fellépő frekvenciacsúcsok átlagát a kérdés utolsó négy szótagjából hasonló módon képzett átlagának függvényében.

4. ábra: A válaszok első három szótagjában fellépő frekvenciacsúcsok átlagát a kérdés utolsó négy szótagjából hasonló módon képzett átlagának függvényében



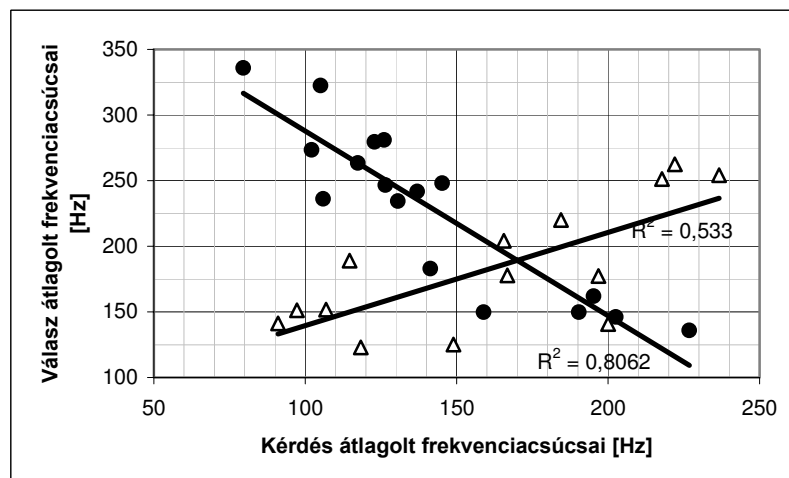
Az 5. ábrán az átlagolás nélkül ábrázolt adatokat olvashatjuk. Jól látható, hogy az átlagolás a keresett összefüggés statisztikai jellemzőit javította.

5. ábra: A válaszok első három szótagjában fellépő frekvenciacsúcsok átlagát a kérdés utolsó négy szótagjából hasonló módon képzett átlagának függvényében



Három, jól elkülöníthető sokaság képezhető az adatokból, ezeket teli körrel, teli háromszöggel, valamint üres négyzettel jelöltük. Több lehetőség is felmerült ezen jelenség magyarázatára (pl.: eldöntendő/kiegészítendő, szótagszám a mondatban stb.). Végül a párbeszéd szereplői-nemében véltük megtalálni a csoportosító faktort. Tehát a faktor a nemek egyezősége, avagy különbözősége. A 6. ábrán láthatjuk az eszerint csoportosított adatokat, a megfelelő lineáris regressziókkal és a (különösen a különböző nemű párokra vonatkozó meggyőzően magas) R^2 értékekkel.

6. ábra: A válaszok első három szótagjában fellépő frekvenciacsúcsok átlagát a kérdés utolsó négy szótagjából hasonló módon képzett átlagának függvényében



Ahol a nemek különböznek, ott az összefüggés fordított, s így írható le matematikailag:

$$y = -1,4085x + 428,75$$

Tehát ezzel az egyenlettel lehet előállítani számítógépesen a női-férfi (különböző neműek) párbeszédben a válasz első három szótagjának F_0Cs -értékének átlagát a kérdés utolsó négy szótagjának F_0Cs -értékének átlagából. Az $R^2=0,8062$.

Ha a nemek egyeznek, akkor az összefüggés egyenes, és így írható le egyenletben:

$$y = 0,7114x + 68,32$$

Tehát, ezzel az egyenlettel lehet előállítani számítógépesen a férfi-férfi, nő-nő (egyező neműek) párbeszédében a válasz első három szótagjának F_0Cs -értékének átlagát a kérdés utolsó négy szótagjának F_0Cs -értékének átlagából. Az $R^2=0,533$.

Ebben az esetben a lineáris regresszió együtthatóinak konfidenciaintervallumában szerepel az $m=1$ és a $b=0$ ($y=m*x+b$) (l. 8. ábra). Tehát ezen adatok alapján nem zárható ki, hogy a válasz első három szótagjának F_0Cs -értékének átlaga megegyezhet a kérdés utolsó négy F_0Cs -értékének átlagával. Bár itt figyelembe kell venni azt, hogy a nők alapfrekvenciája magasabb, mit a férfiaké.

2. táblázat

$y=m*x+b$		Azonos neműek	Különböző neműek
m	Érték	-1,408	0,711
	95%-os	-1,789	1,13
	konfidenciaintervallum	+ -1,028	- 0,293
b	Érték	428,747	68,32
	95%-os	484,719	139,04
	konfidenciaintervallum	+ 372,776	- 2,399

4. Következtetések

A fenti kísérletsorozattal megállapítottuk, hogy a kérdés-válasz harmónia átmenetei milyen paraméterektől függenek. Megállapítottuk, hogy nagy valószínűséggel nem befolyásolja a kérdés utolsó szótagja a válasz első szótagját. Megállapítottuk, hogy a kérdés utolsó szótagján mért F_0V nincs döntő befolyással a válasz első szótagján mért F_0K -ra. Megállapítottuk azt is, hogy a kérdés-válaszban a dallammenetek a mondat végén átlagosan -50% és $+50\%$ közé esnek, tehát az alaphangfrekvenciához képest 50% -val süllyednek le vagy 50% -kal emelkednek fel.

A kísérletek során megállapítottuk, hogy a kérdés utolsó négy szótagján mért F_0Cs -értékek átlaga és a válasz első három szótagján mért F_0Cs -értékek átlag között erős korreláció van. Itt bebizonyítottuk, hogy az összefüggésben **döntő tényező** az, hogy **azonos vagy különböző neműek** társalognak. Az azonos és a különböző neműek párbeszédében megállapítottuk, hogy milyen matematikai összefüggések vannak, és megalkottunk olyan egyenleteket, amelyek által létre lehet hozni a kérdés utolsó négy szótagján mért F_0Cs -értékek átlagából a válasz első három szótagjában mért F_0Cs -értékeket.

További vizsgálatokat tervezünk arra vonatkozóan, hogy a beszéddallam milyen határok között változtatható meg anélkül, hogy az elhangzó szituáció dialógus természetes hangzása, a mondatok egymásból való hangzásbeli következése ne gépi beszédként hasson.

További vizsgálatokat tervezünk arra vonatkozóan is, hogy a fenti szabályrendszert szintetizált beszédben alkalmazzuk, és percepció tesztekkel állapítsuk meg a szabályok működésének eredményét és korlátait.

Irodalom

- Bolla Kálmán 1992. *Szupraszegmentális elemzések*. Egyetemi Fonetikai Füzetek 7. Budapest.
- Couper-Kuhlen – Elizabeth-Selting, Margret (ed.) 1996. *Prosody in Conversation. Interactional Studies*. Cambridge: Cambridge University Press.
- Csúri Bálint 1919. Hanglejtés. *Magyar Nyelvőr* 78. 71–78.
- Deme László 1962. A hanglejtés. In: Tompa József (szerk.): *A mai magyar nyelv rendszere* II. Budapest: Akadémiai Kiadó. 503–522.
- Fónagy Iván – Magdics Klára 1967. *A magyar beszéd dallama*. Budapest: Akadémiai Kiadó..
- Fónagy Iván – Magdics Klára 1963. A kérdő mondatok dallamához. *NyudÉrt.* 40. 89–106.
- Gósy Mária 1993. A kiegészítendő kérdés dallamváltozása. *Magyar Nyelvőr* 117. 443–457.
- Hegedűs Lajos 1930. *Magyar hanglejtésminták grafikus ábrázolása. Kísérletfonetika tanulmány*. Bécs: Collegium Hungaricum füzetek V.
- Horváth Viktória 2004. *Megakadásjelenségek a párbeszédekben*. In: *Beszédkutatás* 2004. Budapest: MTA Nyelvtudományi Intézet. Budapest.
- Laczkó Mária 1993. A tempó és a szünet viszonya a hangos olvasásban. *Beszédkutatás* 1993. Budapest: MTA Nyelvtudományi Intézet.
- Markó Alexandra 2007. A társalgás hangtana. *Magyar Tudomány*..
- Olaszy Gábor 1995. A kérés, a figyelmeztetés, a felszólítás és a kérdés prozódiaja a kijelentő mondat tükrében. *Beszédkutatás* 1995 Budapest: MTA Nyelvtudományi Intézet.
- Olaszy Gábor – Koutny I. 2001. Intonation of Hungarian questions and their prediction from text. In: Puppel Stalislav – Grazina Demenk (eds.): *Prosody 2000*. Poznan. 179–196.
- Olaszy Gábor 2005. A magyar kérdés dallamformáinak és intenzitás-kezelésének fonetikai vizsgálata. *Beszédkutatás* 2002. Budapest: MTA Nyelvtudományi Intézet.
- Varga László 1993. A magyar beszéddallamok fonológiai, szemantikai és szintaktikai vonatkozásai. *Nyelvtudományi Értekezések* 135.
- Varga László 1994. A hanglejtés. In: Kiefer Ferenc (szerk.) *Strukturális magyar nyelvtan 2. Fonológia*. Budapest: Akadémiai Kiadó. 468–549.

Summary

Phonetic analysis of vocalization structure of question-answer pairs in the Hungarian language (using radio-aired material)

Because conversation is the most frequent situation of talking, it is important to analyze its structural components (interrogative, declarative, etc. sentences), and their acoustic parameters, such as pitch (F_0) and intensity (I). According to our hypothesis there is a correspondance between the vocalization structure of the question and that of the answer's. In the present article these correspondances are investigated and a prediction is made regarding the vocalization structure of the answer, based on that of the question's.

In the investigations the maximal frequency value (F_0^{Cs} , KHz) for each syllable was measured. It was found that there is a strong correlation between the average of the maximal frequency (F_0^{Cs}) value of the last four syllable of the question, and that of the first three syllable of the answer. It was also shown that the sex of the conversing parties is a crucial factor determining the parameters of the above mentioned relationship: the corresponding mathematical formulae were established. Using these formulae the average of F_0^{Cs} in the first three syllable of the answer can be predicted from the average of F_0^{Cs} in the last four syllable of the question.

Further investigations are intended to be carried out regarding the applicability of the above relations in synthesized speech, and the verification and bounds of these relations using perception tests.