

Beszéddallam reprodukciója kottakép alapján

Gocsál Ákos^{1,2}, Partos Dorka¹

¹*Pécsi Tudományegyetem Művészeti Kar*

²*ELTE Nyelvtudományi Kutatóközpont*

Abstract

In 1967, Iván Fónagy and Klára Magdics published figures of hundreds of intonation contours, using musical notation, reflecting a wide range of emotional states and attitudes. The purpose of this study was to examine whether they can be reproduced by means of speech manipulation software. The intonation curves of ten sentences, selected from the book, were reproduced in Praat, after musical notes were converted to fundamental frequency values. A pilot study with five sentences (reflecting apathy, joy, sadness, fright, and mock) was conducted to test whether listeners can decode the emotions and attitudes expressed. Thirty-five students (aged 18-27 years), all native speakers of standard Hungarian, indicated on 5-point scales how intensely they felt the sentences they had heard expressed the emotions (1 = not at all, 5 = very strongly). Each sentence was evaluated by all five emotions, and an extra emotion not included in the test sentences (surprise). In general, the random distribution of scores suggested that listeners were unable to identify the intended emotions, but the sentences were perceived as not conveying other qualities. It is concluded that prosodic features other than intonation are also necessary for recognizing emotions.

1. Bevezetés

Jelen műhelytanulmányunkban egy olyan módszert ismertetünk, amelynek alkalmazásával kottaképeken megjelenített beszéddallamot a Praat program segítségével természetes bemondásra illesztünk.

A beszéddallam képi megjelenítése iránti igény nem új a beszédkutatás szakirodalmában. Sir Joshua Steele már 1775-ben kidolgozott egy, a zenei notációhoz hasonló jelrendszert a dallam és a prozódia más elemeinek lejegyzéséhez (idézi Murdoch, 1883), ám ez nem terjedt el. Viëtor (1897:103) basszuskulcsban kottaként ábrázolta a német *du* szó különféle érzelmekkel ejtett változatait, Scripture (1906) pedig már az alaphang kontúráját ábrázoló „beszédgörbéket”

Email addresses: gocsal.akos@pte.hu (Gocsál Ákos), pdorka2002@gmail.com (Partos Dorka)

közölt, amelyek megrajzolásához oszcillogramokon (kimogramokon) mért hullámhosszok alapján számolt frekvenciaadatokat. Jones (1909) a kétféle ábrázolásmódot egyesítette úgy, hogy a lehallgatott beszéd dallamgörbéit az ötvonalas rendszerben, basszuskulcsban vagy violinkulcsban ábrázolta. Ő azonban nem kimogramokat használt, a lejegyzés során a beszéddallam frekvenciáit hangvilla hangjához viszonyítva határozta meg.

A beszéd kutatás hazai szakirodalmában már Regner (1861), Gombocz (1907) és Tolnai (1915) is említi a beszéddallam jelentőségét – Gombocz francia nyelvű mondat kottaképét is közli –, a magyar nyelvre vonatkozóan azonban csak Csűry (1925) munkájában találunk először képi ábrázolást. Csűry ötvonalas, kottaszerű, de nem zenei lejegyzést alkalmazott. Néhány évvel később pedig Hegedűs (1930) Scripture-höz hasonlóan dallamgörbéket rajzolt, azonban ezek elemzésekor egy-egy kiemelt frekvenciaértékhez zenei hangokat is közölt.

A dallamok lejegyzésében jelentős előrelépés történt az 1960-as években. Magdics (1963) részletes elemzést nyújtott a beszéd és a zene összehasonlításáról, és altkulcsban több mondat kottaképét is közölte. Arra a következtetésre jutott, hogy bár a dallam, a ritmus, a hangközök stb. sok szempontból különböznek, mégis e tényezők alapján összehasonlítások végezhetők. Magdicsnak e munkája nyújt elméleti alapot Fónagy és Magdics (1967) nagyszabású munkájához, amely sok száz kottapéldán keresztül mutatja be magyar beszéd dallamformáit. Ezeknek a dallamformáknak a lejegyzése Magdics (1963) példáihoz hasonlóan szintén altkulcsban történt. Munkájukat élesen bírálta Boros (1969) a zenei kontroll hiánya és számos, a lejegyzés során adódott pontatlanság miatt. Bartók (1969) ugyan hitelesnek tartotta a közölt dallamíveket, azonban felfeltette, hogy ezek a dallamok nem feltétlenül tükrözik jól a köznyelvi beszédet, mivel a mondatokat színészek bemondásában rögzítették.

Boros a későbbiekben zenei kották formájában közölt beszéddallamokat (Boros, 1971, 1975), míg Bartók (1974) nyolcvonalas rendszert használt, hozzátéve, hogy ez a vonalrendszer semmi esetre sem tekinthető „zeneinek”. Bartók radikálisan eltérő véleményt fogalmazott meg Fónagy és Magdics munkájával szemben, ugyanis kijelentette, hogy bár a beszédben szubjektív módon meghatározott

hangközök hasonlíthatnak a zenei hangközökre, azokkal sohasem lehetnek azonosak, mert a beszédben zenei hangközök nem léteznek.

A mai beszédelemző szoftverek alkalmazásával könnyen megjeleníthetők a dallamgörbék, így jól támogatható az intonációs rendszerek fonológiai, azaz diszkrét kategóriákkal, szimbólumokkal való leírása (Hirst & Cristo, 1998:16), de a görbék megrajzolásával könnyen tanulmányozhatók a prototipikustól (Markó, 2017), eltérő mintázatok is (Markó, 2013; Tóth, 2016). Napjainkra a beszéddallam zenei megalapozottságú, kottaszerű ábrázolása már visszaszorult a beszéd-kutatás szakirodalmában, azonban a dallamok zenei hangközökkel való párhuzamba állítása időről időre felmerül (Bowling et al., 2010; Robledo et al., 2016; Gocsál & Urbanics, 2022), sőt, ismeretesek olyan szerzők, akik határozottan állást foglaltak a beszéd kottával történő leírása mellett (Chow & Brown, 2018). Ők azzal érvelnek, hogy a kotta az egyetlen olyan eszköz, amely a dallamot és a ritmust egységes rendszerben tudja ábrázolni, illetve csak a kotta segítségével tudja egy beszélő egy másik beszélő prozódiaját reprodukálni.

Az itt ismertetendő módszerünknek is ez a kiindulópontja. Célunk az, hogy a Fónagy és Magdics (1967) által közölt kották alapján szoftveres manipuláció alkalmazásával reprodukáljuk a lejegyzett dallamot, majd validálni szeretnénk azokat, azaz azt szeretnénk megállapítani, hogy a hallgatók reprodukált dallam észlelésekor képesek-e visszakódolni az adott dallam által Fónagy és Magdics (1967) szerint kifejezett érzelmet, attitűdöt. Ugyanakkor nem kívánunk elvi állásfoglalást tenni a zenei lejegyzés helyességéről vagy helytelenségéről, hanem az motivál bennünket, hogy megvizsgáljuk, a Fónagy és Magdics által közölt, kottakép formájában rendelkezésre álló több száz dallamforma gyakorlati célokra felhasználható-e.

2. Az alkalmazott eljárás

2.1. A mondatok kiválasztása

Az eljárás kidolgozásához kiválasztottunk Fónagy és Magdics (1967) kötetéből tíz olyan kottát, amelyek a szöveges ismertetés alapján jelentősen eltérő

érzelmet, attitűdöt, szándékot fejeznek ki. Ezek az alábbiak: 1. felháborodás előkészítő résszel (583. kotta, p. 239.), 2. közöny (364. kotta, p. 192.), 3. öröm (367. kotta, p. 193.), 4. szomorúság (372. kotta, p. 194.), 5. rémület (432. kotta, p. 205.), 6. nyugtatás (445. kotta, p. 208.), 7. gúny (502. kotta, p. 221.), 8. fojtott hangú fenyegetés (550. kotta, p. 232), 9. sértődöttség, duzzogás (612. kotta, p. 245.), 10. elgondolkodás, töprengés (624. kotta, p. 248.). A mondatok kiválasztása önkényesen történt, azonban törekedtünk arra, hogy minél változatosabb érzelmeket és attitűdöket tükrözzenek, és szerepeljenek közöttük alapérzelmek is (öröm, szomorúság).

Az alábbiakban a fenti 1. példán keresztül ismertetjük az eljárást. Első lépésként a MuseScore v3.6.2.5 programban rögzítettük a dallamot (MuseScore, 2021), így az lehallgathatóvá vált (1. ábra).



1. ábra. A *Nahát, ez már mégiscsak sok!* mondat kottaképe (Fónagy & Magdics, 1967:239 alapján).

A következő lépésben táblázatos formában rögzítettük a kottáról leolvasható zenei hangokat, illetve a hozzájuk tartozó frekvenciaértékeket egészen kerekítve (1. táblázat).

1. táblázat. A *Nahát, ez már mégiscsak sok!* mondat zenei hangjai.

szótag	<i>Na</i>	<i>hát</i>	<i>ez</i>	<i>már</i>	<i>még</i>	<i>is</i>	<i>csak</i>	<i>sok</i>
zenei hang	G	G	G	G	E'	D'	Db'	C'
frekvencia (Hz)	196	196	196	196	330	294	277	262

A frekvenciaértékek tanulmányozása során – összevetve azokat Grácsi et al. (2020) által közölt, mai alapfrekvencia-adatokkal – azt valószínűsítettük, hogy a bemondás női beszélőtől származik. A Fónagy és Magdics (1967:314) által adott rövid módszertani leírásból azonban kiderül, hogy az egyes dallamformák

lehallgatását, „kiéneklését”, majd első kottázását követően a lejegyző mindegyik dallamot közös „alaphangra” transzponálta, emiatt a frekvenciaadatok nem adnak támpontot a beszélő nemének legalább bizonyos valószínűséggel történő megítéléséhez. Ez az említett „alaphang” a szerzők közlése szerint a beszélő által a legkisebb erőfeszítéssel képezhető hang volt, amelyet a végső lejegyzéskor minden beszélő esetében az altkulcson a legalsó vonalra (F hang) helyeztek. Ez a transzponálási mozzanat kétségtől összevethetővé teszi a különböző beszélők által produkált dallamokat, azonban – egyetértve Bartók (1969) korabeli kritikai észrevételével, és kiegészítve azt – az „alaphang” itt csak egy technikai fogalom, és semmi esetre sem tekinthető annak az átlagos alapfrekvenciának, amelyet egyébként ilyen esetekben mérni szokás. Ezt alátámasztja az, hogy a kották túlnyomó többségében a jelzett F hang alatt már nem jelentek meg hangok, az átlagos alapfrekvencia viszont éppen attól „átlagos”, mert egy szélesebb terjedelem középértéke, tehát törvényszerű, hogy alatta is előforduljanak hangok. Bartók (1969) úgy vélte, hogy a szerzők a vizsgált beszédszakaszok legmélyebb hangját tekintették „alaphangnak”, azonban ez sem pontos megállapítás. A kötet számtalan olyan kottát közöl, amelyekben egyetlen F hang sincs, ugyanakkor előfordul F alatti hang is (pl. 714/b. kotta, p. 294.).

Mindez tehát kérdéseket vet fel azzal kapcsolatban, hogy a kottákon megjelenített beszéddallamok reprodukciójakor milyen alapfrekvenciát állítson be a kutató, illetve a beszélő létező, „igazi” alapfrekvenciájához a kottán ábrázolt hangokat hogyan viszonyítsa. Jelen munkánkban erre két megoldást mutatunk be.

2.2. A mondatok rögzítése

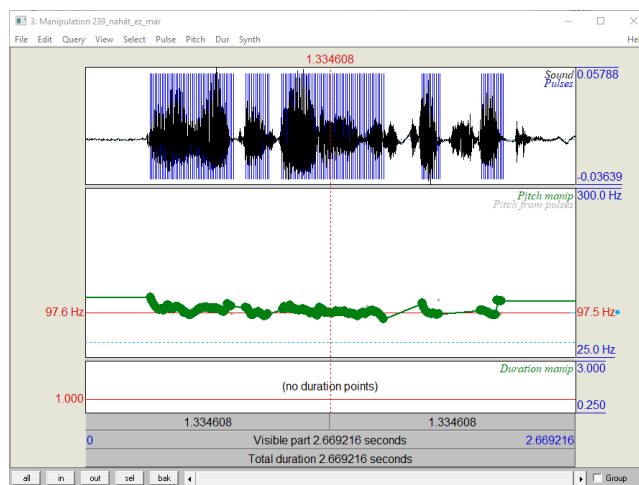
A kiválasztott mondatokat egy magyar anyanyelvű, 53 éves férfi beszélő bemondásában rögzítettük. A beszélő a sztenderdnek tekinthető köznyelvi nyelvváltozatot beszéli, nem dohányzik, beszédhibája, hangképzési rendellenessége, beszédszervi betegsége nincs.

A bemondások rögzítése TASCAM DR-44WL típusú digitális diktafonnal, .wav formátumban (16 bites felbontásban, 44 kHz-es mintavételezéssel), csendes

szobában, zavaró körülmények nélkül történt. A beszélő a kiválasztott mondatokat semleges érzelmi állapotban, dallamvariáció nélkül, egyenletes alaphangfrekvenciával mondta el.

2.3. A dallamgörbe vizsgálata

A fenti eljárással rögzített *Nahát, ez már mégiscsak sok!* (239. kotta) mondatból a Praat 6.0.52. programban (Boersma & Weenik, 2019) létrehoztunk egy manipulációs objektumot. Minden beállítást a Praat által felkínált alapértéken tartottunk (2. ábra).

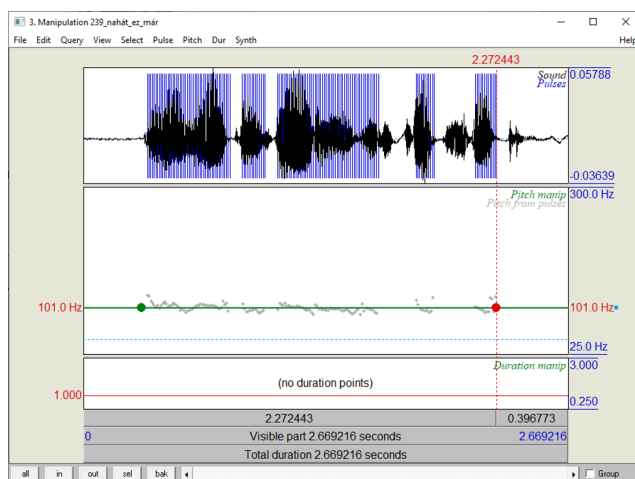


2. ábra. A *Nahát, ez már mégiscsak sok!* mondat a manipulációs ablakban.

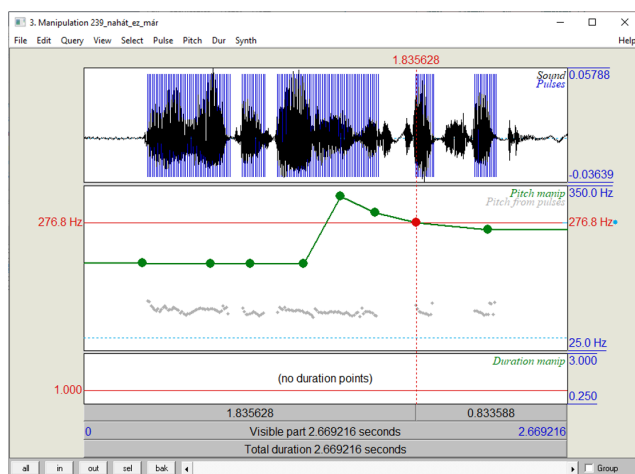
Látható, hogy a beszélő, bár igyekezett dallamvariáció nélkül beszélni, ez nem sikerült teljesen, kisebb ingadozások továbbra is vannak a dallamban. Első lépésként ezeket az ingadozásokat küszöböltük ki, megfelelő beállításokkal monotonná alakítottuk az összes dallamot, majd az alaphangfrekvenciát a vizsgált mondat alaphangfrekvenciájának mediánjára állítottuk (100 Hz, 3. ábra).

Ezt követően először az 1. táblázat szerinti értékekre állítottunk be az egyes szótagokon az alaphangfrekvenciát (4. ábra).

Az így kapott dallamgörbe azonban túl magas, természetellenes hangzású volt, így a férfi beszélő számára természetes hangfekvésre transzponáltuk. A



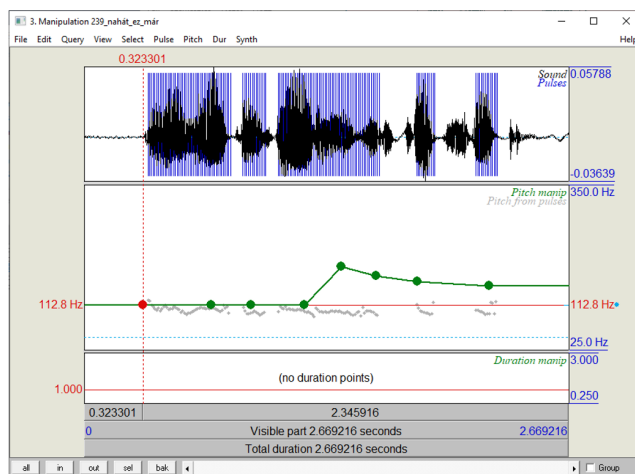
3. ábra. A *Nahát, ez már mégiscsak sok!* mondat egyenletes alaphangon.



4. ábra. A *Nahát, ez már mégiscsak sok!* mondat dallamgörbéje az 1. táblázatban közölt frekvenciaértékek alkalmazásával.

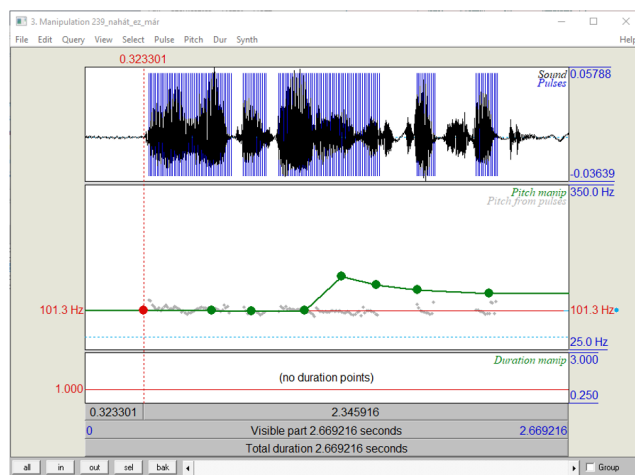
transzponálást két változatban végeztük el. Az első változatnál a szerzők által korábban említett F „alaphangot” vettük alapul (176 Hz), és ezt szállítottuk le a bemondáson mért 101 Hz-es mediánértékre (5. ábra).

A második változatnál a mondat elején, több szótagon is egyenletesen megjelenő G hangot vettük alapul (196 Hz), és ezt szállítottuk le a bemondáson mért



5. ábra. A *Nahát, ez már mégiscsak sok!* mondat első transzponált változata.

101 Hz-es mediánértékre, azaz 11,48 félhanggal transzponáltuk a bemondást lefelé (6. ábra).



6. ábra. A *Nahát, ez már mégiscsak sok!* mondat második transzponált változata.

Valószínűsíthető, hogy mindkét megoldás működőképes, a hallgatók el tudják fogadni. Az első esetben azonban a magasabb hangfekvés nagyobb feszültségi állapotot tükrözhet, a második, mélyebben megvalósuló dallam nyugodtabb beszélő benyomását kelti.

2.4. A módszer tesztelése

A következőkben pilot kutatást végeztünk annak megállapítására, hogy a fenti módszer alkalmazásával rekonstruált beszéddallamok valóban közvetítik-e – Fónagy és Magdics szóhasználatát megtartva – az adott érzelmeket, attitűdöket a hallgatók számára.

A 2.1. alfejezetben említett mondatok közül ötöt választottunk ki a pilot kutatáshoz, ezek dallama közönyt, örömet, szomorúságot, rémületet és gúnyt fejez ki. Két esetben módosítottunk a Fónagy és Magdics által közölt mondatokon azzal a céllal, hogy a hallgatókat a mondatokban előforduló szavak, kifejezések ne befolyásolják a kifejezett érzelem megítélésében. Így az eredeti, közönyt kifejező *Nem érdemes ezzel foglalkozni.* mondat helyett a *Ma is fogunk ezzel foglalkozni.*, illetve a szomorúságot tükröző *Kidobtak az állásomból.* helyett *Beszéltek az állásomról.* mondatokat használtuk. A pilot kutatáshoz használt mondatokat és a kottaképekből visszafejtett frekvenciaadatokat a 2. táblázat tartalmazza.

A frekvenciaértékek beállítását követően minden mondatot egységesen – a fenti második megoldás alapján – 11,48 félhanggal lefelé transzponáltunk, így a kottaképeknek megfelelően a dallam és a hangfekvés is megmaradt.

A kutatáshoz tesztlapot készítettünk, amely nem csak a fenti öt érzelmet, attitűdöt tartalmazta, hanem egy hatodikot, a meglepetést is. A kutatás résztvevői ötfokú skálán jelölték, hogy az adott érzelem, attitűd milyen mértékben tükröződik az elhangzott mondatban (1 = semennyire, 5 = nagyon erősen). A résztvevők mindegyik mondatot értékelték mind a hat érzelemre vonatkozó skálán.

A kutatásban 35 nappali tagozatos egyetemi hallgató vett részt (életkoruk 18–27 év), mindannyian magyar anyanyelvűek, hallószervi betegségről, halláskárosodásról nem számoltak be. A lehallgatás a Pécsi Tudományegyetem Művészeti Karának egyik tantermében, zavaró körülmények nélkül, több kisebb csoportban, számítógépről, jó minőségű, Xiaomi típusú Bluetooth hangszórón lejátszva történt. A kísérletvezető a jelen pilot kutatásban nem szereplő, a 2. fejezetben ismertetett mondat lejátszásával bemutatta a feladatot, ezzel együtt

2. táblázat. A pilot kutatáshoz használt mondatok és frekvenciaadataik.

1. közöny	<i>ma</i>	<i>is</i>	<i>fo</i>	<i>gunk</i>	<i>ve</i>	<i>le</i>	<i>fog</i>	<i>lal</i>	<i>koz</i>	<i>ni</i>
zenei hang	D'	H	H	H	A	A♭	H	A	G	G
frekvencia (Hz)	294	246	246	246	220	233	246	220	196	196
2. öröm	<i>kép</i>	<i>zeld</i>	<i>még</i>	<i>sem</i>	<i>ő</i>	<i>vizs</i>	<i>gáz</i>	<i>tat</i>		
zenei hang	349	294	349	246	262	277	262			
frekvencia (Hz)	F'	D'	F'	H	C'	C#'	C'			
3. szomorúság	<i>be</i>	<i>szél</i>	<i>tek</i>	<i>az</i>	<i>ál</i>	<i>lá</i>	<i>som</i>	<i>ról</i>		
zenei hang	A	F	F	F	G	F	F♭	F		
frekvencia (Hz)	220	175	175	175	196	175	164	175		
4. rémület	<i>be</i>	<i>szél</i>	<i>jen</i>	<i>sán</i>	<i>dor</i>	<i>mi</i>	<i>tör</i>	<i>tént</i>		
zenei hang	C'	C'	C'	C'	C'	D'	C'	C'		
frekvencia (Hz)	262	262	262	262	262	294	262	262		
5. gúny	<i>mi</i>	<i>cso</i>	<i>da</i>	<i>ér</i>	<i>zõ</i>	<i>szív</i>				
zenei hang	G	F	F	G	F	G				
frekvencia (Hz)	196	175	175	196	175	196				

megbizonyosodott arról, hogy a résztvevők mindegyike számára egyértelmű a feladat és a tisztán hallják a lejátszott mondatokat. Egy hallgató a tesztlap megtekintésekor jelezte, hogy nem ismeri a *közöny* szót, számára a kutatás vezetője példákkal illusztrálva elmagyarázta a szó jelentését.

A tesztmondatok lejátszása minden csoportban azonos körülmények között történt. Először a hallgatók a közönyt, majd – egy, a kutatók által előzetesen megállapított, véletlenszerű sorrend szerint – a többi érzelmet, attitűdöt kifejező mondatot hallgatták meg (öröm, szomorúság, rémület, gúny) és értékelték a fenti módon. Minden mondat lejátszása egymás után kétszer történt, az első lehallgatás után a hallgatók rögzítették az adott mondattal kapcsolatos benyomásait, ezt követően még egyszer meghallgathatták ugyanazt a mondatot, korrigálhatták válaszukat.

A számításokat az *R* szoftverben (4.4.1 verzió, R Core Team, 2024) végeztük el a χ^2 (goodness-of-fit) próba alkalmazásával. Minden esetben nullhipotézis-

ként azt feltételeztük, hogy az egyes skálaértékre azonos arányban érkeznek válaszok. Ahol ettől szignifikáns eltérés mutatkozott, a kiugró cellaértékeket a standardizált reziduumok (SR) vizsgálatával kerestük meg. Kiugrónak tekintettük az abszolút értékben 1,96-nál nagyobb SR-értékeket. Egyes esetekben nem szignifikáns χ^2 ellenére is határértéken túli SR-értékek adódtak, ekkor a nem szignifikáns χ^2 értéket tekintettük mérvadónak. A hatás nagyságát a Cramér-féle V segítségével állapítottuk meg (*rcompanion* csomag; Mangiafico, 2024), a kapott V -értékek értelmezését Mangiafico (2016) alapján végeztük.

3. Eredmények

Mindegyik mondat esetében megvizsgáltuk, hogy hány fő választotta az egyes skálaértékeket a különböző érzelmeknél. Előfordult, hogy egy vagy több skálaértéket egyetlen hallgató sem választotta, így a nullhipotézisnek megfelelően a következő elméletileg várt cellaértékekkel dolgoztunk: öt kategória esetében 7, négy kategóriánál 8,75, három kategóriánál 11,67.

3.1. A közönyt kifejező dallammal megvalósított mondat

A skálaértékek eloszlását és a számítások eredményét a 3. táblázat tartalmazza. Itt, illetve a 4., 5., 6. és 7. táblázatokban félkövér kiemelés jelzi azokat az észlelt érzelmeket, amelyek esetében a skálaértékek eloszlása a nullhipotézisnek megfelelően nem különbözik szignifikánsan az egyenletes eloszlástól.

Az észlelt közöny esetében kapott nem szignifikáns p -érték azt mutatja, hogy kiegyenlített azon válaszadók aránya, akik szerint a lejátszott mondat kifejezi a közönyt, illetve akik szerint nem. Az összes többi észlelt érzelem esetében a skálaértékek arányos megoszlását feltételező nullhipotézis elvethető. Az adatok eloszlásából világosan látszik, hogy a válaszadók ezekben az esetekben jellemzően a legalacsonyabb skálaértéket jelölték meg. Ahol szignifikáns eltérés adódott, a V értékek erős összefüggést jeleznek, továbbá az SR-értékek megerősítik az 1-es skálaérték kiugróan gyakori ($> 1,96$), illetve a magasabb skálaértékek elvártól szignifikánsan ritkább ($< -1,96$) előfordulását.

3. táblázat. Közöny.

észlelt érzelem	1	2	3	4	5	χ^2	df	p	V
közöny	2	8	8	11	6	6,285	4	0,178	0,212
SR	-2,11	0,42	0,42	1,69	-0,42				
gúny	27	3	3	2	0	50,829	3	< 0,001	0,695
SR	7,12	-2,24	-2,24	-2,63					
rémület	23	5	4	2	1	47,143	4	< 0,001	0,580
SR	6,76	-0,84	-1,26	-2,11	-2,53				
szomorúság	13	7	10	4	1	12,857	4	< 0,05	0,303
SR	2,53	0	1,26	1,26	-2,53				
öröm	23	7	4	1	0	33	3	< 0,001	0,560
SR	5,56	-0,68	-0,85	-3,02					
meglepetés	28	5	2	0	0	33,686	2	< 0,001	0,703
SR	5,85	-2,39	-3,46						

3.2. Az örömet kifejező dallammal megvalósított mondat

A 4. táblázat foglalja össze az adatokat és a számítások eredményeit.

4. táblázat. Öröm.

észlelt érzelem	1	2	3	4	5	χ^2	df	p	V
közöny	19	8	7	1	0	19,286	3	< 0,001	0,428
SR	4,00	-0,29	-0,68	-3,02					
gúny	28	4	2	1	0	57	3	< 0,001	0,736
SR	7,51	-1,85	-2,63	-3,02					
rémület	18	8	1	5	3	25,429	4	< 0,001	0,426
SR	4,64	0,42	-2,53	-0,54	-1,69				
szomorúság	23	3	7	2	0	32,543	3	< 0,05	0,556
SR	5,56	-2,24	-0,68	-2,63					
öröm	10	6	5	10	4	4,571	4	0,334	0,180
SR	1,26	-0,42	-0,84	1,26	-1,26				
meglepetés	3	3	11	10	8	8,285	4	0,081	0,243
SR	-1,69	-1,69	1,69	1,26	0,42				

A nullhipotézis az öröm és a meglepetés észlelésekor igazolódott, azaz a megfigyelt eloszlások ezekben az esetekben nem különböznek szignifikáns mértékben az adatok elvárt, egyenletes eloszlásától. A hallgatók egy része szerint tehát örömet és meglepetést fejez ki a lejátszott mondat, más részük szerint viszont nem. A meglepetés esetében azonban a p -érték tendenciaszerű kapcsolatra utal, a 3-as érték közel szignifikáns mértékben gyakrabban, az 1-es és 2-es érték viszont

közel szignifikáns mértékben ritkábban fordult elő, mint az elméletileg várt gyakoriság. A többi esetben viszont elvethető a nullhipotézis, azaz az elhangzott mondatot egyértelműen nem közönyösnek, gúnyosnak, rémületet vagy szomorúságot kifejezőnek ítélték a hallgatók. A szignifikáns esetekben az 1-es skálaérték kiugróan gyakori előfordulását igazolták a magas SR-értékek ($> 1,96$), a V pedig erős hatásra utal.

3.3. A szomorúságot kifejező dallammal megvalósított mondat

Az 5. táblázat foglalja össze az adatokat és a számítások eredményeit.

5. táblázat. Szomorúság.

észlelt érzelem	1	2	3	4	5	χ^2	df	p	V
közöny	11	4	10	3	7	7,142	4	0,178	0,128
SR	1,69	-1,26	-1,26	-1,69	0				
gúny	23	3	4	4	1	46,571	4	$< 0,001$	0,576
SR	6,76	-1,69	-1,26	-1,26	-2,53				
rémület	12	12	8	2	1	16	4	$< 0,01$	0,338
SR	2,11	2,11	0,42	-2,11	-2,53				
szomorúság	8	8	6	11	2	6,285	4	0,178	0,212
SR	0,42	0,42	-0,42	1,69	-2,11				
öröm	26	5	3	1	0	46,257	3	$< 0,001$	0,663
SR	5,56	-0,68	-0,85	-3,02					
meglepetés	21	5	5	2	2	36,286	4	$< 0,001$	0,509
SR	5,91	-0,84	-0,84	-2,11	-2,11				

A nullhipotézis a közöny és a szomorúság esetében tartható meg, azaz a hallgatók egy része szerint a szándékunk szerint szomorúságot kifejező dallammal kimondott mondat kifejezi a közönnyt és a szomorúságot, más részük szerint viszont nem, és arányuk kiegyenlített. A szignifikáns p -értékek arra utalnak, hogy a hallgatók szerint a mondat egyértelműen nem fejez ki gúnyt, rémületet, örömet, meglepetést. Ezekben az esetekben az SR-értékek az 1-es skálaérték kiemelkedően magas, míg a magasabb skálaértékek elvártánál szignifikánsan ritkább előfordulását igazolták. A V a szignifikáns esetekben erős összefüggést mutat.

3.4. A rémületet kifejező dallammal megvalósított mondat

A 6. táblázat foglalja össze kapott eredményeket.

6. táblázat. Rémület.

észlelt érzelem	1	2	3	4	5	χ^2	df	p	V
közöny	4	3	8	7	13	8,857	4	0,064	0,251
SR	-1,26	-1,69	0,42	0	2,53				
gúny	11	5	8	10	1	9,42	4	0,051	0,259
SR	1,69	-0,84	0,42	1,26	-2,53				
rémület	24	4	4	3	0	35,514	4	< 0,01	0,581
SR	5,95	-1,85	-1,85	-2,24					
szomorúság	23	9	3	0	0	18,057	2	< 0,001	0,508
SR	4,06	-0,95	-3,10						
öröm	30	2	2	1	0	68,886	3	< 0,001	0,81
SR	8,29	-2,63	-2,63	-3,02					
meglepetés	23	8	1	3	0	33,914	3	< 0,001	0,568
SR	5,56	-0,29	-3,02	-2,24					

A rémületet kifejező mondat esetében két nem szignifikáns, de tendenciaszerű összefüggés adódott. A hallgatók tendenciaszerűen inkább közönyösnek érezték a mondatot, a gúny esetében viszont az 5-ös skálaértéktől eltérő vélemények voltak tendenciaszerűen túlsúlyban. A hallgatók szerint a mondat egyértelműen nem fejez ki rémületet, szomorúságot, örömet, meglepetést. A szignifikáns esetekben a V erős összefüggést mutat.

3.5. A gúnyt kifejező dallammal megvalósított mondat

A 7. táblázat foglalja össze kapott eredményeket.

Az adatok alapján az állapítható meg, hogy a hallgatók nem szignifikánsan, de tendenciaszerűen gúnyosnak ítélték a szándékunk szerint gúnnal ejtett mondatot. Az összes többi esetben szignifikánsan elutasítóak voltak, azaz a mondat nem fejezett ki közönyt, rémületet, szomorúságot, örömet, meglepetést a kísérleti személyek számára. A V -értékek erős összefüggést mutatnak.

3.6. Az eredmények összefoglalása

Általános megállapításként azt fogalmazhatjuk meg, hogy a feltételezett érzelmek – bár felismerésük nem volt egyértelmű – általában nagyobb mértékben

7. táblázat. Gúny.

észlelt érzelem	1	2	3	4	5	χ^2	df	p	V
közöny	10	12	9	2	2	12,571	4	< 0,05	0,299
SR	1,26	2,11	0,84	-2,11	-2,11				
gúny	7	4	8	3	13	8,857	4	0,064	0,251
SR	0	-1,26	0,42	-1,69	2,53				
rémület	31	2	1	0	0	51,235	2	< 0,001	0,868
SR	7,15	-3,39	-3,75						
szomorúság	23	4	4	4	0	30,943	3	< 0,001	0,542
SR	5,56	-1,85	-1,85	-1,85					
öröm	22	3	4	5	1	41,429	4	< 0,001	0,544
SR	6,33	-1,69	-1,26	-0,84	-2,53				
meglepetés	12	8	10	2	3	10,857	4	< 0,05	0,278
SR	2,11	0,43	1,26	-2,11	-1,69				

tükröződtek a lejátszott mondatokban, mint a többi öt érzelem vagy attitűd többsége. A gúny esetében volt a felismerés a legpontosabb, itt közel szignifikáns eredményt adott a számítás. A közönyös, az örömteli és a szomorú dallammal ejtett mondatok esetében a nullhipotézis igazolódott, azaz a skálaértékek eloszlása nem különbözött szignifikánsan a véletlen eloszlástól. Amikor azonban a hallgatók arról nyilatkoztak, hogy ugyanezek a mondatok milyen mértékben tükröznek más érzelmeket, a legtöbb esetben alacsony skálaértékek adódtak, azaz a hallgatók elutasították, hogy ezeket a további érzelmeket az adott dallammal megvalósított mondat kifejezné. De hasonló, elutasító eredményt kaptunk a – szándékunk szerint – rémületet kifejező mondat érzelmi töltetének visszakódolásakor is: a válaszokból egyértelműen az derül ki, hogy a mondat a hallgatók szerint nem tükröz rémületet.

4. Következtetések

Műhelytanulmányunkban annak a lehetőségét vizsgáltuk, hogy a Fónagy Iván és Magdics Klára által 1967-ben közölt kottaképek alapján szoftveres úton rekonstruálhatók-e beszéddallamok. A bemutatott módszerrel ez technikailag megvalósítható, azonban kérdéses, hogy a hallgatók megbízhatóan vissza tudják-e kódolni a dallamok által, az említett szerzők szerint kifejezett érzelmeket, atti-

tűdöket. Pilot kutatásunk eredményei azt mutatják, hogy önmagában a dallam rekonstrukciója nem ad elegendő támpontot ahhoz, hogy a hallgató egyértelműen azonosítsa a meghatározott érzelmet, ahhoz azonban általában igen, hogy milyen érzelmet, attitűdöt biztosan *nem* fejez ki. Az érzelmek megfelelő visszakódolásánál megfigyelt bizonytalanság arra utalhat, hogy az érzelmet tükröző dallam bár lényeges, de nem egyetlen prozódiai eleme az érzelmes beszédnek, és vélhetően önmagában még nem elegendő az érzelmek pontos észleléséhez. Emellett bizonyos esetekben tévesztéseket is megfigyeltünk. Az örömet kifejező mondat tendenciaszerűen inkább meglepetést tükrözött a hallgatóknak. A rémületet kifejező pedig szintén tendenciaszerűen inkább a közöny érzetét keltette, de a gúny elutasítása sem volt egyöntetű, csak tendenciaszerű. A tévesztések okainak feltárása további feladatokat jelent. Kutatásunkban más prozódiai és akusztikai tényezőket, paramétereket nem változtattunk meg, márpedig ismeretes az intenzitás, az energia- vagy a tempóviszonyok szerepe a különböző érzelmek kifejeződésekor (Scherer, 2003, 2013). Ezek beemelése a további kísérletekbe feltehetően javítja az érzelmek felismerésének sikerességét.

Mindezek mellett az sem zárható ki, hogy az 1967-ben rögzített beszéd-dallamokat a mai hallgató kissé máshogy észleli, mint a korabeli beszélők, de az is elképzelhető, hogy a zenei hangokat alaphétközre átfordító módszerünk sem hibátlan. Hiányossága még a kutatásunknak, hogy nem alkalmaztunk semleges érzelmi töltetű mondatokat, amelyekkel összehasonlításokat lehetett volna végezni. Erre Fónagy és Magdics munkája lehetőséget teremt, mivel minden esetben közölték a semleges érzelmi állapotban megvalósított mondatok dallamának kottaképét. További kutatásokban erre ki kell térni.

Amennyiben sikerül további fejlesztőmunkával a hallgatók számára felismerhetővé tenni a kották formájában lejegyzett érzelmeket, akkor javaslatot tehetünk a közölt dallamok különféle alkalmazásaira. Chow és Brown (2018) is említ néhány ilyen alkalmazást, például a gépi beszéd javítását, vagy az intonáció tanítását az idegennyelv-oktatásban. Ehhez hozzátehetjük még a művészi beszéd tanításának támogatását, vagy a beszélő érzelmi állapotának, attitűdjének észlelésével kapcsolatos kísérleteket. Elképzelhető ugyanis, hogy ugyanazon

mondat dallamában egy adott helyen már félhangnyi különbség is hordozhat olyan jelentéstöbbletet, amely alapvetően más kontextusba helyezi a mondatot. Az itt vázolt módszer ilyen kutatásokra is alkalmas, és a beszédprodukción és a beszédészlelés eddig még kevésbé kutatott területeire is bepillantást nyújthat.

Hivatkozások

- Bartók, J. (1969). A beszéddallam lejegyzésének kérdéseiről. In D. Pais, & L. Benkő (Eds.), *Dolgozatok a hangtan köréből. Nyelvtudományi Értekezések 67* (p. 88–92). Budapest: Akadémiai Kiadó.
- Bartók, J. (1974). Egyéni és társadalmi érvényű elemek a köznyelvi hanglejtésben. In *Általános Nyelvészeti Tanulmányok X. A nyelv hangdomíniuma* (p. 5–20). Budapest: Akadémiai Kiadó.
- Boersma, P., & Weenik, D. (2019). Praat: Doing phonetics by computer. v. 6.0.52. URL: www.praat.org.
- Boros, R. (1969). A magyar beszéddallam jegyzése. In D. Pais, & L. Benkő (Eds.), *Dolgozatok a hangtan köréből Nyelvtudományi Értekezések 67* (p. 76–87). Budapest: Akadémiai Kiadó.
- Boros, R. (1971). Néhány hétköznapi beszéddallam. *Magyar Nyelvőr*, 95, 234–236.
- Boros, R. (1975). Beszéddallamok. *Magyar Nyelvőr*, 99, 41–46.
- Bowling, D., Gill, K., Choi, J., Prinz, J., & Purves, D. (2010). Major and minor music compared to excited and subdued speech. *The Journal of the Acoustical Society of America*, 127, 491–503. doi:10.1121/1.3268504.
- Chow, I., & Brown, S. (2018). A musical approach to speech melody. *Frontiers in Psychology*, 9, 247. doi:10.3389/fpsyg.2018.00247.
- Csúry, B. (1925). A szamosháti nyelvjárás hanglejtésformái. *Magyar Nyelv*, XXI, 1–21.

- Fónagy, I., & Magdics, K. (1967). *A magyar beszéd dallama*. Budapest: Akadémiai Kiadó.
- Gocsál, Á., & Urbanics, T. (2022). Vocal pleasantness ratings explained by a musical approach. In M. Nourbakhsh, & N. Chahartagh (Eds.), *Proceedings Book. The Second International Conference on Laboratory Phonetics & Phonology (ICLPP)* (p. 59–66). Tehran, Iran: Booy-e Kaqaz (BOOKA) Publication.
- Gombocz, Z. (1907). A mondat zenei hangsúlyáról. *Urania*, 8, 129–131. URL: http://real-j.mtak.hu/10730/1/MTA_Urania_1907.pdf. URL:.
- Grácz, T., Gósy, M., Krepsz, V., Markó, A., Huszár, A., Damásdi, N., & Gocsál, Á. (2020). Az alapfrekvencia jellemzői az életkor és nem függvényében. In A. Fóris, A. Bölcskei, J. Bóna, T. Grácz, & A. Markó (Eds.), *Nyelv, kultúra, identitás. Alkalmazott nyelvészeti kutatások a 21.századi információs térben. III. Fonetika*. Budapest: Akadémiai Kiadó. URL: https://mersz.hu/hivatkozas/m675nyki3f_23 elérés forrás.
- Hegedűs, L. (1930). *Magyar hanglegjtésformák grafikus ábrázolása*. Királyi Magyar Egyetemi Nyomda.
- Hirst, D., & Cristo, A. (1998). A survey of intonation systems. In D. Hirst, & A. Cristo (Eds.), *Intonation Systems. A Survey of Twenty Languages* (p. 1–44). Cambridge: Cambridge University Press.
- Jones, D. (1909). Intonation curves. a collection of phonetic texts, in which intonation is marked throughout by means of curved lines on a musical stave. URL: <https://ia800903.us.archive.org/23/items/intonationcurves00jonerich/intonationcurves00jonerich.pdf> uRL:.
- Magdics, K. (1963). From the melody of speech to the melody of music. *Studia Musicologica Academiae Scientiarum Hungaricae*, 4, 325–346. URL: <http://www.jstor.com/stable/901395>. URL:.

- Mangiafico, S. (2016). Summary and analysis of extension program evaluation in r, version 1.20.07. URL: <https://rcompanion.org/handbook> revised 2024.
- Mangiafico, S. (2024). rcompanion: Functions to support extension education program evaluation. r package version 2.4.36. URL: <https://cran.r-project.org/web/packages/rcompanion/index.html>.
- Markó, A. (2013). Kérdő megnyilatkozások a spontán beszédben. *Beszéd kutatás*, (p. 42–59).
- Markó, A. (2017). Hangtan. In G. Nagy (Ed.), *Nyelvtan* (p. 75–203). Budapest: Osiris Kiadó.
- Murdoch, J. (1883). *A Plea for Spoken Language*. Cincinnati, New York: Van Antwerp, Bragg & Co. URL: <https://play.google.com/books/reader?id=WmQQAAAAAYAAJ&pg=GBS.RA2-PA54&hl=hu>.
- MuseScore, B. (2021). Musescore 3.6.2.5. URL: <https://musescore.org/hu>.
- R Core Team (2024). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: <https://www.r-project.org/>.
- Regner, T. (1861). A magyar nyelv kiejtése. *Akadémiai Értesítő, II*, 363–447. URL: http://real-j.mtak.hu/35/1/AkademiaiErtesito_1861-1862_Nyelv.pdf. URL:.
- Robledo, J., Hurtado, E., Prado, F., Román, D., & Cornejo, C. (2016). Music intervals in speech: Psychological disposition modulates ratio precision among interlocutors’ nonlocal f0 production in real-time dyadic conversation. *Psychology of Music*, 44, 1404–1418. doi:10.1177/0305735616634452.
- Scherer, K. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40, 227–256.
- Scherer, K. (2013). Vocal markers of emotion: Comparing induction and acting elicitation. *Computer Speech and Language*, 27, 40–58.

- Scripture, E. (1906). Researches in experimental phonetics. the study of speech curves. URL: <https://archive.org/details/researchesinexp01scrigoog>
uRL:.
- Tolnai, V. (1915). Adatok a magyar hanglejtéshez. *Magyar Nyelv*, 11, 51–59.
- Tóth, A. (2016). Kérdő funkciójú megnyilatkozások kisiskolások beszédében. *Beszéd kutatás*, (p. 43–58).
- Viëtor, W. (1897). Kleine phonetik des deutschen, englischen und französischen. URL: https://openlibrary.org/works/OL13122732W/Kleine_Phonetik_des_Deutschen_englischen_und_franz%C3%B6sischen
uRL:.