

A FilmHír adatbázis fejlesztése régi filmhíradókban elhangzó beszéd vizsgálatához

Gocsál Ákos^{1,2}

¹*PTE Művészeti Kar, Zeneművészeti Intézet*

²*ELKH Nyelvtudományi Kutatóközpont*

Abstract

This paper presents the first steps of the development of a speech database that supports the research of the spoken material in the early sound newsreels produced in Hungary. Since there was no recording protocol, and the recordings were not made in a controlled environment, the development of this database differs from that of other, purpose-built databases, such as the BEA Hungarian Spontaneous Speech Database, in several ways. There is no balanced selection of speakers or speech styles, and the speech material may contain elements unpredictable to the developer. In the first phase of development, speech materials of 140 newsreels, made in 1931 and 1932, were extracted and annotated (only pause-to-pause sections). A table was created with basic metadata. It was revealed that the newsreels contain a variety of speech styles, such as conversations, public speeches, narrations, commands, and interviews. In the second phase, further metadata were added, including background noises, music, utterances in foreign languages, distortions, etc. Research experiences and possibilities offered by this database are also discussed as well as directions for its further development.

1. Bevezetés

Ebben a tanulmányban egy olyan adatbázis összeállításával kapcsolatos eddigi eredményeket és tapasztalatokat ismertetjük, amely a magyar filmhíradók – nyelvészeti, fonetikai szempontból mindeddig feldolgozatlan – beszédanyagának kutatását teszi lehetővé.

A nemzetközi és a hazai szakirodalomban is sok évtizedes hagyománya van a fonetikai kutatásokat segítő adatbázisok, korpuszok használatának. Ezek többféle szempont alapján is jellemezhetők és csoportosíthatók. Néhány példát idézünk a magyar szakirodalomból. Életkor, nem, iskolai végzettség szerint

Email address: gocsal.akos@pte.hu (Gocsál Ákos)

nagy változatosságot mutató felnőtt beszélők közreműködésével készült a magyar spontánbeszéd-adatbázis (BEA), amely többféle beszéd típus tanulmányozását is lehetővé teszi (Gósy et al., 2012). Más adatbázisok specifikus csoportok, például gyerekek, tinédzserek beszédének jellegzetességeinek kutatását segítik, magyarra a GABI (Bóna et al., 2014) és a Tini BEA (Gyarmathy & Neuberger, 2015) adatbázis említendő. Más adatbázisok fonációs rendellenességgel beszélők, depressziósok, Parkinson-kórosok beszédét tartalmazzák (Sztahó et al., 2019), de ismeretesek olyanok is, amelyeket speciális célokra, például zajos, vagy érzelmes beszédet automatizáltan felismerő, feldolgozó módszer fejlesztéséhez állítottak össze (Sztahó et al., 2011). Szintén hazai fejlesztők munkája a HuComTech multimodális korpusz, amely a hanggal együtt videofelvételeket is tartalmaz a beszédeseményekről. Az igen gazdag, 47 szintű annotáció az akusztikai és nyelvi jellemzők mellett a nemverbális jelek lejegyzését – testtartás, gesztusok, mimika, tekintet stb. – is tartalmazza, így az egyes közlések, társalgások még tágabb kontextusban kutathatók, mint ha csak a hang állna rendelkezésre (Hunyadi et al., 2018).

Az érzelmes beszéd kutatásához készített adatbázisokkal kapcsolatban egy igen lényeges kérdéssel találkozta a kutatók. Felmerült ugyanis, hogy jelentősen eltérnek egymástól a mesterséges – azaz eljátszott vagy tudatosan előidézett – helyzetben, illetve valós körülmények között, spontán módon keletkezett adatok (Campbell, 2000; Campbell et al., 2006), így ennek megfelelően az adatbázisok is két fő csoportra – „mesterséges” és „valós életbeli” – oszthatók (Pittermann et al., 2010, 42). Hasonló, de további részletekkel kiegészített felosztást közölt Aubergé et al. (2003). Ők az érzelmes beszédet tartalmazó korpuszokat három, egymásra merőleges dimenzió szerint osztályozták. A korpuszok készülhetnek *in vivo* vagy *in vitro* módszerekkel. Az előbbi a kutatói beavatkozás nélküli, az utóbbi a laboratóriumi, tudatosan beállított körülmények közötti adatrögzítésre vonatkozik. A második jellemző a beszédjellemzők – interakciós helyzet, nyelvi tartalom, fonetikai tulajdonságok – fölötti kutatói kontroll jelenléte vagy hiánya. A harmadik jellemző pedig az, hogy a beszélők eljátszák-e az adott érzelmet, vagy pedig a beszéd autentikusan tartalmazza-e azt. Pittermann

et al. (2010, 43) mindezek mellett további csoportosítási szempontként megemlíti, hogy professzionális színészek vagy naiv beszélők bemondásában készült-e a felvétel, illetve maga a nyilatkozat tartalmaz-e érzelmi töltetet vagy szemantikailag semleges-e.

A mesterséges vagy valós életbeli beszédesemények használata, és az ebből adódó különbségek problematikája azonban túlmutat az érzelmes beszédből épített adatbázisokon. Más céllal is állítanak össze valós életből vett beszédekből – például parlamenti vitákból (Solberg & Ortiz, 2022), telefonos ügyfélszolgálatok beszélgetéseiből (Vicsi & Sztahó, 2009) vagy televíziós műsorokból (Teleki et al., 2005; Lleida et al., 2020) – adatbázisokat, és rendszeresen felmerülő kérdés, hogy mennyire tekinthetők természetesnek az adatbázisba beemelt beszédmin-ták. Campbell (2000) szerint a televíziós műfajok esetében a hírműsorokban bemutatott interjúk általában természetesebbek, mint a talk-show-típusú műsorokban és valóságshow-kban elhangzó beszédek, ugyanis ezekben az esetekben előfordulhat a helyzetek tudatos megjátszása. Grimm et al. (2008) viszont beszámolnak egy érzelmes beszédet tartalmazó, német talkshowadásokból összeállított korpuszról, amely előre nem megírt, a műsor vendégei között megvalósult autentikus társalgásokat tartalmaz. A természetesség problémájával foglalkozott még Devillers & Vidrascu (2007) az általuk összeállított, 688, segélyhívó vonalon lezajlott beszélgetésből álló korpusz bemutatásakor. A szerzők rámutattak, hogy vészhelyzetben, amikor valaki mentőt hív, a beszédben tükröződő érzelmek természetesen adódnak, nem megjátszottak.

A valós életből vett, autentikus helyzetekben rögzített beszédesemények természetességének azonban „ára” van, mégpedig éppen a kontroll, azaz előzetesen meghatározott szempontok, egységes körülmények, felvételi protokoll hiánya, s ezt a kutatónak figyelembe kell vennie (Douglas-Cowie et al., 2003). Mivel ennek vannak különféle technikái, a kétféle – „mesterséges” és „valós életbeli” – adatokból felépített adatbázisok között nincs olyan éles határ, mint amilyenre a fentiek alapján következtetni lehetne. A kontrollált körülmények között összeállított adatbázisok fejlesztői számos olyan módszert alkalmaznak, amellyel a felvételi körülményeket közelítik az autentikus, természetes helyzetekéhez. Ide

sorolhatók például azok a meghatározott témából kiinduló, de kötetlen társalgások, amelyek jól modellálják a laboratóriumon kívül megvalósuló helyzeteket. Ilyen társalgásokat találunk a BEA adatbázisban (Gósy et al., 2012; Neuberger et al., 2014), vagy egy külföldi példát is említve, a Cieri et al. (2004) által összeállított telefonos adatbázisban, amelynek egymást nem ismerő beszélői egy véletlenszerűen kiválasztott témáról folytattak rövid beszélgetéseket. Ugyanakkor az előzetes kutatói kontroll, tudatos tervezés nélkül rögzített beszédmintákból is összeállítható olyan adatbázis, amelyek több szempontból is kontrolláltnak tekinthetők. Ilyenek például a rádióbemondók hírolvasásai, amelyeknél – még akkor is, ha a kutató nincs ilyen információk birtokában – feltételezhető, hogy minden esetben hangszigetelt helyiségben, professzionális eszközökkel, zavaró külső hatások nélkül történt a felolvasás rögzítése.

Mindezek alapján belátható, hogy elengedhetetlen feladat az adatbázis fejlesztésének dokumentálása, az egyes elvi megfontolások, döntések, fejlesztési szempontok precíz rögzítése és mindezek közzétevése felhasználók számára. Jól példázza ezt a BEA adatbázis ismertetése (Gósy et al., 2012). Munkájukban a szerzők részletesen bemutatták, hogy a beszédanyagok rögzítése milyen körülmények között, milyen protokoll szerint történt. A beszélők az interjúkészítőtől többféle feladatot kaptak, így különféle beszéd típusok rögzítése vált lehetővé. A fejlesztők törekedtek a nők és a férfiak, illetve a különböző életkorú beszélők arányos megjelenítésére, továbbá tendenciaszerűen figyeltek az iskolázottság alakulására is. A hangrögzítést minden esetben ugyanabban a zajszigetelt helyiségben, professzionális felszereléssel végezték, így a felvételi körülmények minden adatközlőnél megegyeztek. Különbségek legfeljebb az egyes beszélők között lehettek. Ezeknek az információknak a birtokában a felhasználó megismerheti az adatbázis által nyújtott lehetőségeket, annak esetleges korlátait, így ezek tudatában választhatja ki a kutatási céljainak megfelelő beszédmintákat, értelmezheti kutatási eredményeit.

Az előzőekben felvázolt problémákkal, dilemmákkal a FilmHír adatbázis fejlesztésének kezdeti szakaszától kezdődően találkoztunk. A filmekben talált beszédanyag jelentős részét valós életbeli események során rögzítették, így a kuta-

tói kontroll hiánya, az előadódó adatok esetlegessége és rendezetlensége, a beszéd típusok sokfélesége miatt egyedi megoldást kellett kidolgozni ahhoz, hogy a tervezett adatbázis jól átlátható, könnyen lekérdezhető legyen. Az alábbiakban a fejlesztési folyamat részleteit mutatjuk be, amelyek nem csak a felhasználás és a továbbfejlesztés lehetőségeit vetítik előre, hanem kiindulópontként szolgálhatnak más, hasonlóan nem kutatási céllal rögzített beszédanyagok adatbázisban való feldolgozásához is.

2. A FilmHír adatbázis felépítése

A magyar filmhíradók az interneten bárki számára szabadon hozzáférhetők (<https://filmhiradokonline.hu/>). A filmhíradókon végzett kutató-fejlesztőmunkát megelőzően tájékoztunk azok felhasználhatóságáról. A honlapon közzétett tájékoztató (<https://filmhiradokonline.hu/intro.php>), illetve a Magyar Nemzeti Filmintézet munkatársával történt konzultáció megerősítette, hogy oktatási, kutatási, közismereti célra felhasználhatók a Filmhíradók Online anyagai.

Az első filmhírek előzetes megtekintése során világossá vált, hogy az adatbázis fejlesztési feladatai és alkalmazási lehetőségei sok szempontból jelentősen eltérnek a tervezett, protokoll szerint rögzített adatbázisokéitól. Négy olyan szempontot emelünk ki, amelyek mentén a legjelentősebb különbségek megfogalmazhatók:

1. A FilmHír adatbázis felépítése nem előre meghatározott protokoll szerint történik, így számítani kell arra, hogy véletlenszerűen bármilyen beszéd típus és bármilyen demográfiai jellemzővel rendelkező beszélő előfordulhat. Ebből adódóan a beszélőket, beszédmintákat illetően nem beszélhetünk reprezentativitásról, de még kiegyensúlyozottságról sem.
2. A filmhíradóknál számos, a felvételek akusztikai minőségét kedvezőtlenül befolyásoló tényezővel kell számolni. A kezdetleges hangrögzítési technológia, a hangsávot ért sérülések, filmszakadások, torzulások, a különféle, ki nem szűrt háttérzajok következtében akusztikai szempontból is rendkívül heterogén a filmhírek beszédanyaga.

3. Maga a rögzítés is változatos körülmények között, nem kutatói kontroll mellett történt, emiatt a spontaneitás és természetesség szempontjából sem kezelhetők egységesen a beszédminták. Több filmben megfigyelhető, hogy a beszélők vélhetően nem is voltak tudatában annak, hogy hangjukat rögzítik, míg máskor kifejezetten a kamerába beszéltek. Más esetekben pedig, például az ünnepi szónoklatok esetében az feltételezhető, hogy akkor is ugyanígy, vagy nagyon hasonló módon valósult volna meg a beszédesemény, ha a híradó nem rögzítette volna. Ezekben az esetekben tehát a beszédesemények a rögzítés tényétől függetlenül adottak voltak, autentikus, természetes helyzetekben valósultak meg, a kamera csak mellékesen volt jelen.
4. A meghatározott protokoll alapján készült, előzetesen megtervezett adatbázisok általában tetszőlegesen bővíthetők további beszélőkkel vagy beszéd típusokkal. A FilmHír számára azonban véges számú felvétel áll rendelkezésre. A fejlesztés során azonban törekszünk a teljességre, azaz valamennyi fellelhető filmhír beszédanyagának feldolgozására. Adatbázisunk ebből a szempontból hasonlít az ómagyar korpuszra, amelyek fejlesztői célul tűzték ki, hogy korpuszuk valamennyi fennmaradt nyelvemléket tartalmazza, ugyanakkor ők is említették a reprezentativitással, kiegyensúlyozottsággal kapcsolatos, korábban említett problémát a korlátozott nyelvi anyag miatt (Simon & Sass, 2012).

Az adatbázis felépítéséhez a filmhírek feldolgozását kronologikus sorrendben terveztük, így a jelen állapotban a legkorábbi, 1931. szeptember – 1932. december közötti hangos híradókkal foglalkozunk. Az egyes hírek megkülönböztetése két számadattal történt. Az első számadat az adott híradó (általában 6–9 hírből álló hírblokk) száma, a második pedig a híradón belüli filmhír száma. A legelső hangos hír ennek megfelelően a 398/1. számmal azonosítható. A teljes azonosító tartalmaz még néhány karaktert (MVH, Magyar Világhíradó), azonban ezt nem mindenhol tüntettük fel, mivel 1944 októberéig minden híradó esetében azonos volt.

Az adatbázis felépítésének volt egy első, feltáró jellegű fázisa, amely konkrét kutatási feladathoz szolgáltatott adatokat. A régi filmhíradókban hosszú évtizedeken keresztül alkalmaztak bemondókat a hírek felolvasására, így kézenfekvőnek tűnt a hírolvasás tempójának vizsgálata. A beszéd tempóértékei egyébként is könnyen vizsgálhatók más célból készített felvételeken, mivel tempó minden beszédeseményben jelen van, szemben egyes ritkán előforduló nyelvi struktúrákkal, továbbá előhívásukhoz nem szükségesek külön kutatói instrukciók, illetve viszonylag kevés annotációs munkát igényelnek (Seifart, 2021). A legkorábbi, 1931-32-ben készült hangos filmhírek megtekintése során azonban bebizonyosodott, hogy azok igen heterogén beszédanyagot tartalmaznak, a narrációs szövegek felolvasásán kívül számos más beszéd-típus is előfordul bennük. Ezért az egyes bemondások tempóértékein kívül azonosítottuk a beszéd-típusokat, és más adatokat is rögzítettünk, amelyekből világossá vált, hogy egy így felépített adatbázis számos további kutatási feladat megoldására lehet még alkalmas (Gocsál, 2022).

A fejlesztés jelenleg is zajló második fázisában a teljes eddigi adatbázist átvizsgáltuk és kiegészítettük olyan adatokkal, amelyek segítségével jóval pontosabban szűrhető az adatbázis, így kutató-fejlesztő feladatokhoz célzottan kereshető beszédminták.

A két fázis ismertetéséhez Drude et al. (2014) munkáját vesszük alapul, akik a korpuszfejlesztés során keletkező adatok három típusát különböztették meg. Az elsődleges adatok maguk a hang- vagy videófelvetelek. A másodlagos adatokat az annotáció során rögzített szövegek, illetve olyan időfüggő szimbólumok alkotják, amelyek a vizsgált beszédminta releváns tulajdonságait kódolják. A harmadik adattípushoz pedig a metaadatok tartoznak. Ezek az adatok az elsődleges és a másodlagos adatok nem időfüggő, általános tulajdonságait jelentik, például a rögzített esemény mi volt, hol, mikor, történt, ki beszél rajta stb., de ide tartozik a felvétel formátuma, kódolása, struktúrája, forrása, azonosító kódja és a helye az adatbázisban.

2.1. A fejlesztés első fázisa

2.1.1. Az elsődleges adatok

Az adatbázist alkotó elsődleges adatok maguk a filmhíradók. Mivel az adatbázis célja elsősorban a nyelvi-kommunikációs jelenségek tanulmányozásának támogatása, a filmhíradókból kinyerhető, megállapítható adatoknak csak egy meghatározott, az említett kutatási terület szempontjából releváns részhalmazával foglalkozunk. Értelemszerűen ide tartoznak a beszédet tartalmazó hangsvókok, de a kommunikatív aktust kísérő, és a filmen megjelenített minden olyan szöveges és képi információ is, amely segítségével megállapítható az adott beszédeseemény kontextusa, azonosítható a beszéd típus, esetleg a beszélő. Az adatbázis felépítése során fontos teendő volt az információhordozó egységek meghatározása, amelyek az adatbázis egyes rekordjait alkotják. Más kutatásban az egyedi híreket tekintették egységnek (Pincemin et al., 2020), azonban a jelen esetben ez egyelőre nem tűnt célszerűnek, ugyanis egy híren belül gyakran többféle beszédeseemény is megvalósult, és több beszélő is megszólalt. Így az adatbázis számára definiáltuk a „közlési egység” fogalmát, azaz az (1) egy filmhíren belül (2) ugyanazon beszélő (3) ugyanazon beszéd típushoz tartozó bemondását tekintettük az adatbázis elemi egységének, azaz rekordnak.

Döntést igényelt az is, hogy szerepeljen-e minden előforduló beszédeseemény az adatbázisban. A fejlesztés első fázisában kizártuk az idegen nyelvű és a csoportosan, tömegesen mondott szövegeket (pl. éljenzés), illetve a gyenge technikai minőségű, érthetetlen vagy erősen zajos bemondásokat, tehát csak a viszonylag tiszta, érthető bemondásokkal foglalkoztunk.

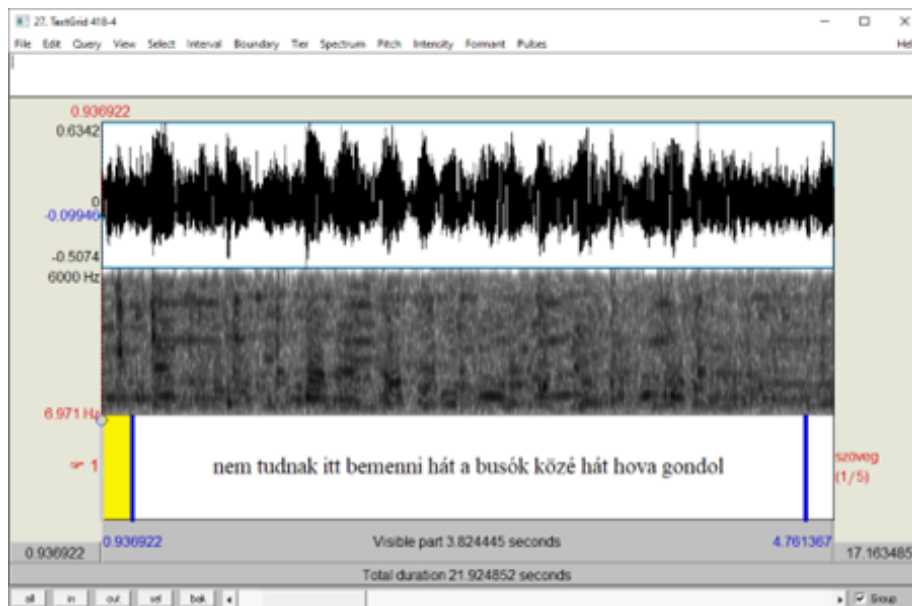
A vizsgált időszakból 61 híradó 441 filmhíre maradt fent. Ezekben a fenti elvek alapján összesen 258 közlési egységet találtunk. A filmhírek beszédet tartozó részleteit külön hangfájlokban, *.wav* formátumban elmentettük. Ezek formátuma *xxx-x-y.wav* volt, ahol *xxx-x* a híradó száma, *y* pedig a híradón belüli filmhír száma volt.

2.1.2. A másodlagos adatok

Az egyes közlési egységek annotálását a *Praat* programban végeztük (Borsma & Weenink, 2019). A fejlesztés első fázisában csak a szünettől szünetig tartó beszédszakaszokat jelöltük, és csak elhangzó szövegeket jegyeztünk le, felhasználva, de több esetben kiegészítve, korrigálva a filmhiradokonline.hu oldalon közzölt szövegeket. Ahol filmszakadás miatt ugrás volt az elhangzó szövegben, ott megjelöltük a szakadás előtti, illetve utáni, egészben elhangzó szó határait, ezzel a szakadást tartalmazó rövid részt kihagytuk a további feldolgozásból. A szöveg lejegyzéshez a BEA spontánbeszéd-adatbázis annotációjának első változatát vettük alapul (Gyarmathy et al., 2014), azaz fonológiai alapon, a helyesírási szabályoknak megfelelően rögzítettük az elhangzó beszédet. Az idegen szavakat, illetve a feltűnően sajátos ejtésű magyar szavakat kiejtés szerint írtuk, nagybetűket és írásjeleket nem használtunk. Ez az egyszerűsített lejegyzés már lehetővé tette a beszédszakaszok időtartamainak és az elhangzó beszédhangok számának meghatározását, így első kutatási feladatunkhoz az artikulációs tempó értékeit ki lehetett számítani. Ugyanakkor részletesebb annotálásra egyelőre nem vállalkoztunk, részben annak időigényessége miatt, részben pedig azért, mert nem ismerve a beszédanyagot, bizonytalan volt, hogy milyen mélységben, miket érdemes egyáltalán lejegyezni. Az 1. ábrán látható a 418/1. sz. filmhírből kiemelt *Nem tudnak itt bemenni, hát a busók közé, hát hova gondol* mondat megjelenítése. Az oszcillogram és a spektrogram is jól mutatja az igen erős háttérzaj jelenlétét, azonban a szöveg érthetőségét ez nem akadályozta. A beszédminőségnek ez a szakasza olyan kutatási feladatokra jól használható, amelyek nem igénylik az alacsony zajszintet, vagy épp ellenkezőleg, kifejezetten szükségesek hozzá, például beszédfelismerő rendszerek teszteléséhez.

2.1.3. A metaadatok

Ibrus & Ojamaa (2018) szerint a metaadatok a filmhírt és az általa megjelenített történeti eseményeket kulcsszavakból álló verbális nyelvvé fordítják át. A metaadatok sajátos metanyelvet alkotnak. Ez a metanyelv azonban nem szigorúan standardizált elemekből épül fel, ugyanis ebben az esetben nem lehetne



1. ábra. A 418/4. sz. filmhír lejegyzésének részlete.

új jelentéseket konstruálni, illetve magát az adatbázist sem lehetne fejleszteni különféle kutatási célokra. Mindebből az következik, hogy egy filmhír-adatbázis fejlesztésekor a metaadat-állomány létrehozása nagyban függ a fejlesztő szubjektív megítélésétől – mit tart fontosnak –, illetve attól, hogy a későbbi felhasználók igényeiről hogyan gondolkodik. A metaadatoknak számos típusa létezik. Egy más céllal, a spanyol köztelevízió adásaiból összeállított adatbázis fejlesztésekor rögzítették a beszélőket, beszédfordulókat, továbbá azt is, hogy a felvétel városi vagy falusi környezetben, illetve belső térben vagy külső helyszínen készült, osztott képernyő (több kamera képe is látszik-e), meleg vagy hideg időjárás volt, illetve nappal vagy éjszaka készült-e a felvétel (Leida et al., 2020).

A jelen adatbázis metaadat-struktúrájának kialakításakor a korábbiakhoz hasonló okok miatt egyszerű, de kutatási célra már használható rendszert igyekeztünk kialakítani. Egy Excel-táblában minden egyes, közlési egységre vonatkozóan az alábbi adatokat rögzítettük:

- a közlési egység azonosítói
 - sorszám az adattáblában
 - a honlapon közölt azonosító
 - a filmhíradó száma
 - a filmhír száma
- a filmhír megjelenésének ideje
 - évszám
 - hónap
- a beszélő személy
 - neve, ha ismert
 - születési éve, ha ismert
 - életkora, ha ismert
- az annotálás állapota
 - x, ha elkészült
- beszéd típus
- a filmhír témája néhány szóban
- megjegyzés
- link a filmhírhez

A beszéd típus feltüntetése csak első megtekintés alapján történt. A megjegyzéseknél azokat a lényeges információkat tüntettük fel, amelyek nem szerepeltek külön oszlopban, például azt, ha gyermek volt a beszélő, vagy hogy feltételezhetően megtanult szöveget mondott el.

Az első fázis során 441 filmhírt tekintettünk meg, ezek közül 140-ben találtunk feldolgozható beszédeseeményeket. Összesen 258 közlési egységet adatoltunk a fenti módszer alkalmazásával, ezek szünetek nélküli, összesített időtartama 3221,37 másodperc volt.

2.2. A fejlesztés második fázisa

A tizenöt hónapnyi híryanagy feldolgozása során számos olyan megfigyelést tettünk és tapasztalatot szereztünk, amely birtokában felmerült, hogy érdemes a teljes adatbázist újragondolni, más elvek alapján felépíteni.

2.2.1. Az elsődleges adatok

Az első fázisban kizárólag a jól érthető, magyar nyelvű beszédrészleteket emeltük ki a filmhírekből, és ezekből állítottuk össze az adattáblát. A fejlesztés során azonban több olyan adatfajta is előadódott, amelyekről úgy véltük, hogy rögzítésük jelentősen bővítheti az adatbázis felhasználhatóságát. Ide tartozik például az idegen nyelvű bemondások, a különböző torzítások, hibák, vagy az aláfestő zenék, különféle háttérzajok megjelenítése. Így a második fázis során azt az alapelvet követtük, hogy az adatbázis minden filmhírral kapcsolatban tartalmazzon valamilyen információt, akkor is, ha beszéd nincs benne, mivel ez is informatív lehet a kutató számára. Felülvizsgálva az előző gyakorlatot, ebben az esetben az adatbázist alkotó egységnek, rekordnak az egyedi filmhírt tekintettük, a jövőbeni kutatóra bízva azt, hogy saját kutatási feladatának megfelelően kiemelje abból a szükséges közlési egységeket.

Ebben a fázisban kétféle hangfájlt használtunk. A teljes híradót tartalmazó fájl *xxxx.wav* formátumban mentettük el, ahol *xxxx* a híradó száma volt, pl. *0405.wav*. A magyar beszédet tartalmazó filmhíreket pedig a könnyebb kezelhetőség miatt – mivel itt az elhangzó szövegeket részletesen lejegyeztük – külön fájlokban mentettük el *xxxx-y.wav* formátumban, ahol *y* az adott filmhír híradón belüli száma volt, pl. *0405-1.wav*.

2.2.2. A másodlagos adatok

A teljes híradókat tartalmazó hangfájlok annotálása két sávban történt. Az elsőben jelöltük a filmhír számát a híradón belül, a másodikban pedig a hallható hangeseményre jellemző címkéket jegyeztünk le az alábbiak szerint.

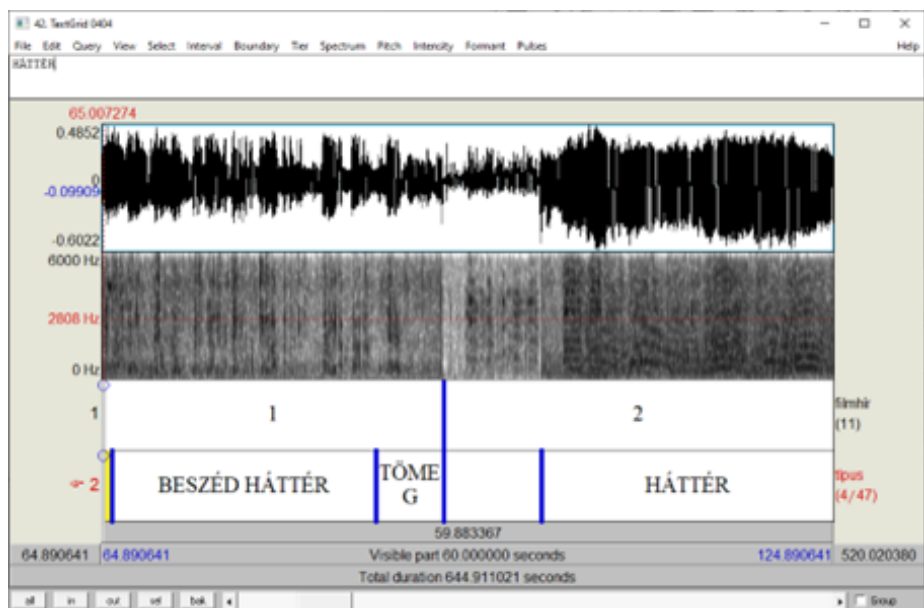
- BESZÉD: magyar nyelvű beszédesemény

- IDEGEN: idegennyelvű beszédesemény
- ÉNEK
- ZENE: a hír alatt megszólaló, a helyszínen rögzített vagy utólag alákevert zene (a hír főcíme alatti felvezető zenét nem jegyeztük le)
- TÖMEG: egyszerre, tömegesen mondott beszéd, pl. éljenzés, bravózás, skandálás
- TORZ: sérült, erősen zajos, torzult, érthetetlen felvétel
- SZAKAD: ugrás a hangsávban, filmszakadás következtében hiányzó rész miatt (ennek határait az előtte, ill. utána egészben elhangzó szó határaival jelöltük ki)
- HÁTTÉR: háttérzörejek, zajok

A teljes híradókhöz készült annotáció a hangokat nagyobb blokkokban, az elhangzó szöveg lejegyzése nélkül mutatja, és elsősorban a híradón belüli tájékozódást szolgálja, a megfelelő részlet kiválasztását segíti. Pontos mérések ezek finomításával a konkrét kutatói igények alapján lehetségesek. A 2. ábra szemlélteti mindezt a 404. sz. híradónál. Az 1. filmhírben Ripka Ferenc főpolgármester tart beszédet szűkebb körben, így a beszéd mellett az utcai zajok miatt a HÁTTÉR címkét is megjelenítettük. A TÖMEG címke a közönség éljenzésére utal. Ezt követően látható a 2. filmhír, amely rendőrlovak utcazajhoz szoktatásáról szól. Az elején nem jegyeztük le a főcím alatti zenét, ezért látható egy rövid kihagyás a 2. sávon. A HÁTTÉR címke itt is utcazajt jelez, de ebben a hírben nincs beszéd. Az első hírt külön fájlként (0404-1.wav) is elmentettük és a szövegét lejegyeztük, míg a másodikat nem, mivel abban nem volt beszéd.

A külön elmentett filmhírek beszédanyagát a korábbiakban ismertetett elvek alapján, azaz elsősorban a helyesírási szabályok szerint, de egyedi esetekben a kiejtést követve jegyeztük le. Alapesetben egy annotációs szintet és egy „egyéb” feliratú sávot alkalmaztunk, amelyben a beszéd alatti zenét, háttérzajt jeleztük. Egyszerre beszélés esetén további annotációs sávot adtunk hozzá.

A 3. ábra egy beszédet tartalmazó, külön elmentett, részletesebben annotált fájl mutat. A 406/3. sz. hír arról tudósít, hogy dr. Sipőcz Jenő polgármes-

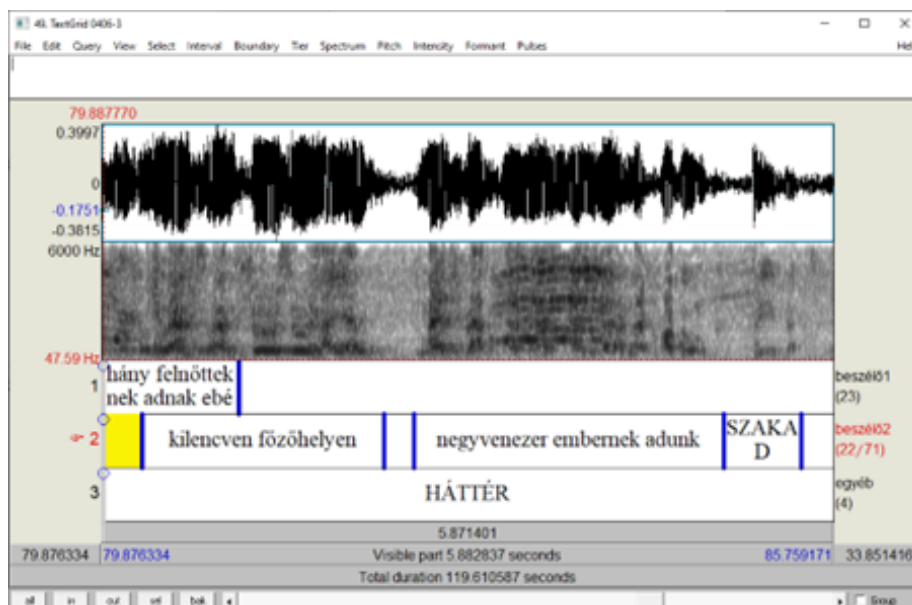


2. ábra. A 404. sz. híradó annotációjának részlete.

ter meglátogatott egy budapesti csecsemőgondozó intézményt, ahol a szociális ellátásról nyilatkozott. Ebben a hírben interjúkészítő is megszólal, és egyszerre beszélés is előfordul, emiatt két beszédsvot használtunk. Háttérzajként végig gyermekzsivaj és egyéb zajok hallatszanak. Megfigyelhető itt egy filmszakadás következtében adódó ugrás. Emiatt az előző szakaszt az utolsó egészben elhangzó szó (*adunk*) határánál lezártuk.

2.2.3. A metaadatok

A fentiek alapján látható, hogy számos új metaadat rögzítése vált szükségesé, így az adattáblát jelentősen át kellett strukturálni. Az adatokat tartalmazó táblázat sorai a rekordok, azaz az egyedi filmhírek. Az első oszlopok az egyes filmhírek azonosítását szolgálják. Rögzítettük a filmhír, a híradó, a filmhír híradón belüli sorszámát, továbbá a hangfájl nevét. Feltüntettük, hogy rendelkezésre áll-e az annotációt tartalmazó textgrid-fájl, illetve a híradó bemutatásának idejét is rögzítettük. Az 1. táblázat az adattábla részletét mutatja. A táblá-



3. ábra. A 406/3. sz. filmhír annotációjának részlete.

zatból megállapítható, hogy a 403. sz. híradó hét filmhírt tartalmazott, ebből kettőben volt magyar nyelvű beszéd.

1. táblázat. A filmhírek azonosítóit tartalmazó adattábla részlete

| sorszám | híradó | hír | hangfájl | annotálva | év | hónap |
|---------|--------|-----|----------|-----------|------|----------|
| 25 | 403 | 1 | 0403 | 1 | 1931 | november |
| 26 | 403 | 2 | 0403-2 | 1 | 1931 | november |
| 27 | 403 | 3 | 0403-3 | 1 | 1931 | november |
| 28 | 403 | 4 | 0403 | 1 | 1931 | november |
| 29 | 403 | 5 | 0403 | 1 | 1931 | november |
| 30 | 403 | 6 | 0403 | 1 | 1931 | november |
| 31 | 403 | 7 | 0403 | 1 | 1931 | november |

A táblázat következő oszlopaiban a beszédesemény jelenlétét, illetve néhány jellemzőjét tüntettük fel. Többféle megoldás kipróbálását követően ezeknek az adatoknak a bináris ábrázolását láttuk legcélravezetőbbnek, ugyanis így világosan átláthatók az egyes tulajdonságok, és ezek kombinációival egyszerűen beál-

líthatók szűrések. A „beszéd” feliratú oszlop azt jelzi, hogy van-e beszédesemény az adott filmhírben. A „csak beszéd” arra utal, hogy háttérzörej vagy aláfestő zene nélküli-e a beszéd. A beszéd alatti aláfestő zene és a háttérzaj jelenlétét külön oszlopokban tüntettük fel. Rögzítettük, hogy volt-e tömeges beszéd (pl. éljenzés), megszólal-e több beszélő (pl. interjúkészítő és interjúalany), előfordul-e egyszerre beszélés, illetve a beszédesemény magyar vagy idegen nyelvű-e. A 2. táblázat illusztrálja mindezt az előzőekben bemutatott rekordokkal (az első néhány oszlop takarásával).

2. táblázat. A filmhírek metaadatait tartalmazó adattábla részlete

| sorsz. | híradó | hír | beszéd | csak beszéd | aláfestő zene | háttérzaj | tömeg | több beszélő | egysz. | magyar | idegen |
|--------|--------|-----|--------|-------------|---------------|-----------|-------|--------------|--------|--------|--------|
| 25 | 403 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 26 | 403 | 2 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 27 | 403 | 3 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 |
| 28 | 403 | 4 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 29 | 403 | 5 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| 30 | 403 | 6 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| 31 | 403 | 7 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |

Megfigyelhető, hogy két filmhír (403/1, 403/3) nem tartalmazott beszédet, csak háttérzajt. A 403/2. és 403/5. sz. filmhíreknél a „csak beszédet” és a „háttérzajt” is megjelöltük. Ennek az az oka, hogy ezekben a hírekben a narrátor hangja tisztán és háttérzajjal is hallható. Három filmhírnél adatoltunk több beszélőt, egy esetben találtunk egyszerre beszélést. Az öt, beszédet tartalmazó filmhír közül kettőben magyar, háromban idegennyelvű szöveget találtunk. A 2. táblázatban terjedelmi okok miatt nem tüntettük fel, de az adattábla tartalmaz még oszlopokat annak megjelenítésére, hogy milyen nyelvű volt a szöveg. Az eddig feldolgozott anyagban német, angol, francia és latin nyelven hangoztak el beszédek.

Rögzítettünk még további kiegészítő információkat. Külön oszlopban megjelenítettük, hogy a filmhír tartalmazott-e a helyszínen játszott zenét (ez nem

azonos az utólag alákevert, aláfestő zenével), a közlés tartalmazott-e paralingvisztikai elemeket (pl. hűmmögést, sóhajtást, csettintést), illetve a film technikai minőségével kapcsolatban két lényeges információt tüntettünk fel, a filmszakadást és a torzulást (3. táblázat, szintén a korábbi oszlopok kitakarásával).

3. táblázat. A filmhírek azonosítóit tartalmazó adattábla részlete

| sorszám | híradó | hír | zene | paralingvisztikai | szakadás | torz |
|---------|--------|-----|------|-------------------|----------|------|
| 25 | 403 | 1 | 1 | 0 | 0 | 0 |
| 26 | 403 | 2 | 0 | 0 | 0 | 0 |
| 27 | 403 | 3 | 0 | 0 | 1 | 0 |
| 28 | 403 | 4 | 1 | 0 | 0 | 0 |
| 29 | 403 | 5 | 1 | 0 | 0 | 1 |
| 30 | 403 | 6 | 0 | 0 | 0 | 1 |
| 31 | 403 | 7 | 1 | 0 | 0 | 1 |

Mindezek mellett továbbra is feltüntettük az adattáblában az első fázisnál már említett módon a beszélő személyét, életkorát, a beszédtypust, a filmhír témáját, illetve a filmhírhez mutató linket.

A jelenleg is folyamatban lévő második fázis során eddig 61 filmhír feldolgozása történt meg, ami az 1931 szeptembere és decembere közötti filmhíradók híranyagát jelenti.

3. Kutatási tapasztalatok

Az eddigiekben egy kutatás készült az első fázisban összegyűjtött adatok, illetve további mérések alapján (Gocsál, 2022). Ebben a kutatásban az elhangzó beszédek artikulációstempó-értékeit vizsgáltuk. Az artikulációs tempó biztonságosan mérhető, mivel a beszédszakaszok hossza és az elhangzó beszédhangok száma könnyen megállapítható. Mivel a protokoll hiánya miatt a beszédanyag strukturálatlan, elvégeztük a felvételek beszédtypusok szerinti, az adatbázisban első megközelítésben rögzítetteknek pontosabb kategorizálását is. Fő kategóriáknak a Wacha (1974) és Krepsz (2016) által közölt típusokat – spontán beszéd,

felolvasás, reprodukzív vagy interpretatív beszéd, félreproduktív beszéd – vettük alapul, majd ezekhez kvalitatív módszerrel, iteratív eljárással olyan alkategóriákat dolgoztunk ki, amelyek a konkrét beszédeseményeket pontosabban jellemzik. Így meghatároztuk, hogy a filmhírekben a legnagyobb számban szónoklatok, szűkebb körben elmondott beszédek, társalgások, narrációk, bejelentések, vezényszavak hangzottak el. Ezekre a beszéd típusokra külön-külön kiszámítottuk az artikulációstempó-értékeket, és a vezényszavak kivételével minden esetben 10–12 hang/s közötti mediánértéket kaptunk. A Kruskal–Wallis-próba nem igazolt szignifikáns különbséget a szűkebb körű beszédek, társalgások, narrációk és bejelentések tempóértékei között ($H(3) = 3,968, p = 0,265$), a szónoklatok beemelésével azonban szignifikáns eltérés mutatkozott ($H(4) = 13,192, p = 0,01$) A vezényszavak esetében valamivel 8 hang/s fölötti medián adódott.

48 beszélő személyazonosságát, és külön adatgyűjtéssel az életkorát is meg tudtuk állapítani, így a mai szakirodalmi adatokhoz hasonlóan kimutatható volt 90 évvel ezelőtti beszélők esetében is az idősebbek lassabb beszéde, azaz az életkor és az artikulációs tempó szignifikáns összefüggése ($r = -0,404, p < 0,01$).

Kvalitatív eljárással – kategóriák felállításával, ezek ciklikus felülvizsgálásával – a filmhíreket téma szerint is csoportosítottuk, így hét hírkategóriát (politikai-közéleti, gazdasági, sport, kiállítás-tudomány-kultúra, technikai újdonságok, hétköznapi életképek, külföldi hírek) különböztettünk meg. Ezek a kvalitatív kutatási eredmények – az adatbázisba felvett elemek bármely szempont szerinti kategorizálása – visszaforgathatók, azaz metaadatként beépíthetők az adatbázisba, így további kutatásokhoz szűrési feltételként alkalmazhatók.

Az adatbázis előfordulási gyakoriságok összesítésére is alkalmas. Meg tudtuk vizsgálni, hogy az áttekintett 15 hónapos időszakban hogyan alakult a szóbeli közléseket tartalmazó filmhírek száma, illetve a leggyakrabban előforduló beszéd típusok, a narrációk és a szónoklatok esetében bizonyos ingadozás mellett lassan emelkedő tendenciát figyeltünk meg. Ez arra utal, hogy a híradók szerkesztői egyre tudatosabban használtak beszédet a hírekben.

Más akusztikai-fonetikai mérést, adatgyűjtést nem végeztünk, azonban feltűnt, hogy a viszonylag gyengébb minőség ellenére is sok esetben jól megjeleníthető a dallamgörbe, mérhetőek az f_0 -értékek, továbbá megállapíthatók a formánsértékek is. Ez további kutatási lehetőségeket nyújt a korabeli beszéd tanulmányozásához. Korlátozó tényező ugyanakkor, hogy a spektrogram csak kb. 6 kHz-ig jelenik meg, így egyes kutatási feladatok elvégzésére – pl. réshangoknál a spektrális eloszlás súlypontjának (Centre of Gravity, CoG) vizsgálata, vö. Gráczy et al. (2020) mai beszélőkre – erősen korlátozottan vagy sehogy nincs lehetőség.

4. További lehetőségek, fejlesztési irányok

Tapasztalataink igazolják Wallis (2014) megállapítását, azaz az annotálás során a fejlesztő nem csak információkat – pl. mondathatárokat – tesz hozzá a szöveghez, hanem ezzel párhuzamosan kritikailag reflektál az annotálás kerekeire, módszereire is. A reflexió során nem csak azt vizsgálja, hogy hasznos-e bizonyos adatokat lejegyezni, vagy helyes volt-e a lejegyzés, hanem azt is, hogy milyen kihívásokat jelent az előre nem várt, újabb tények lejegyzése. Ebből adódóan az adatbázis fejlesztése újabb és újabb felülvizsgálatokból álló ciklikus folyamat, különösen egy ilyen, eleve strukturálatlan, protokoll nélkül, más céllal rögzített beszédanyag esetében. Mindezt a fejlesztés itt ismertetett két fázisa is jól illusztrálja.

Mivel előre megjósolhatatlan, hogy milyen konkrét kutatási feladatokhoz fogják használni az adatbázist, gondolni kell arra, hogy a kutatók saját céljainak megfelelően hogyan fejleszthetik azt tovább, ahogy az észet filmhír-adatbázis esetében Ibrus & Ojamaa (2018) közölték. Előfordulhat, hogy a fonáció típusa lesz az érdekes, ebben az esetben külön annotációs sáv hozható létre (Seppänen et al., 2003), míg más esetekben, például a beszélő dialektusának megjelenítéséhez külön metaadat rögzítése lehet célszerű (Solberg & Ortiz, 2022). Az eddig használt Excel-tábla egyelőre megfelel ez utóbbi célnak, hosszabb távon, nagyobb mennyiségű adat kezeléséhez egy rugalmasan alakítható relációs

adatbázis kiépítése tűnik célszerűbbnek. Jelenlegi terveink szerint az adatokat átemeljük egy Access-adatbázisban definiált táblákba. Ezzel a megoldással a lekérdezések átláthatóbbá válhatnak, miközben a felhasználó saját kategóriákkal, vagy akár további adattáblákkal kiegészítve, tetszőlegesen fejlesztheti tovább az adatbázist. Hátránya ugyanakkor, hogy az Access használata kevésbé elterjedt, mint az Excelé, így tervezzük egy olyan felhasználói dokumentáció elkészítését is, amely az alapoktól kezdődően bemutatja az adatbázis működését és fejlesztési lehetőségeit. Sikeres tesztelést követően a Nyelvtudományi Kutatóközpont honlapján tervezzük az adatbázis közzétételét a szakmai érdeklődők számára, a többi nyelvi adatbázishoz hasonló módon.

Az akusztikai-fonetikai kutatások mellett az adatbázis más nyelvészeti szakterületek kutatását is segítheti. Csak néhány példát említve: a korabeli szóhasználat, a sajátos megszólítás- és említésformák, formális és informális élethelyzetek elemzése szociolingvisztikai, dialektológiai, sőt, nyelvtörténeti szempontból is értékes információkat nyújthat.

Számos külföldi kutató alkalmazott hasonló korpuszokat beszédtechnológiai fejlesztésekhez is. Ezek célja olyan szövegfelismerő rendszerek előállítására, amelyek a legváltozatosabb helyzetekben is robosztus megoldást tudnak nyújtani, például filmek feliratozásához (Álvarez et al., 2022). Koržinek et al. (2019) 1945 és 1962 közötti filmhíradók szövegének automatizált lejegyzéséhez állítottak össze adatbázist. Ebben a gyenge hangminőség mellett kihívást jelentett a számos, ma már nem használt kifejezés is. A Carrive et al. (2021) által ismertett adatbázis fejlesztői pedig francia filmhíradók kutatásához többek között automatizált beszéd- és arcfelismerő szoftvereket alkalmaztak és gazdag metaadat-állományt állítottak elő, amivel az adatbázisuk többféle tudományterületen változatos kutatási feladatok megoldását segítheti.

Nem találtunk erre vonatkozóan említést a szakirodalomban, de lehetőséget látunk kriminalisztikai fonetikai eszközök alkalmazására is, mivel a torzult vagy a lehallgatás során érthetetlen beszédrészekből történeti szempontból adott esetben értékes információk nyerhetők ki. Egy példát említve: a 431. számú, egyetlen hírből álló híradóban 1:57-től azt láthatjuk, hogy az akkoriban min-

denki által jól ismert pilóta, Endresz György tragikusan végződött repülőútja előtt elbúcsúzik a budapesti közönségtől. A hír 2:31–2:35 perc közötti részében érthetetlen beszédfoszlányt hallunk. Ha valamely zaj- vagy torzításcsökkentő eljárás alkalmazásával (Fejes, 2021) egy hangszakértő meg tudja állapítani, hogy mi hangzott el – például a beszélő utalást tett saját egészségi állapotára, fáradtságára, vagy egy elhanyagolhatónak tűnő technikai problémára –, mindez új megvilágításba helyezheti a szerencsétlenség körülményeivel kapcsolatos történettudományi ismereteket. De a kriminalisztikai fonetikai eljárások és a beszédtechnikai fejlesztések tesztelésére is lehetőséget nyújt az adatbázis, mivel a bemondások igen változatos háttérhangok, zörejek, zenék mellett hangoznak el.

Szintén történettudományi jelentősége lehet az orvosi diagnosztikát támogató, beszéd alapú osztályozó eljárások alkalmazásának egyes korabeli közéleti szereplők beszédén (pl. Sztahó et al., 2019). Elképzelhető, hogy ezek segítségével az adott beszélő valamely, mindeddig nem ismert betegsége valószínűsíthetővé válik, ami új támpontot adhat döntéseinek, magatartásának újraértelmezéséhez.

A beszélő személyre és az aláfestő zenére vonatkozó metaadatok használata más jellegű kutatásokat is lehetővé tesz. Lehnert (2019) az 1950-es és 60-as években készült német filmhíradókban alkalmazott narrációs hangokat és aláfestő zenéket vizsgálta. Felfigyelt arra, hogy a II. világháború utáni években a korábbi szigorú, tekintélyt parancsoló, kiabáló hang lassan eltűnt, és helyette kevésbé merev, kevésbé stilizált beszédstílust alkalmaztak, amivel a néző számára hitelesebbé vált a hír. Az NDK-ban két női és három férfi bemondó jelent meg, mindannyiuk hangja jellegzetesen eltérő karakterű volt. A női bemondók nem csak a családdal, gyerekekkel, hanem ipari termeléssel kapcsolatos híreket is felolvastak, ami megfelelt a korabeli ideológiának, ti. nők is dolgoztak ipari munkahelyeken. Az NSZK-ban ugyanakkor csak rövid ideig alkalmaztak női hírolvasókat a filmhíreknél, aminek az oka feltehetően a hagyományosnak tekinthető szerepfelfogás volt.

Lehnert (2019) ugyanebben a munkájában a filmhírek zenei aláfestését is elemezte. Egy 1962-es filmhíradóban a munkások termelésben elért sikereiről

szóló hírben büszkeséget sugalló narrációs hangot és zenei aláfestést használtak, míg ugyanebben a híradóban más híreknél jazzt vagy dél-amerikai zenét alkalmaztak. A zenei aláfestés azonban nem csak a hangulatát adhatja meg a filmhírnek, de sajátos értelmezési kereteket is nyújthat az elhangzó szöveg, illetve képekben közölt információk befogadásához. A szerző talált példát arra, hogy egy, akkoriban jól ismert játékfilm zenéjét vették át a híradó szerkesztői egy filmhírhez, ami tudat alatt sajátos asszociációkra készíthette a nézőket, így rejtett üzenetet közvetíthetett számukra.

Végül még egy példát említve: egy másik kutatásban Gagnepain et al. (2020) televíziós hírekből állítottak össze korpuszt, amelyen egy beszédfelismerő algoritmus, majd szövegbányászati eljárás alkalmazását követően témamodellezést végeztek. Ezzel azt kívánták feltárni, hogy a médiában milyen kollektív sémák, narratívák jelennek meg egy adott, múltbéli témával kapcsolatban, amelyek egy társadalom közös emlékezetét formálják.

A hangos filmhíradók feldolgozása és adatbázisban való rögzítése új lehetőségeket nyit a beszéd, a nyelv, a nyelvhasználat kutatásához. Az itt összegzett eredmények azt is mutatják, hogy egy ilyen adatbázis a beszédeseményeket tágabb kontextusban kezelve szemiotikai, médiatörténeti, történeti, és más irányú kutatások számára is számtalan lehetőséget kínál. Úgy véljük, az itt bemutatott és a későbbiekben továbbfejlesztendő adatbázis is hozzájárulhat ilyen irányú hazai kutatásokhoz.

Köszönetnyilvánítás

A kutatást a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal NKFIH-FK-128814 számú pályázata támogatta.

Hivatkozások

Álvarez, A., Arzelus, H., Torre, I. G., & González-Docasal, A. (2022). Evaluating novel speech transcription architectures on the spanish RTVE2020 database. *Applied Sciences*, 12. URL: <https://doi.org/10.3390/app12041889>.

- Aubergé, V., Audibert, N., & Rilliard, A. (2003). Why and how to control the authentic emotional speech corpora. In *Proc. 8th European Conference on Speech Communication and Technology (Eurospeech 2003)* (pp. 185–188). URL: https://www.isca-speech.org/archive/eurospeech_2003/auberge03_eurospeech.html.
- Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer. version 6.0.52. <http://www.praat.org/>.
- Bóna, J., Imre, A., Markó, A., Váradi, V., & Gósy, M. (2014). GABI – Gyermeknyelvi beszédadatbázis és információtár. *Beszéd kutatás*, 22, 246–251.
- Campbell, N. (2000). Databases of emotional speech. In *Proceedings of ISCA workshop on speech and emotion*. URL: https://www.isca-speech.org/archive_open/archive_papers/speech_emotion/spem_034.pdf.
- Campbell, N., Devillers, L., Douglas-Cowie, E., Aubergé, V., Batliner, A., & Tao, J. (2006). Resources for the processing of affect in interactions. In *Proceedings of the International Conference on Language Resources and Evaluation 2006* (pp. XXV–XXVIII). URL: <http://www.lrec-conf.org/proceedings/lrec2006/>.
- Carrive, J., Beloued, A., Goetschel, P., Heiden, S., Laurent, A., Lisena, P., Mazuet, F., Meignier, S., Pincemin, B., Poels, G., & Troncy, R. (2021). Transdisciplinary analysis of a corpus of French newsreels: The ANTRACT Project. *DHQ: Digital Humanities Quarterly*, 15. URL: <http://digitalhumanities.org/dhq/vol/15/1/000523/000523.html>.
- Cieri, C., Miller, D., & Walker, K. (2004). The Fisher Corpus: A resource for the next generations of speech-to-text. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)* (pp. 69–71). European Language Resources Association (ELRA). URL: <http://www.lrec-conf.org/proceedings/lrec2004/pdf/767.pdf>.

- Devillers, L., & Vidrascu, L. (2007). Real-life emotions detection on human-human medical call center interactions. In C. Manfredi (Ed.), *Models and analysis of vocal emissions for biomedical applications: 5th international workshop. December 13–15, 2007* (pp. 143–146). Firenze: Firenze University Press. URL: https://www.isca-speech.org/archive_v0/maveba_2007/papers/mv07_143.pdf.
- Douglas-Cowie, E., Campbell, N., Cowie, R., & Roach, P. (2003). Emotional speech: Towards a new generation of databases. *Speech Communication*, *40*, 33–60.
- Drude, S., Trilsbeek, P., Sloetjes, H., & Broeder, D. (2014). Best practices in the creation, archiving and dissemination of speech corpora at the language archive. In S. Ruhi, M. Haugh, T. Schmidt, & K. Wörner (Eds.), *Best Practices for Spoken Corpora in Linguistic Research* (pp. 183–207). Newcastle upon Tyne: Cambridge Scholars Publishing. URL: <https://www.mpi.nl/publications/item2051339/best-practices-creation-archiving-and-dissemination-speech-corpora>.
- Fejes, A. (2021). A kriminialisztikai hangtechnikai szakértői tevékenység. In B. Harmati, E. K. Szitkay, A. L. Pap, & B. Papp (Eds.), *Honestas, Humanitas, Humilitas: Ünnepi kötet Kerezszi Klára tiszteletére* (pp. 49–59). Budapest: L'Harmattan Kiadó. URL: http://real.mtak.hu/135138/1/Honestas_beliv_press.pdf.
- Gagnepain, P., Vallée, T., Heiden, S., Decorde, M., Gauvain, J. L., Laurent, A., Klein-Peschanski, C., Viader, F., Peschanski, D., & Eustache, F. (2020). Collective memory shapes the organization of individual memories in the medial prefrontal cortex. *Nature Human Behaviour*, *4*, 189–200. doi:<https://doi.org/10.1038/s41562-019-0779-z>.
- Gocsál, Á. (2022). Beszédtípusok és artikulációs tempóik a korai hangos filmhíradókban. In L. Csárdás, & J. Bóna (Eds.), *Sokszínű beszédtudomány*. Buda-

- pest: Akadémiai Kiadó. URL: <https://mersz.hu/csardas-bona-sokszinu-beszedtudomay>.
- Gósy, M., Gyarmathy, D., Horváth, V., Grácsi, T. E., Beke, A., Neuberger, T., & Nikléczy, P. (2012). BEA: Beszélt Nyelvi Adatbázis. In M. Gósy (Ed.), *Beszéd, adatbázis, kutatások* (pp. 9–24). Budapest: Akadémiai Kiadó.
- Grácsi, T. E., Csapó, T. G., Deme, A., Bartók, M., & Markó, A. (2020). Az /s/ és /ʃ/ hangok akusztikai és artikulációs elkülönítése gyermekeknél: Két longitudinális esettanulmány. In J. Bóna, & V. Krepsz (Eds.), *Nyelvfejlődés csecsemőkortól kamaszkorig* (pp. 45–66). Budapest: ELTE Eötvös Kiadó.
- Grimm, M., Kroschel, K., & Narayanan, S. (2008). The Vera am Mittag German audio-visual emotional speech database. In *2008 IEEE International Conference on Multimedia and Expo* (pp. 865–868). Hannover, Germany: IEEE. doi:<https://doi.org/10.1109/ICME.2008.4607572>.
- Gyarmathy, D., & Neuberger, T. (2015). Egy hiánypótló adatbázis: A Tini BEA. *Beszédkutatás*, 23, 209–222.
- Gyarmathy, D., Neuberger, T., & Grácsi, T. E. (2014). Lejegyzési útmutató a BEA spontánbeszéd-adatbázis háromszintű annotálásához. *Alkalmazott Nyelvtudomány*, 14, 35–41.
- Hunyadi, L., Váradi, T., Szekrényes, I., Kovács, G., Kiss, H., & Takács, K. (2018). Human-human, human-machine communication: On the HuComTech multimodal corpus. In *Selected papers from the CLARIN Annual Conference 2018*. URL: <https://ep.liu.se/ecp/159/006/ecp18159006.pdf>.
- Ibrus, I., & Ojamaa, M. (2018). Newsreels versus newspapers versus metadata: A comparative study of metadata modelling the 1930s in estonia. *VIEW Journal of European Television History and Culture*, 7, 123–137. doi:<https://doi.org/10.25969/mediarep/14758>.

- Koržinek, D., Wołk, K., Brocki, L., & Marasek, K. (2019). Automatic transcription of the polish newsreel. *Poznan Studies in Contemporary Linguistics*, 55, 183–209. doi:<https://doi.org/10.1515/psic1-2019-0008>.
- Krepsz, V. (2016). Fonetikai hasonlóságok és különbözőségek a beszéd típusokban. In J. Bóna (Ed.), *Fonetikai olvasókönyv* (pp. 175–188). Budapest: ELTE Fonetikai Tanszék.
- Lehnert, S. (2019). Music and voice in German newsreels of the 1950s/1960s. In *Proceedings of RE:SOUND 2019 – 8th International Conference on Media Art, Science, and Technology* (pp. 91–98). doi:<http://dx.doi.org/10.14236/ewic/RESOUND19.15>.
- Lleida, E., Ortega, A., Miguel, A., Bazán-Gil, V., Pérez, C., Gómez, M., & de Prada, A. (2020). *RTVE2020 Database Description*. URL: <http://catedrartve.unizar.es/reto2020/RTVE2020DB.pdf>.
- Neuberger, T., Gyarmathy, D., Grácsi, T. E., Horváth, V., Gósy, M., & Beke, A. (2014). Development of a large spontaneous speech database of agglutinative Hungarian language. In P. Sojka, A. Horák, I. Kopeček, & K. Pala (Eds.), *Text, Speech and Dialogue* (pp. 424–431). Cham: Springer International Publishing. doi:https://doi.org/10.1007/978-3-319-10816-2_51.
- Pincemin, B., Heiden, S., & Decorde, M. (2020). Textometry on audiovisual corpora experiments with txm software. In *15th International Conference on Statistical Analysis of Textual Data JADT 2020*. Toulouse, France. URL: http://lexicometrica.univ-paris3.fr/jadt/JADT2020/jadt2020_pdf/PINCEMIN_HEIDEN_DECORDE_JADT2020.pdf.
- Pittermann, J., Pittermann, A., & Minker, W. (2010). *Handling Emotions in Human-Computer Dialogues*. Springer.
- Seifart, F. (2021). Combining documentary linguistics and corpus phonetics to advance corpus-based typology. In G. Haig, S. Schnell, & F. Seifart (Eds.), *Doing corpus-based typology with spoken language data: State of the*

- art* (pp. 115–139). Honolulu: University of Hawai'i Press. URL: <https://scholarspace.manoa.hawaii.edu/handle/10125/74659>.
- Seppänen, T., Toivanen, J., & Väyrynen, E. (2003). Mediateam speech corpus: A first large finnish emotional speech database. In *ICPhS-15* (pp. 2469–2472). URL: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_2469.pdf.
- Simon, E., & Sass, B. (2012). Nyelvtechnológia és kulturális örökség, avagy korpuszpítés ómagyar kódexekből. *Általános Nyelvészeti Tanulmányok*, 24, 243–264.
- Solberg, P. E., & Ortiz, P. (2022). The norwegian parliamentary speech corpus. *ArXiv*, . doi:<https://doi.org/10.48550/ARXIV.2201.10881>.
- Sztahó, D., Imre, V., & Vicsi, K. (2011). Automatic classification of emotions in spontaneous speech. In A. Esposito, A. Vinciarelli, K. Vicsi, C. Pelachaud, & A. Nijolt (Eds.), *Analysis of Verbal and Nonverbal Communication and Enactment. The Processing Issues. Lecture Notes in Computer Science, vol. 6800* (pp. 222–239). Berli, Heidelberg: Springer-Verlag.
- Sztahó, D., Kiss, G., Tulics, M. G., & Vicsi, K. (2019). Betegségek automatikus szétválasztása időben eltolt akusztikai jellemzők korrelációs struktúrája alapján. In *XV. Magyar Számítógépes Nyelvészeti Konferencia* (pp. 209–212). URL: <http://acta.bibl.u-szeged.hu/59086/>.
- Teleki, C., Velkei, S., Tóth, S. L., & Vicsi, K. (2005). Development and evaluation of a hungarian broadcast news database. In F. Augusztinovicz, A. B. Nagy, & Z. Hunyadi (Eds.), *Proceedings of Forum Acusticum 4th European Congress on Acoustics* (pp. 2659–2662). Budapest: OPAKFI.
- Vicsi, K., & Sztahó, D. (2009). Ügyfél érzelmi állapotának detektálása telefonos ügyfélszolgálati dialógusban. In *VI. Magyar Számítógépes Nyelvészeti Konferencia* (pp. 217–225). URL: <http://acta.bibl.u-szeged.hu/58711/>.

Wacha, I. (1974). Az elhangzó beszéd főbb akusztikus stíluskategóriáiról. *Általános Nyelvészeti Tanulmányok*, X, 203–216.

Wallis, S. A. (2014). What might a corpus of parsed spoken data tell us about language? In L. Veselovská, & M. Janebová (Eds.), *Complex Visibles Out There. Proceedings of the Olomouc Linguistics Colloquium 2014: Language Use and Linguistic Structure* (pp. 641–662). Olomuc: Palacký University. URL: <https://corplingstats.wordpress.com/2014/06/24/corpus-language/>.