



Szerkesztette:

MARKÓ  
ALEXANDRA

BESZÉD • KUTATÁS • ALKALMAZÁS

# TANULMÁNYOK A BESZÉDTUDOMÁNY ALKALMAZÁSAINAK KÖRÉBŐL

 ELTE  
EÖTVÖS  
KIADÓ

# TANULMÁNYOK A BESZÉDTUDOMÁNY ALKALMAZÁSAINAK KÖRÉBŐL



# TANULMÁNYOK A BESZÉDTUDOMÁNY ALKALMAZÁSAINAK KÖRÉBŐL

SZERKESZTETTE:  
MARKÓ ALEXANDRA

BUDAPEST, 2021



A kötet megjelenését a Tématerületi Kiválósági Program Támogatta.

Szerkesztette: Markó Alexandra

A szerkesztő munkatársai: Balogh Roland, Juhász Kornélia, Juhász Milán,  
Kaposvári Boglárka, Tóbisz-Hertelendy Réka, Weidl Zsófia

A tanulmányokat a témakörben jártas szakemberek lektorálták.

© Szerzők, 2021

© Szerkesztő, 2021

ISBN 978-963-489-358-5

ISBN 978-963-489-359-2 (pdf)

ISSN 2064-4442



[www.eotvoskiado.hu](http://www.eotvoskiado.hu)



Felelős kiadó: az ELTE Bölcsészettudományi Kar dékánja

Projektvezető: Urbán László

Kiadói szerkesztő: Tihanyi Katalin

Tördelés: Farkas Milán

Borítóterv: Csele Kmotrik Ildikó

Nyomdai kivitelezés: Multiszolg Bt.

# Tartalom

(Sorozat)szerkesztői előszó .....	7
A hallás, a beszédészlelés és a beszédmegértés vizsgálata .....	9
<i>Gósy Mária</i>	
A diszfónia és automatikus felismerése .....	35
<i>Tulics Miklós Gábriel</i>	
A depresszió automatikus becslése a beszéd akusztikai-fonetikai jellemzői alapján .....	65
<i>Kiss Gábor</i>	
A nyelv és a beszéd sajátosságai neurodegeneratív kórképekben .....	87
<i>Hoffmann Ildikó – Svindt Veronika – Bóna Judit</i>	
Az orofaciális miofunkcionális diszfunkció és a nyelvlökéses nyelés .....	115
<i>Havadi-Nagy Marian – Deme Andrea</i>	
Idegennyelv-elsajátítás és akcentus .....	147
<i>Juhász Kornélia</i>	
Hogyan beszélünk a gyerekekhez? .....	175
<i>Kohári Anna</i>	
Egyéni variabilitás a beszédben .....	201
<i>Grácsi Tekla Etelka – Krepsz Valéria – Huszár Anna</i>	
Beszélőazonosítás a kriminalisztikában .....	229
<i>Tatár Zoltán – Varga Zoltán – Főző Eszter</i>	
Fonetika és társalgáskutatás: a beszélőváltások kérdésköre .....	251
<i>Horváth Viktória – Bóna Judit – Dér Csilla Ilona – Gyarmathy Dorottya – Hámori Ágnes – Huszár Anna – Krepsz Valéria – Weidl Zsófia</i>	



## (Sorozat)szerkesztői előszó

A *Beszéd – Kutatás – Alkalmazás* című tudományos könyvsorozat tizenegyedik kötete ismét különféle témájú tanulmányokat összefogó kötet, melynek a megszerkesztésére a sorozatszerkesztő vállalkozott. E gyűjtemény célja az, hogy az oktatásban alkalmazható friss, aktuális áttekintést adjon a beszéd tudománnyal kapcsolatos alkalmazásorientált kutatásokról, ezek módszertanáról, olykor módszertani nehézségeiről, eredményeiről és az alkalmazások köréről. Ennélfogva a jelen áttekintésbe olyan tudományterületek is bekerültek, vagy kaptak nagyobb súlyt, amelyek korábban csak egy tanulmány erejéig szerepeltek a sorozat első tíz kötetében. Éppen annak érdekében, hogy a friss tudományos eredmények is reprezentációhoz juthassanak, a tanulmányok túlnyomó részéből most is folyó tudományos projektekbe, zajló vagy éppen csak lezárult (doktori) kutatásokba kap betekintést az olvasó. Az eredmények egy része itt olvasható először magyar nyelven.

A hallás, a beszédészlelés és a beszédmegértés működésével, eltéréseivel és vizsgálati módszereivel ismert meg Gósy Mária tanulmánya. A humán feldolgozásból a gépi azonosításra tér át az ezt követő két tanulmány, mindkettő a beszéd mint biomarker szerepét vizsgálja különféle beszédben is azonosítható eltérések kapcsán. Tulics Miklós Gábor kutatása a beszéd egyik leggyakoribb elváltozását, a diszfónia megjelenését és automatikus azonosításának módszereit mutatja be. Kiss Gábor egy, a beszédben is nyomot hagyó állapot, a depresszió jelenlétének és a depresszió súlyosságának automatikus becsüléséről ad számot írásában. Ezek mellett neurodegeneratív kórképeknek is lehet hatásuk a beszédre. A leggyakoribb ilyen betegségek nyomán kialakuló nyelvi és beszédbeli sajátosságokat Hoffmann Ildikó, Svindt Veronika és Bóna Judit tanulmánya mutatja be. Talán nem is gondolnánk rá, de akár még a nyelésünk működési formájának is lehet hatása a beszédünkre, különösen a beszédhangok képzésére. Ilyen összefüggések is feltárnak Havadi-Nagy Marian és Deme Andrea tanulmányából, amely az orofaciális miofunkcionális diszfunkció jelenségét járja körül. Mindezek a hatások és még végtelenül sok tényező járul hozzá ahhoz, hogy a beszédünk egyedi és egyéni. Ilyen faktor lehet az is, hogy nemcsak az anyanyelvünkön szólalunk meg, hanem idegen nyelveket is beszélünk. Mindannyiunk hétköznapi tapasztalata, hogy sok esetben megállapítható, hogy a beszélő az általa éppen használt nyelvet második (vagy többedik) nyelveként tanulta. Arról, hogy milyen okok állnak ennek a tapasztalatnak a háttérében, Juhász Kornélia összefoglaló tanulmánya tájékoztat. Máshogyan beszélünk akkor is, amikor kisgyerekekhez szólunk. Kohári Anna azt mutatja be, hogy milyen jellemzőkben tér el a gyermekekhez és a felnőttekhez szóló beszéd, és mi minden állhat mindennek a háttérében a különböző elméletek szerint. A beszéd egyéni mintázataihoz természetesen sok további tény hozzájárul, ilyen például az életkor. Elsősorban az életkorral és a nemmel kapcsolatos változatosságot tárja elénk Grácsi Tekla Etelka, Krepsz Valéria és Huszár Anna tanulmánya, akik egy adatbázis folyamatban lévő építése kapcsán a beszéd egyéni változásait is



vizsgálják. Ez a témakör pedig elvezet bennünket ahhoz a kérdéshez, hogy mi alapján és hogyan azonosítható a beszélő. Ezt a kérdést a kriminalisztikai alkalmazások kapcsán több szempontból is bemutatja Tatár Zoltán, Varga Zoltán és Főző Eszter írása. Végül a leghétköznapibb beszédhelyzet, a társalgás vizsgálatának pragmafonetikai dimenzióit tárja elénk a beszélőváltások sokrétű elemzése kapcsán Horváth Viktória, Bóna Judit, Dér Csilla Ilona, Gyarmathy Dorottya, Hámori Ágnes, Huszár Anna, Krepsz Valéria és Weidl Zsófia.

A kötetnek nem titkolt célja, hogy a felsőoktatás minden szintjén segítséget nyújtson a korszerű ismeretek közléséhez, átadásához és befogadásához. A tanulmányok szerzői úgy állították össze és fogalmazták meg az itt szereplő szövegeket, hogy azok az egyetemi oktatásban akár kötelező vagy ajánlott olvasmányként is szerepelhessenek, a hallgatók önállóan fel tudják őket dolgozni, és a tanultakról számot tudjanak adni akár referátum, akár dolgozat vagy vizsgakeretében. A tanulmányok éppen ezért az átlagos és általános műveltségi szintből kiindulva fogalmazzák meg az adott témakör tudományos hátterét, mutatják be az alkalmazás használati körét, célcsoportját. Kitérnek az esetleges módszertani kihívásokra, és ha a terület jellege ezt megkívánja, megjelölik a további kutatás kérdéseit, irányait. Némelyik tanulmány olyan témákat és kérdéseket is felvet, amelyekről a hallgatók önállóan gondolkodhatnak, vagy azt csoportban megvitathatják, esetleg projektmunkában is kidolgozhatók.

A tanulási célok szempontjából releváns elméleti és gyakorlati vonatkozású fogalmak esetében a szerzők ügyeltek arra, hogy ezeket világosan körülhatárolják, tanulható, definíciószerű állításokat fogalmazva meg. Ezeket a kulcsszószerű elemeket *Fogalmak* cím alatt gyűjtöttük össze az egyes tanulmányok végén. Az online kiadás előnye, hogy e fogalmak a szövegben automatikusan kereshetők, mind a definíció, mind annak kontextusa segítheti a feldolgozást.

A felvonultatott kutatások sokszínűek, és korántsem merítik ki azoknak az alkalmazásoknak a körét, amelyek a beszédtudományhoz kapcsolódóan a mindennapi életünkben is felbukkannak. Éppen ezért tervezzük a gyűjtemény folytatását egy újabb kötetben, újabb területeket bemutató írásokkal.

A *Beszéd – Kutatás – Alkalmazás* című sorozatot 2013-ban indította útjára az ELTE Eötvös Kiadó. Kifejezetten online és ingyenesen elérhető kötetek megjelenését terveztük, így a sorozat a kezdettől az ELTE READER tartalmi között szerepel, miközben a kötetek – a hagyományos könyvformátumot jobban kedvelő olvasók kedvéért – nyomtatott formában is elérhetők. Sokat köszönhetünk a mindenkori támogatóinknak, hiszen a megjelentetés költségeinek átvállalásával lehetővé tették, hogy ezek a friss tudományos tartalmak az olvasók számára ingyenesen elérhetők legyenek. A jelen kötet létrejöttét a Tématerületi Kiválósági Program támogatta.

*Markó Alexandra*  
a sorozat és a kötet szerkesztője

# A hallás, a beszédészlelés és a beszédmegértés vizsgálata

Gósy Mária

NYELVTUDOMÁNYI KUTATÓKÖZPONT  
ÉS ELTE EÖTVÖS LORÁND TUDOMÁNYEGYETEM

## 1. Bevezetés

A szóbeli kommunikáció körfolyamat, a beszélő személy közli a szándéka szerinti gondolatait (ez a beszédprodukciónak), a hallgató személy pedig feldolgozza azokat (ez a beszédfeldolgozás<sup>1</sup>). Majd a szerepek felcserélődnek, a hallgató lesz a beszélő, aki válaszol, reagál, az iménti beszélő pedig hallgatóként megérti az elhangzott közléseket. A mindennapokban feltételezzük, hogy ez a kommunikációs folyamat megfelelően működik, a hallgató pontosan érti és értelmezi az elhangzottakat. Felnőttek esetében ez rendszerint így is van; ha mégsem, akkor a hallgató jelzi, jelezheti a problémát, visszakérdez, rákérdez, ismétlést vagy megerősítést kér stb. Gyermekkorban ez ritkábban következik be, annál is inkább, mivel számos esetben a gyermek nincs is tudatában annak, hogy téves volt a beszédmegértése (vagy annak egy része), illetve hogy valamelyik folyamat, esetleg több is, hibásan működött.

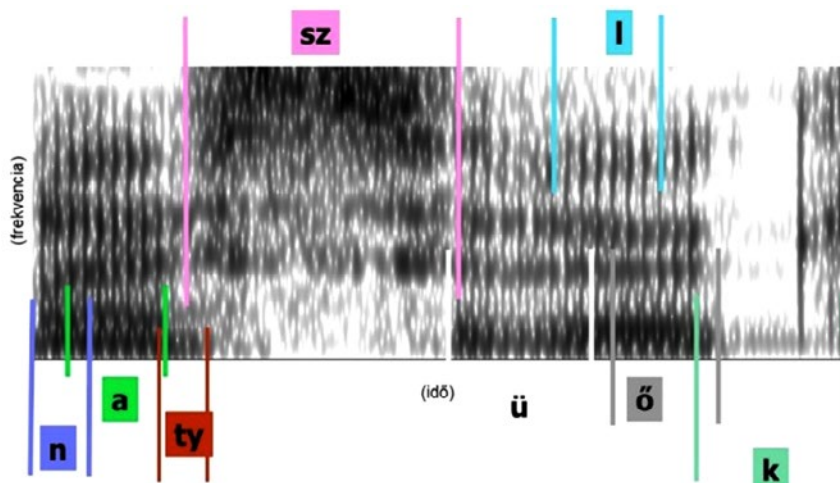
A beszédprodukciónak tervezési folyamatai a beszéd szándékától a kiejtésig rejtetten működnek, a létrejött beszéd azonban hallható; megítélhető és minősíthető az, hogy a beszélő mit és hogyan mondott. Az elhangzott beszéd számos szempontból elemezhető. Ezzel szemben a beszédfeldolgozás teljes egészében rejtetten működik, csak bizonyos reakciók alapján lehetnek bizonytalan ismereteink arról, hogy a hallgató mit észlelt, hogyan értette és miként értelmezte az elhangzott közlést. A mechanizmus vizsgálata csak közvetve lehetséges (Gósy 2005). A beszédprodukciónak különféle artikulációs és akusztikai jellemzőit hosszú ideje kutatják, a beszédfeldolgozás mechanizmusát csak néhány évtizede. Ennek ellenére nagyon gazdag ismeretanyag gyűlt össze, amely magyarázatot ad arra, hogy miként vagyunk képesek az elhangzott beszéd megértésére átlagos és nehezített (pl. zajos) kommunikációs körülmények között is.

---

<sup>1</sup> A beszédfeldolgozás mechanizmusának megnevezésére többféle terminus is ismeretes. A *beszédfeldolgozás* és a *dekódolás* a mechanizmus egészére utalnak. A *beszédészlelés* (illetve a *beszédpercepció*), a *beszédmegértés* és az *értelmezés* általában egyes folyamatokra hivatkoznak. A *beszédmegértés* terminust szokták a teljes folyamat megnevezésére is használni.

## 2. A beszédfeldolgozás folyamatai

A beszédfeldolgozás során a hallgató az artikuláció következtében létrejövő, folyamatos akusztikai jelből diszkrét nyelvi egységeket azonosít (beszédhangokat, szótagokat, szavakat, mondatokat). A kiinduló kérdés az, hogy ebből az akusztikai jelsorozatból hogyan dekódolja a hallgató az abban kódolt üzenetet (pl. PISONI–REMEZ 2005; HEALD–NUSBAUM 2014). Az 1. ábra egyetlen szó példájával illusztrálja a beszédészlelés alapproblémáját. Az ábra hangszínekének vízszintes tengelye az időt, függőleges tengelye a frekvenciát szemlélteti (a feketedés mértéke az intenzitásváltozásokra utal). A rárajzolt vonalak az egyes beszédhangok akusztikai lenyomatának határait illusztrálják, amelyek alapján kiolvashatók egyúttal a koartikulációs megvalósulások is (együttes ejtések, egymásra hatások). Látható például, hogy a szókezdő nazális mássalhangzó ejtése még nem fejeződött be, de a követő magánhangzó artikulációja már megkezdődött.



1. ábra

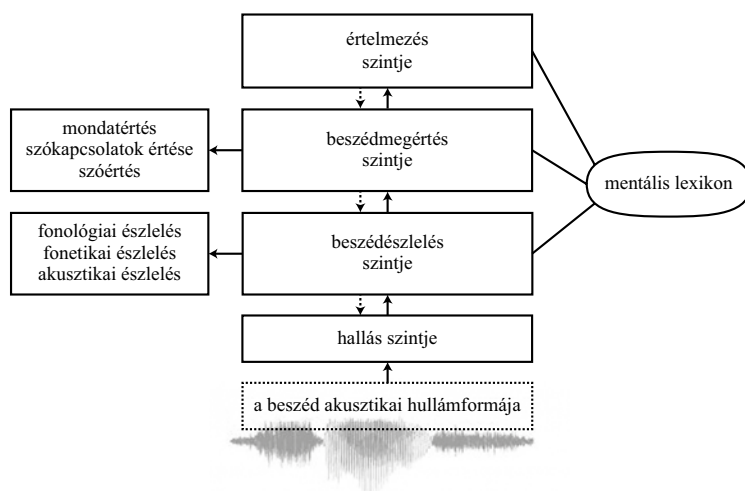
*A nagyszülők (ejtés: natyszülők) szó kiejtése alapján készült hangszínekép és a betűkkel jelölt beszédhangok határai a koartikulációs sajátosságokkal*

A kutatók meghatározták a beszédfeldolgozás univerzális modelljét, amely a nyelvspecifikus sajátosságokkal kiegészülve teszi lehetővé, hogy megértsük, miként nyerjük ki a nyelvi (szemantikai, fonológiai, grammatikai stb.) információkat a beszédakusztikai jelből (Gósy 2005; Csépe 2007; Pléh 2013). A 2. ábra a beszédfeldolgozásnak a legtöbbször által elfogadott, **hierarchikus** felépítésű **modelljét** szemlélteti. Ez az univerzális modell nyelvfüggetlen. Attól válik nyelvspecifikussá, hogy beépülnek az adott nyelvre jellemző sajátosságok az egyes szinteken zajló folyamatok működésébe. A mechanizmus felső(bb) szintjeinek szemantikai és szintaktikai elemzései visszahat(hat)nak az alsóbb szintek műveleteire a feltételezett interaktív

kapcsolatok mentén (ezt szemléltetik a lefelé mutató szaggatott nyilak). Ez az interaktív kapcsolat segíti az átmeneti dekódolási nehézségek azonnali megoldását.

A beszédfeldolgozás bemenete, azaz első szintje a **hallás**. Meghatározása szerint környezetünk hangjelenségeinek a feldolgozását jelenti, beleértve a beszéd azonosítását is, de nem jelenti a beszéd észlelését vagy megértését (csak felismerjük, hogy beszédet hallottunk, és nem más hangjelenséget vagy zajt). A hallással kapcsolatban két fontos dolog hangsúlyozandó. (i) A hallás épsége nélkül a ráépülő beszédfeldolgozási folyamatok csak részlegesen működnek (vö. nagyot-hallók beszéde). (ii) Az ép hallás még nem jelenti azt, hogy a ráépülő beszédészlelési, beszédmegértési és értelmezési folyamatok megfelelően funkcionálnak (a hallás ugyanis nem azonos a ráépülő folyamatokkal).

A beszédészlelés folyamatában három alapszintet különítünk el, az akusztikait, a fonetikait és a fonológiai (Gósy 1989, 2005). Az első szinten, az **akusztikai észlelés** során történik a beszéd akusztikai jellemzőinek felismerése (pl. az időviszonyok, a kváziperiodikus vagy a zörejszerű összetevők, a frekvenciasávok, intenzitásviszonyok azonosítása). A **fonetikai észlelés** szintjén az időben változó frekvencia- és intenzitásviszonyokat a hallgatónak meg kell feleltetnie az adott nyelv beszédhangjainak (pl. magánhangzó vagy mássalhangzó, zöngés zárhang vagy zöngétlen zárhang hangzott-e el). Az észlelt jelenségek aktiválják a neurális reprezentációkat. Nézzünk egy példát! Az elsődleges elemzéssel meghatározzuk, hogy az észlelt hang aperiodikus (nem szabályosan ismétlődő) összetevőkből állt, zörejeket tartalmazott, jellemzően 2000 Hz és 3000 Hz között, majd a következő szinten ezen elemzési eredmény alapján azonosítjuk a zöngétlen, posztalveoláris *s* réshangot. A **fonológiai észlelés** szintjén a felismert beszédhangok kijelölik az adott fonémát, itt történik a koartikulációs folyamatok megfejtése (pl. a *szabaccság* ejtés fonémái: *szabadság*), valamint a nyelvre jellemző egyéb fonológiai folyamatok felismerése.



2. ábra

A beszédfeldolgozás univerzális, hierarchikus modellje a hallástól az elhangzott beszéd értelmezéséig

A beszédészlelési folyamatok működtetése során a hallgató szegmentálja (elhatárolja) az egymást követő beszédhangokat, felismeri az adott nyelvre jellemző, perцепциót érintő szabályszerűségeket. Hangsúlyozzuk, hogy a jelentésnek azonban ebben még nincs szerepe. Ez azt jelenti, hogy amikor a beszédészlelés során felismerjük az elhangzott  $t + i + g + r + i + s$  hangsort, akkor az a folyamat működésének eredménye. Ha az észlelt hangsor alapján rájövünk, hogy ez voltaképpen a *tigris* nevű 'csíkos nyamacska', az már szóértés.

A beszédészlelés alapfolyamataihoz továbbiak is csatlakoznak: a szeriális, a transzformációs, a vizuális, a ritmusészlelés és a beszédhang-megkülönböztetés. A **szeriális észlelés** (sorozatészlelés) a beszédhangok egymásutániségének felismerését, illetve a szavak megfelelő sorrendjének az azonosítását biztosítja. A **transzformációs észlelés** teszi lehetővé, hogy az azonos, illetve hasonló jellemzőket általánosítani tudjuk, ami alapot ad a beszédhangok és (az alfabetikus ábécékben) a betűk összefüggésének felismerésére. A **vizuális észlelés** elsősorban a látható artikuláció (pl. ajakműködés) beépítését jelenti a beszédhangok felismerésébe (nem csak a hallássérültek esetében). A **beszédrítmus perцепciójának** szoros kapcsolata az időviszonyokkal. A **beszédhang-differenciálás** révén különböztetjük meg az eltérő hangminőségeket (pl. a zöngés és a zöngétlen mássalhangzókat).

A **beszédmegértés** is több folyamatból tevődik össze. Egyrészt felismerjük a jelentést, másrészt a grammatikai szerkezetet, azaz a morfológiai, szintaktikai struktúrákat, és összekapcsoljuk a különböző elemzési eredményeket. Mindennek révén értjük meg a szavakat, a különféle morfémaik eredményezte különbségeket, a szavak kapcsolatait, a több szóból álló szerkezeteket. Például megértjük az elhangzott *mókus* szót, de ha azt halljuk, hogy *mókusok*, az már más szó, más a morfológiai struktúrája és a jelentése, más a feldolgozási folyamata. A *mókusokat* szó esetében pedig már három morfémat kell azonosítanunk. Kérdés, hogy a gazdag morfológiájú szavak esetén vagy a mondatfelismerésben a jelentés vagy a szerkezet azonosítása az elsődleges. Az előbbi példát ismételve, ha elhangzik az, hogy *mókusok*, akkor vajon a 'mókus' mint jelentés vagy a többes szám ténye lesz az elsődleges elemzési eredmény? A kutatások szerint a szerkezet az elsődleges, a jelentés a másodlagos, bár a felismerésük között eltelt idő a másodperc töredéke (GÓSY 2005; MARTIN et al. 2017).

A hosszabb szövegek feldolgozásához bonyolultabb műveletek integrálására van szükség az elhangzottak értelmezésének szintjén (pl. PLÉH 2013). Itt történik a megértett közlés(ek) összekapcsolása az emlékezetben már korábban tárolt ismeretekkel és/vagy tapasztalatokkal. Mindehhez megfelelő szókincs, életkor-specifikus lexikális hozzáférés, szövegtagolás, logika is szükséges (pl. CSÉPE 2007; PLÉH 2013; LIBÁRDI 2017; HAGOORT–BECKMAN 2019). A feldolgozási folyamatokban érvényesülnie kell az együttműködési elvnek, különféle kognitív működéseknek (pl. a szövegreprezentáció kialakulásának), a figyelemnek és az adott életkorban elvárt háttértudásnak (világismeretnek) is. (A szövegértés rendszerint komplexebb feldolgozást igényel, mint egy rövid, mondatszerű közlés megértése, értelmezése; olykor azonban egyetlen rövid közlés is nagyon bonyolult feldolgozási műveletek sorát teszi szükségessé.)

Említettük a **mentális lexikont**, ami voltaképpen egy tárolórendszer, amelyben a nyelvre, illetőleg a nyelvhasználatra vonatkozó egységek és szabályok találhatók, beleértve például

a pragmatikai jellemzőket is (GÓSY 2005; LIBÁRDI 2017; MARTIN et al. 2017; HAGOORT 2019). Tartalmazza a nyelv fonémáit, beszédhangkészletét, a mentális reprezentációkat, neurális spektrogramokat, szavakat, szókapcsolatokat, összekapcsoló elemeket, különféle fonológiai, grammatikai és egyéb szabályokat. A mentális lexikont a beszédészlelés szintjétől kezdve aktiváljuk, a lexikális hozzáférés (a szükséges nyelvi egységek elérése és aktiválása) a dekódolásban folyamatos.

### 3. A beszédfeldolgozás fejlődése

A beszédfeldolgozás fejlődéséhez nélkülözhetetlen az ép hallás, ami már a megszületéstől funkcionál, de a mielinizáció (ez azt jelenti, hogy az axonoknak, a sejt kimenő részeinek a felülete egy viaszos jellegű anyaggal, a mielinnel vonódik be, amely szigetelést és gyors üzenetközvetítést biztosít) még évekig folytatódik a hallási feldolgozás magasabb szintjein (BOOTHROYD 1997; KEMALOĞLU et al. 2005). A csecsemő a kezdetektől figyel a beszédre, amire erőteljesebb mimikával és mozgással válaszol. A harmadik hónaptól már keresi a hang forrását, később a fajtát a hangforrás irányába fordítja. Az öt hónapos csecsemő felismeri a legközelebbi családtagok hangját (FENT 2007). A folyamatos gögicselés jelzi azt, hogy a beszéd elsajátítása megindult, az első szavak, majd a szavak összekapcsolása a második életévben az ép hallás és a jó beszédészlelés és beszédmegértés jelei. Az óvodáskorú gyermekek hallása sok tekintetben megközelíti a felnőttekét, a gyakorlottság és a hallási figyelem olykor korlátozó tényezők; a beszédfeldolgozásuk mintegy hároméves koruktól tesztelhető.

A gyermek az anyanyelv-elsajátítás során fokozatosan válik képessé arra, hogy elkülönítse és felismerje a beszédhangokat, megfeleltesse őket a fonémáknak, szegmentálja a nyelvi jeleket, működtesse a lexikális hozzáférés folyamatát, azonosítsa a jelentés és a grammatikai forma összefüggéseit (GÓSY 2005; CSÉPE 2007; BAVIN–NAIGLES 2015). A fejlődés során a tapasztalt információk sorozata tárolódik az idegrendszerben, és ezeket az információkat használja fel a gyermek a beszédpercepció fejlődés különböző szakaszaiban. A beszédészlelési folyamatok egy része megelőzi a beszédmegértést, a fejlődés azonban gyakorlatilag párhuzamosan megy végbe. A beszédpercepció, illetve a beszédmegértés és az értelmezés kölcsönösen feltételezik egymást.

A fokozatosság és az ugrásszerű fejlődés egyaránt jellemző lehet. A mássalhangzók megkülönböztetése például 4 éves korra általában jól működik, az időtartamuk pontos azonosítása azonban csak 6-7 éves korra várható el. A szókincs bővülése az első három évben relatíve gyors, majd életkoronként eltérő ütemű növekedést mutat. Ugyanez tapasztalható a beszédprodukción is: a beszédhangok artikulációs gesztusainak elsajátítása például relatíve fokozatos, a határozóragok használata ugyanakkor meglehetősen rövid időszakokban (néhány hét, néhány hónap alatt) következik be (BUNTA et al. 2016). Mindez nem független a beszédészlelés és

a beszédmegértés fejlődésétől, annak egyfajta tükréként is értelmezhető. Tipikus fejlődés esetén a gyermek jól dekódolja az elhangzott közléseket, beszédmegértése és értelmezése (a kapcsolódó egyéb folyamatokkal együtt) életkora kognitív szintjének, illetve világismeretének megfelelően működik (Gósy 2005).

Mi az oka annak, hogy az utóbbi évtizedekben a beszédfeldolgozás vizsgálatának szükségessége fokozottan előtérbe került? A tudományos érdeklődésen túl, a gyakorlat számos ténye hívta fel a figyelmet arra, hogy a gyermekek beszédészlelésének és beszédmegértésének az ellenőrzése nem megkerülhető. Növekvő tendenciát mutatnak az anyanyelv-elsajátítási problémák (beleértve a megkésett beszédfejlődést), az írott anyanyelv (olvasás, írás, helyesírás) megtanulásának nehézségei, ez utóbbit a nemzetközi olvasásfelmérések sokszor szerény teljesítményei is alátámasztják. A tanulási problémák egyre többfélék, és a háttérükben számos, egymásnak sokszor ellentmondó okot próbálnak a szakemberek meghatározni (KORHONEN et al. 2014). Fontos annak megismerése például, hogy a környezeti nyelvi ingerek bizonyos változásának milyen a hatása a beszédfeldolgozásra (és tágabban a nyelvfeldolgozásra), ilyen például a vizuális ingerek túlsúlya a verbális ingerekkel szemben, avagy az elektronikus eszközök használata (mobiltelefon, tablet) egyre fiatalabb életkortól. Növekszik a kétnyelvű gyermekek és az ikrek száma, akiknek a nyelvsajátítása, illetve a dekódolási folyamataik vizsgálata különösen fontos.

Számos teszt áll rendelkezésre a gyermekek nyelvi teljesítményének vizsgálatára, főként a beszédprodukciónak (pl. JUHÁSZ 1999). A beszédészlelés és a beszédmegértés teljesítményének megítélése a rejtetten működő folyamatok miatt egyszerű megfigyeléssel lehetetlen, a tapasztalások sokszor félrevezethetnek, nehéz elkülöníteni a hallásproblémát a beszédfeldolgozás zavarától. Ez a mechanizmus csak célzott és átfogó teszteléssel ismerhető meg. A felszínen ugyanakkor számos jelenség utalhat arra, hogy a beszédpercepció elmaradottan működik, esetleg egyes folyamatokban zavar van. Ha a gyermektől gyakran tapasztalható reakcióhiány vagy téves reakció, jellemzően visszakérdez az elhangzottakkal kapcsolatban, kerüli a mesehallgatást, figyelmetlennek vagy életkorához képest túlzottan játékosnak tűnik, viselkedési (magatartási) zavar jellemzi, a szókincese feltűnő hiányosságokat mutat, akkor ezek jelezhetik, hogy a háttérben a beszédfeldolgozás nehézségei húzódnak meg. Utalhat erre még a domináns kéz használatának késői kialakulása, a megkésett, illetve lassú beszédfejlődés, a beszédprodukciónak hiba, az olvasás, írás, helyesírás, a memoriterek, az idegen nyelv tanulásának nehezítettsége, a házi feladatok rendszeres „elfelejtése”, általában a lassú feladatmegoldás. Ezek a „tünetek” még tovább is folytathatók. Hangsúlyozzuk, hogy egyik sem utal biztosan a beszédfeldolgozás problémájára (hiszen az egyes jelenségek más okokra is visszavezethetők), ezért fontos a célzott tesztelés; annak mielőbbi eldöntése, hogy a gyermek beszédészlelése, beszédmegértése, az értelmezési folyamatok és a lexikális hozzáférés életkor-specifikusan működnek-e.

## 4. A beszédfeldolgozás zavarai

Minél komplexebb egy mechanizmus, annál több helyen állhat elő benne valamilyen zavar. A hallás zavara lehet veleszületett probléma, de kialakulhat bármely életkorban. A **nagyot-hallás** lehet **vezetési** típusú, ekkor a hallójáratban, a dobhártyán vagy a hallócsontokon, illetve a dobüregben van kóros elváltozás. Az ilyen hallászavart mutató gyermek kissé torzítva és halkabban hallja a beszédet. Az **idegi eredetű** vagy percepciós típusú **halláscsökkenést** a belsőfül elváltozásai idézik elő, az elhangzó beszéd jellegzetesen halkabb lesz a hallássérült számára, és esetükben hiányoznak a beszédhangok magas frekvenciás összetevői. Károsodhatnak az agyi pályák és központok, valamint a hallóideg is. A **kevert típusú halláscsökkenés** esetén az előzőekben ismertetettek együttesen jelentkeznek (PYTEL 1996; FENT 2007).

A hallás károsodása értelemszerűen nem teszi lehetővé a beszédészlelés és a beszédmegértés megfelelő működését. Az elvárt folyamatműködés, valamint a visszacsatolás hiánya korlátozza a beszédfeldolgozást a gyermekeknél, ami különösen kockázatos éppen a nyelvcsajátítás időszakában. A nagyothalló gyermekek beszédészlelése rendszerint nagyobb mértékű zavart, illetőleg elmaradást mutat, mint a beszédmegértésük (bár óriásiak az egyéni különbségek). A beszédfeldolgozás fejlődését nemcsak a fennálló (akár korrigált, pl. hallókészülék) hallászavarok, hanem az átmenetiek is negatívan befolyásolják. Ezek leggyakrabban felső légúti hurutos megbetegedések, valamint a megnagyobbodott orrgaratmandula következtében alakulnak ki (BALÁZS–GÓSY 1994). Ezek a különböző mértékű (vezetési típusú) halláskárosodások rendeződnek ugyan, de a fennállásuk alatt a nyelvi fejlődés különböző mértékű megakadása, illetve stagnálása következik be, és következményként kialakulhatnak a beszédfeldolgozással kapcsolatos elmaradások.

A beszédészlelés és a beszédmegértés bármelyik folyamata eltérhet az életkornak megfelelő működéstől ép hallás és életkornak megfelelő értelmi képesség esetén is, csak ezt sokkal nehezebb felismerni, mint például egy beszédhibát. A beszédpercepciós elmaradás, illetve zavar okai három kategóriába sorolhatók: organikus eredet (neurológiai kórképek, agyi működés, beszédszervek, hallás), funkcionális eredet (idegrendszeri, feldolgozási műveletek) és szocio-kulturális eredet (ingerszegény környezet, kódváltó környezet, egyéb nemkívánatos környezeti tényezők).

A környezeti tényezők közül érdemes kiemelni a különféle zajokat, a beszédtempó növekedését, a felnőtt beszédminta megakadásjelenségeinek gyakoriságát és a verbális kommunikáció relatív ritkaságát. A zajszennyezés környezeti ártalom, ami negatív hatással lehet a hallásra és a beszédfeldolgozásra (együttesen és külön-külön is). A gyermekek sokszor vannak zajos környezetben, mind a családban, mind az óvodában, iskolában, a közlekedés során; a szabadidős tevékenységek egy része szintén erős zajjal jár<sup>2</sup>. A zajos környezetben másképpen működnek

<sup>2</sup> Vö. *Gyermekek zajterhelése, hallásvédelmé* c. konferencia (Budapest, 2016. november 10.) előadásaival (<http://www.hermanottointezet.hu/node/800209>).



a beszédfeldolgozási folyamatok, különösen gyermekek esetében, akik a kevesebb beszédrutin miatt alig-alig tudnak kompenzálni. Ennek következtében az agyi reprezentációk kialakulása késhet, szelektív, illetve bizonytalan lehet, ami visszahat a mindennapi kommunikációra. Az emberek beszédtempója az utóbbi évtizedekben felgyorsult (bár természetesen vannak kivételek), ami azt jelenti, hogy a 21. századi gyermekeknek általában rövidebb idő alatt kell működtetniük a beszédpercepciósi folyamataikat, mint elődeiknek. FÓNAGY és MAGDICS mérései szerint (1960) az átlagos, köznyelvi beszédtempó 11,35 beszédhang/s, ez ma akár 2-3 beszédhanggal is gyorsabb lehet másodpercenként (ami mintegy 20 szónyi különbséget jelent percenként). A relatíve gyors beszédtempó miatt a hallásalapú feldolgozás hibás belső reprezentációk kialakulásához, a beszédhangok téves azonosításához, valamint bizonytalanabb beszédmegértéshez vezethet a gyermekeknél. Tegyük hozzá, hogy a zajos környezet és a gyors környezeti beszédtempó gyakran együttesen jellemzik a kommunikációs helyzeteket.

A megakadásjelenségek (GYARMATHY 2017) szerves részei beszédünknek, kisgyermekektől is hallhatók. Úgy tűnik azonban, hogy az előfordulásuk a felnőttek megnyilatkozásaiban – bár beszélőtől, beszédstílustól, beszédhelyzettől stb. függenek – megnövekedett az utóbbi évtizedekben, és ez néha szaggatottá, ezért nehezebben feldolgozhatóvá teszi a folyamatos beszédet. Nem mindegyik gyermek képes kiküszöbölni, avagy kompenzálni a beszédfeldolgozását esetleg ily módon zavaró jelenségeket.

A szülők felteszik néha a kérdést, hogy mennyit beszélgessenek a gyermekeikkel. Nem lehet megmondani, hogy egy adott gyermeknek (különböző életkorokban) milyen időtartamú és jellegű verbális kommunikációra van szüksége ahhoz, hogy az anyanyelv-elsajátítása zökkenőmentes legyen (a válasz tehát általában az, hogy „minél többet”). Az viszont tény, hogy az iskolai órákon relatíve keveset beszélnek a tanulók. A 14–18 évesek között végzett felmérés szerint a diákok beszédének összes időtartama egy hét alatt (a tanórákon) átlagosan 6 perc volt (BALATONI 1999). Ha valaki ritkábban vesz részt beszélgetésekben, akkor nyilvánvalóan csökken az az idő, amit beszélésre és beszédmegértésre fordít, ezáltal redukálódik a jártassága e folyamatok működtetésében.

A zavarok helye és mértéke erősen különböző lehet. Előfordulhat egy gyermeknél életkornak megfelelő beszédészlelés és atipikus beszédmegértés, és megfordítva, avagy tipikus mondatértés és nem megfelelő szövegértés. A legsúlyosabb helyzet az, amikor mind a beszédészlelés, mind a beszédmegértés érintett, több folyamat elmaradást vagy zavart mutat. Érdeemes újra hangsúlyoznunk, hogy a mindennapi kommunikáció során a gyermek beszédfeldolgozási nehézségei rejtve maradhatnak, mivel a beszédhelyzet, a beszédpartnerek és a nyelvi kontextus segítséget jelenthetnek a bizonytalan működések esetén. A gyermek kitalálja, valószínűsíti az elhangzottakat, kompenzál. Ez azonban csak átmenet, célzott fejlesztés nélkül a problémák nem oldódnak meg.

## 5. A GOH beszédhallás-szűrő eljárás

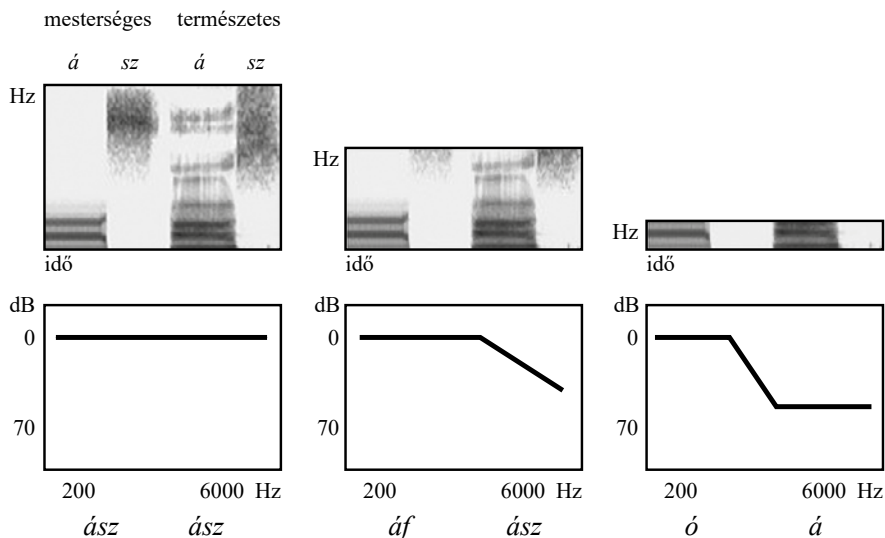
Számos hallásvizsgáló eljárás létezik gyermekek számára, már újszülöttkortól, de ezeknek különféle korlátaik vannak, amelyek egy része a gyermek fiatal korából adódik (PYTEL 1996; FENT 2007). Ahhoz, hogy megbízható eredményt kapjunk a gyermek hallásáról, különösen szűrési céllal, egyszerű és relatíve gyors eljárást kell alkalmaznunk, tekintettel a gyermek életkorára, figyelmére és kognitív képességeire. A **tisztahang-audiometriával** (különböző frekvenciájú szinuszhangok meghallása a leghalkabb elhangzáskor) végzett vizsgálatok a gyermek szubjektív közreműködését igénylik egy, a fiatalabbak számára viszonylag nehéz feladatban, ezért az eredmény ingadozó (pl. BISHARA et al. 1999).

A beszéd alkalmas lehetne a hallásvizsgálatok anyagául, ha nem lenne akusztikailag redundáns, ami azt jelenti, hogy több információt tartalmaz, mint ami feltétlenül szükséges (az artikuláció természetes következményeként). Ha előállíthatók olyan szavak, amelyek csak az akusztikai kulcsokat tartalmazzák, vagyis azokat a legjellemzőbb frekvencia-összetevőket, amelyek meghatároznak egy beszédhangot, akkor ezek alkalmasak lennének a hallás vizsgálatára. Ezek a szavak nem tartalmazznak ugyanis akusztikai többletinformációkat, de ép hallás esetén biztosítják az adott beszédhangok, illetve a hangsorok felismerését. Az egyes beszédhangok akusztikai kulcsait percepciós tesztekkel határozták meg (pl. OLASZY 1985; GÓSY 1989). Ezek a kísérleti eredmények tették lehetővé azt, hogy olyan mesterséges (szintetizált) szavakat hozzanak létre, amelyek csak az akusztikai kulcsokat tartalmazták (GÓSY et al. 1987; GÓSY 2008). Ez adta a GOH beszédhallás-szűrő eljárás anyagát, amelyet az OVEIII (típusú) beszéd-szintetizátorral (és számítógépes ellenőrzéssel) hoztak létre (LILJENCRAANTS 1968).

A **mesterségesen előállított** (szintetizált) **szavakkal** történő hallásszűrés elve a természetes ejtésű és a mesterségesen előállított szavak akusztikai redundanciájának (többletinformációjának) különbségén alapszik<sup>3</sup>. Ez a különbség magyarázza azt, hogy míg a természetes ejtésű szavakat a nagyothallók – a különböző mértékű és típusú halláscsökkenések függvényében – felismerik, addig az ugyanazon, csak akusztikai kulcsokat tartalmazó, szintetizált szavakat nem (HAZAN–FOURCIN 1983). A 3. ábra összefoglaló magyarázatot ad a **GOH-eljárás** működésére hangszíneképek és audiogramsémák segítségével. Az ábra felső része három képet tartalmaz, amelyek az *ász* szó alapján készült akusztikai lenyomatok; a bal oldali mindig a mesterségesen előállított, a jobb oldali a természetes ejtésű szó hangszíneképe (az idő a vízszintes, a frekvenciajellemezők a függőleges tengely mentén láthatók). A hangszíneképpárok jól szemléltetik a szintetizált hangsorok csökkent redundanciájú akusztikai szerkezetét a természetesekéhez képest. A mesterségesen előállított változatban a *sz* intenzív zörejkomponensei 6000 Hz-nél, az *á* intenzív összetevőinek értékei 700 Hz-nél és 1300 Hz-nél vannak. A természetes ejtésű hangsorok sokkal több akusztikai összetevőt tartalmaznak; a *sz* zörejkomponensei mintegy

<sup>3</sup> A mesterségesen előállított (szintetizált) szavak felismerésén alapuló hallásszűrési eljárás 1984-ben szabadalmi oltalmat kapott.

4000 Hz-től 8000 Hz-ig láthatók (kevésbé változó intenzitással); az *á* is mutat (változó intenzitású) összetevőket a felsőbb frekvenciákon is.



3. ábra

*A mesterségesen előállított hangsorok működési elve a beszédhallás vizsgálatára (GOH-eljárás)*

A kétféle hangminőségű *ász* hangsort alul áteresztő szűrővel megszürtük (ez azt jelenti, hogy egy adott frekvenciaérték fölött az összetevők gyakorlatilag nem hallhatók). A szűrt hangsorokról készült hangszíneképeket mutatja a második és a harmadik kép. A második képen 5000 Hz fölött nincs intenzív összetevő, a harmadik képen az erőteljesebb szűrés következtében csak az 1000 Hz alatti összetevők őrződtek meg. Jól látható, hogy mindkét felül áteresztő szűrés eredményeképpen jóval kevesebb észlelhető komponens tartalmaz a szintetizált *ász* hangsort, mint a természetes változat.

Az alsó képsor séma-audiogramokat szemléltet, a második és a harmadik kép a felül áteresztő szűréseknek megfelelő halláskárosodást imitálnak. A középső képen 5000 Hz-nél, a harmadik séma-audiogramon 1000 Hz-nél kezdődik az imitált halláscsökkenés. Ez azt jelenti, hogy az a feltételezett nagyothalló, akinek az audiogramja megfelel a középső vagy a harmadik képen láthatóknak, az elhangzott *ász* hangorból a felső képsor hangszíneképein megjelenő összetevőket képes csak meghallani. Mindkét halláscsökkenés esetében sokkal kevesebb áll rendelkezésére a szintetizált *ász* esetében, mint a természetes ejtésűnél. Az ábra alsó sorában az észlelt hangsorok írott formája látható. Szűrés nélkül, ép hallás esetén mind a szintetizált, mind a természetes ejtésű *ász* pontosan azonosítható; ezt szemléltetik a bal oldali képek. A középső séma-audiogram, illetve a megmaradó komponensek észlelése a szintetizált *ász* esetében az *áf*, a természetes ejtés esetében még az eredeti *ász* hangsort eredményezi. A súlyosabb

halláscsökkenésben a mássalhangzó felismerése már egyik hangsorváltozat esetén sem lehetséges; az *á* magánhangzó természetes ejtésben azonban még azonosítható. A szintetizált hangsor akusztikai maradványa csupán egy *ó* magánhangzó azonosítását teszi lehetővé.

A GOH hallásszűrő készülék (4. ábra) nyelvi anyaga összesen 80, egy szótagból álló szót tartalmaz (40 szolgál a nem beszédhibás és további 40 szó a beszédhibás gyermekek vizsgálatára). A szavak egy részének a jelentése ismert a gyermekek számára, más részük nem, ez azonban nem okoz problémát a visszamondáskor (az anyanyelv-elsajátítás során a gyermek sokszor hall és észlel értelmetlen hangsorokat, amelyekhez a jelentés később társul).



4. ábra

*Az érintőgombos GOH-készülék a fülhallgatóval*

Egy fül hallásának a vizsgálatára 10 szót kell a gyermeknek megismételnie. A feladat egyszerű: az elhangzott szót kell az elhangzás után azonnal visszamondani. A teszt hároméves kortól elvégezhető; ha a gyermek hajlandó az együttműködésre, akkor fiatalabb életkorban is használható. Ötévesekkel a tesztelési idő a magyarázattal együtt mintegy 10 perc, vagyis a hallásszűrés gyors, a feladat mindennapi, és az eredmény megbízható.

A hallásszűréshez kétféle hangerő áll rendelkezésre a környezeti zajszint függvényében, és kiválaszthatók a szükséges szósorok. Az értékelés egy előre nyomtatott tesztlapon történik, ahol a vizsgáló jelöli a gyermek válaszait, függetlenül attól, hogy helyes volt-e, vagy sem (5. ábra). A tesztlap szóanyagát ép halló és nagyothalló gyermekek ismétléseiből állították össze (vö. Gósy 2008). A vizsgáló a jelölések alapján megállapítja, hogy hallászavar vagy globális beszédészlelési probléma valószínűsíthető-e a gyermeknél, és dönt a további szakvizsgálatok szükségességéről. (Az adott életkorban elvárt helyes ismétlések számát és az értékelés módját a GOH-eljárás használati utasítása tartalmazza.) A GOH-eljárás megbízhatóan jelzi

a gyermekek hallásának épségét, illetve az esetleges halláscsökkenést. Ép hallás esetén pedig információt nyújt a globális beszédészlelési teljesítményről. A GOH használata és az adatok értékelése könnyen megtanulható.

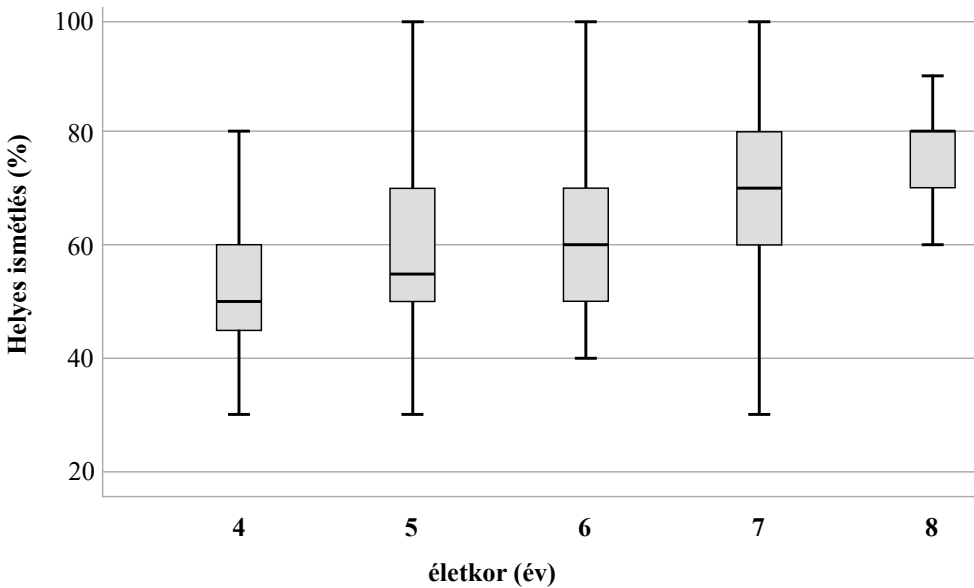
Sorszám	I. Ép hallás	II. Beszédészlelési zavar	III. Enyhe hallászavar	IV. Hallászavar
0.	mos	mus bus bos vus vos mosz los lus nos mös mes musz	busz bosz buf vot vut mo tuf mot bot	pop u op up -
1.	sír	síp sí szíl sél sűr sé sér szír szér	fúr fül fut hit hil fot fup fop hú hó ér	úr út ór ót tú tó ú ó -
2.	bab	dad bob dab bag gab bob a gad dag dod dog	pap pat pot top kap tap kup pup tup kop	tú tó út ót ó ú -
3.	szűz (szül)	zűsz sziz szű szűsz síz szí szé	só sú fül fűt fűz főz fő fűfű fő	hó hú tó tú ó ú -
4.	ágy	ág	áj aj ad ód	tó ót tú út a ó -
5.	sző	szőr szőn szóm szől szörp szűr szű szül sző	fő fű fő ü só sú ló	hó hú tó tó tú ó -
6.	bor	dur bul por pur pol pul dor	pú tú	ó ú -
7.	csók	sók csót csó só sűk csűk	fők fűk fő fű fű fut	tó tú út ót ó ú -
8.	kút	út kúk tút tú tuk pukk ú	ó ú	ő -
9.	ész	tész pész mész méz ész	és ős ős éf óf őf é	út ú ó -
Név:		Szül. idő:		Dátum:

### 5. ábra

*A GOH-eljárás tesztlapjának egyik szósora. A bekarikázott hangsorok egy adott gyermek ismétléseit szemléltetik. A harmadik és a negyedik oszlopba került jelölések azt mutatják, hogy a gyermek hallása az adott frekvencián eltér az ép hallástól*

A GOH-eljárással kapott adatokat összevetették a tisztahang-audiometriával és tympanometriával (objektív vizsgálat, amely a középfül működéséről nyújt információt) végzett diagnosztikai eredményekkel ép halló és hallássérült gyermekek esetében (BALÁZS–GÓSY–SZABÓ 1997; GÓSY 2008). Az eredmények minden esetben alátámasztották, hogy a GOH-val kapott adatok megegyeznek a klinikai vizsgálatok eredményeivel.

644 egynyelvű gyermeknek (négyévestől nyolcévesig) a mesterségesen előállított szavakra adott ismétléseit dolgozták fel (GÓSY–KREPSZ 2018). Az eredmények azt mutatták, hogy az életkor előrehaladtával a szintetizált hangsorok felismerése egyre pontosabbá vált. A gyermekek hibái csökkentek, az akusztikai kulcsok felismerése biztosította a beszédhangok és a hangsorok mentális reprezentációinak aktiválását. A 6. ábra a két fülben hallott szavak helyes ismétléseinek összegzett átlagát szemlélteti az életkor függvényében. Az egyes csoportokban relatíve nagy a szóródás, vagyis minden életkorban előfordultak olyan gyermekek, akik igen jó, illetve gyenge teljesítményt nyújtottak (az utóbbi legritkábban a nyolcéveseknél fordult elő).



6. ábra

*A GOH-eljárással (mesterségesen előállított hangsorok ismétlése fülhallgatón keresztül) kapott eredmények 4 és 8 év közötti gyermekeknél*

A szerzők elemezték, hogy milyen arányban volt tapasztalható ép hallás és ép globális beszédészlelés, illetve halláscsökkenés és beszédészlelési zavar a gyermekeknél. A jobb fülben hallott hangsorok ismétlései azt mutatták, hogy a négy- és ötévesek mintegy 80%-ának volt ép a hallása és a globális beszédészlelése; különböző mértékű halláscsökkenés néhány százalékban fordult elő, a globális beszédmegértés gyengébb teljesítménye 15% körülnek adódott. A hét- és nyolcévesek esetében 70% és 80% közé esett azoknak az aránya, akik ép hallást és jó beszédészlelést mutattak, halláscsökkenés mintegy 2%-ban volt kimutatható. 20% fölé emelkedett azonban a gyenge globális beszédészlelési teljesítményt mutatók aránya. A leggyengébb eredményeket a hatéveseknél tapasztalták, náluk alig több 60%-nál az ép hallást és jó beszédészlelést mutatók aránya. Globális beszédészlelési zavar 30%-ban fordult elő, a halláscsökkenés előfordulása elenyésző volt. A bal fülbe érkezett hangsorok ismétlései hasonló eloszlásokat mutattak, mint a jobb fül esetében. A globális beszédészlelés több gyermeknél mutatott valamivel gyengébb teljesítményt a jobb fülben hallottak ismétléseihez képest.

## 6. A GMP-diagnosztika

A gyermekek beszédészlelési és beszédmegértési folyamatainak megítélésére fejlesztették ki a **GMP-diagnosztikát** (1985 és 1995 között), amely egy sztenderdizált teszteljárás<sup>4</sup> (Gósy 1995/2006). Célja a magyar anyanyelvű gyermekek beszédfeldolgozási teljesítményének vizsgálata és minősítése. A sztenderd értékek révén megállapítható, hogy az adott gyermek teljesítménye megfelel-e az életkorának, vagy nem, a diagnosztika továbbá információt nyújt az elmaradás vagy a zavar mértékéről is. A GMP 20 különböző tesztet tartalmaz, közülük az egyik – a korábbiakban tárgyalt – GOH-eljárás (GMP1), egy másik teszt pedig a nem beszélő gyermekek beszédfeldolgozásának vizsgálatára irányul (GMP20). A diagnosztika 18 tesztje vizsgálja az akusztikai, fonetikai, fonológiai, szeriális, transzformációs, vizuális és ritmusészlelést, a hangdifferenciálást, a mondatértést, a szövegértést, a mentális lexikon aktiválását, a rövid idejű verbális és vizuális memóriát, a szavak centrális szintézisét, az agyfélteke-dominanciát, a kéz- és szemdominanciát, valamint az irányfelismerést. A teszt eszközigényes (vö. 7. ábra); a felvétele mintegy fél órán át tart, az időtartam függ a gyermek életkorától. Az óvodásokkal 16, az iskolásokkal 17 teszt felvétele szükséges (beleértve a GOH-eljárást a hallás épségének szűrésére).



7. ábra

*A GMP-diagnosztika eszközei (képfüzet a mondatértés vizsgálatához, beszédanyagok adathordozón, cipzár és kapcsok a kezesség megállapításához, képsor a rövid idejű vizuális memória teszteléséhez, színes kockák a transzformációs észlelés vizsgálatához)*

<sup>4</sup> Az alapesztet bárki használhatja, a teljes diagnosztikát csak a minisztérium által akkreditált tanfolyamot sikeresen elvégzők, ők jogosultak az eredmények diagnosztikus megfogalmazására is.

A GMP-diagnosztikához egy négyoldalas alaptesztlap és további kiegészítő tesztlapok tartoznak, amelyek egyrészt segítik, irányítják a tesztfelvételt, másrészt ezekre kell rögzíteni a gyermek válaszait, reakcióit, döntéseit, meghatározott jelrendszer alapján. A diagnosztika eredményei a felvételt követően azonnal értelmezhetők és a szülővel, pedagógussal megbeszélhetők. A gyermek teljesítménye alapján megtervezhető a célzott fejlesztés (GÓSY–IMRE 2007). A megalapozott diagnosztikára épülő egyéni és csoportos fejlesztés igen jó eredményeket igazolt (pl. SZÁNTÓ 2017).

Az egyes **tesztek anyagának** és a **tesztfelvétel módszerének** alábbi leírásában ismertetjük a GMP valamennyi tesztjét, de nem térünk ki az egyes életkorokban meghatározott sztenderd értékek bemutatására, azok a diagnosztika útmutató könyvében megtalálhatók (GÓSY 1995/2006). A diagnosztika tesztjeit illesztették a hierarchikus beszédfeldolgozási modell szintjeihez, illetve folyamataihoz (vö. 2. ábra), az ott zajló életkor-specifikus dekódolási műveletek minősítésére. A GMP2–6, valamint a 12, 15 és 19-es tesztek rögzített hanganyagok (férfi beszélő bemondásában), amelyeket részben adathordozóról (DVD, tablet, laptop, asztali számítógép), de szabad hangtérben hall a gyermek, részben pedig fülhallgatón keresztül (ilyenek a GOH, a GMP15 és a GMP19).

A beszédészlelést tíz különböző teszttel vizsgálják. Ezeknek a beszédanyagoknak egy része célzottan torzított, hogy hangsúlyozottan az adott percepció szint működéséről nyújtson információt. A beszédészlelést vizsgáló tesztek szavait és mondatait nem kell a gyermeknek megértenie, képesnek kell lennie azonban a hibátlan ismétlésükre az életkorában elvárt mértékben. A GMP2 a beszédészlelés akusztikai-fonetikai szintjének vizsgálatára irányul. A gyermeknek 10 db, ún. fehér zajjal elfedett (jel/zaj viszony = 4 dB) mondatot kell megismételnie. Példák: *Rakjátok össze a játékokat!*, avagy *A repülőgép most szállt le*. Ugyanilyen fehér zajjal, a mondatokkal megegyező módon elfedett szavakat tartalmaz a GMP3, például *meggy, étterem, ablak*. A feladat itt is az elhangzás utáni azonnali ismétlés. A GMP4-ben 10 db szűk frekvenciás mondat azonosítása történik (2200 Hz és 2700 Hz közötti, 36 dB/oktáv meredekségű sáváteresztő szűrő alkalmazásával létrehozott mondatanyag). A teszt információt nyújt a fonetikai észlelésről, a beszédhangok, hangkapcsolatok és hangsorok felismeréséről. Példák: *A tejfölt a macska itta meg.*, *Dobd a papírt a szemétkosárba!*. A gyermek feladata a mondatok egyszerű hallás alapján történő azonnali ismétlése. A GMP5-ös tesztben 10, kismértékben gyorsított, többféle fonológiai koartikulációs jelenséget tartalmazó mondat azonnali ismétlése történik (a gyorsítás mértéke 20%, műszeresen). Példák: *Az irigység rossz tulajdonság.*, *A katonák felesküdték a zászlóra*.

Természetes ejtésben, férfihanggal rögzített mondatokkal, torzításmentesen vizsgálják a mondatok azonosítását a 3 év körüli, valamint a súlyos elmaradást mutató gyermekek esetében (GMP6). Példák: *Most nyílnak a hóvirágok.*, *Megettétek az ebédet?*. Az anyanyelv-elsajátításban nagy a szerepe az ajakműködésnek, valamint egyéb, látható artikulációs mozgásnak. Ez a vizuális információ segíti a gyermeket az artikuláció elsajátításában és a beszédészlelésben, amit a GMP7-es teszt elemez. A gyermeknek a vizsgáló szájmozgása alapján kell állatneveket felismernie (pl. *oroszlán, egér*). A beszédritmus észlelését időmértékes verssorok ismételtetésével vizsgálják (GMP14). A szeriális észlelés (GMP10) diagnosztizálása értelmetlen hangsorok visszamondatásával történik. Példák: *galalajka, zseréb, vucsityó*. A GMP17-es teszt a beszédhangmegkülönböztetés működésének megismerésére irányul. A vizsgáló értelmetlen, egy és két



szótagú hangsorparókat mond a gyermeknek, amelyek vagy azonosak (pl. *hem* vs. *hem*), vagy csak egy fonológiai jegyben különböznek egymástól (pl. *nőcs* vs. *nőcs*, *aszá* vs. *asá*, *gev* vs. *bev*). A gyermek feladata annak eldöntése, hogy egyforma vagy nem egyforma hangsorokat hallott. A transzformációs észlelés (GMP18) feladatmegoldása arra ad választ, hogy képes-e a gyermek az elhangzásnak megfelelő beszédhang-differenciálást és sorozatészlelést egy vizuális/taktilis transzformáció során megvalósítani. Az eredmény rávilágít arra, hogy miként felelteti meg (majd) a gyermek az anyanyelvi beszédhangokat a betűknek.

A beszédmegértést a GMP-ben a szóaktiválás, valamint a mondat- és a szövegértés tesztjeivel vizsgálják. Egy gyermek szókincsének a becslése csaknem lehetetlen, a benyomások nemegyszer tévesek lehetnek. A GMP11-es teszt a lexikális hozzáférésről, azaz a szóaktiválásról nyújt információt. A gyermeknek meghatározott idő alatt a *ma*- és *ke*- szókezdetekkel kell szavakat előhívnia. A szövegértési teszt (GMP12) a szemantikai szerkezetek értelmezésének, a részletek azonosításának, az ok-okozati viszony felismerésének, illetve az értelmezés működésének vizsgálatára irányul. A gyermek egy rövid (közel 2 perces) mesét/történetet hall, amelynek az elhangzása után 10 kérdést kell megválaszolnia (a vizsgáló több mese közül választhat). A mondatértés vizsgálata (GMP16) az adott életkorban szükséges szemantikai sajátosságok és a grammatikai struktúrák feldolgozásába enged betekintést. Ehhez a teszthez 10 darab színes rajzpár tartozik; az egyes párok csak egyetlen dologban különböznek egymástól (a 7. ábrán látható egy rajzpár). Az adott képpárhoz tartozó tesztmondat mindig az egyik képre igaz (pl. *A nyuszi kergeti az oroszlán.* vs. *A nyuszi kergeti az oroszlánt.*). A gyermeknek az adott tesztmondat jelentését ábrázoló képre kell rámutatnia.

A folyamatok egyfajta vezérlését, illetve a gyermeknek az ezekben nyújtott teljesítményét két dichotikus teszt vizsgálja. A centrális működésről nyújt felvilágosítást az, hogy a gyermek képes-e a két fülébe egy időben érkező, de különböző nyelvi információt szintetizálni (GMP15). A beszédanyag öszetett szavakból áll (pl. *kávéfőző*). Az ösztetétel első szavát a gyermek a jobb, a másodikat a bal fülében hallja, egyidejűleg. A feladata az ösztetett szó felismerése és kimondása a két (különböző fülben) hallott szó szintetizálása alapján. A lateralizáció vagy agyfélteke-dominancia kialakulásának, illetőleg kimutathatóságának késése összefügg a beszédpercepció zavarokkal, vizsgálata hatéves kortól szükséges. A beszédanyag 30, két szótagú, páronként hangzásukban hasonló szóból áll (pl. *mozdony*, *macska*). A gyermeknek a két fülébe egy időben érkező, különböző szavakat meg kell ismételnie (GMP19).

A diagnosztika a beszédfeldolgozáshoz kapcsolódó, abban szerepet játszó folyamatokat is ellenőrzi. A rövid idejű verbális memória (GMP8) megfelelő működése szükséges a beszéd-észleléshez, a helyes artikulációhoz, a rövid idejű vizuális memóriával (GMP9) együtt az írott nyelv megtanulásához. Az előbbihez a vizsgáló 12 darab egy, két és három szótagú szót mond (pl. *szív*, *labda*, *hőember*), az utóbbi tesztben pedig 12 darab kis, színes képet mutat egy lapon a gyermeknek (l. a 7. ábrán). A feladat a hallott szavak, illetve a látott képek felidézése és felsorolása. A GMP13-as tesztben ellenőrzik a gyermek kéz- és szemdominanciáját, illetve három egyszerű feladattal az irányfelismerését. Tipikus fejlődés esetén 3 éves korra általában egyértelműen kialakul a domináns kéz. Ha a gyermek 5 vagy 6 éves korában is még bizonytalan a domináns kezének használatában, annak súlyos következményei lehetnek a beszédpercepcióra, az irányfelismerésre, illetve az írott nyelv megtanulására.

Nem vagy alig beszélő, 3-4 éves gyermekek beszédpercepció folyamatát próbáljuk megítélni a GMP20-as teszttel. Ekkor a beszédindításhoz és a beszédinduláshoz szükséges alapfolyamatokról igyekszünk információt kapni (a gyermek beszédhallása, egyes percepció folyamatok működése, szituatív beszédmegértési teljesítmény, mondatértés, kooperációs készség).

A GMP-diagnosztika tesztjeinek sorrendjét a vizsgáló szabadon választhatja meg, azoknak szigorúan kötött sorrendje nincsen. A tesztek egymásutániságát a gyermek életkorára és feltételezett beszédfeldolgozási nehézségeire tekintettel kell megtervezni.

## 7. Eredmények a GMP-diagnosztikával

Míthogy a beszédfeldolgozási folyamatok rejtetten működnek, ezért az esetleges elmaradások vagy zavarok nehezen felismerhetők. A felszínen tapasztalható jelenségek több okra is visszavezethetők, ezért csak valószínűsítik (többek között) a beszédészlelés és a beszédmegértés elmaradását az anyanyelv elsajátítása során. A látszólag tipikus fejlődést mutató gyermekek beszédfeldolgozásának vizsgálata éppen azért fontos, mert sok esetben fennállnak az elmaradások, zavarok, és akár hosszú távú következményes problémákat is okozhatnak, de fel sem merül, hogy ezek a nehézségek léteznek. Több kutatás foglalkozott a nyelvi zavart nem mutató gyermekek beszédészlelésével és beszédmegértésével (pl. MARKÓ 2007; GÓSY–HORVÁTH 2006, 2011; IMRE 2007). A látszólag tipikusan fejlődő gyermekek egy részénél is igazoltak azonban beszédészlelési és beszédmegértési nehézségeket. Az 1. táblázat 4 és 9 év közötti, csoportonként 50 tipikus fejlődésű gyermek három beszédészlelési folyamatban nyújtott eredményeit összegzi (GÓSY 2012). Az átlagos értékek kisebb-nagyobb mértékű elmaradásokat mutatnak (a 4 és 5 évesek GMP2-ben nyújtott teljesítményét kivéve), ami azt jelzi, hogy minden korcsoportban akadtak olyan gyermekek, akiknek a teljesítménye elmaradt az elvárt szinttől.

### 1. táblázat

*Tipikusan fejlődő gyermekek teljesítménye az akusztikai-fonetikai (GMP2), a fonológiai (GMP5) és a szeriális (GMP10) észlelésben*

Életkori csoportok	GMP2 (%)		GMP5 (%)		GMP10 (%)	
	átlagérték	elvárt érték	átlagérték	elvárt érték	átlagérték	elvárt érték
4 évesek	63,4	50	44,1	50	64,7	80
5 évesek	76,6	70	61,2	70	75,4	90
6 évesek	79,8	90	66,6	90	75,6	100
7 évesek	81,3	100	74,7	100	80,8	100
8 évesek	86	100	79,9	100	84,3	100
9 évesek	90,5	100	80,6	100	88,2	100

Különösen nagy mind az elméleti, mind a gyakorlati jelentősége annak, hogy az ún. **rizikó-gyermekeknek** a beszédfeldolgozási folyamatait teszteljék. Ilyenek a hallássérültek, a megkésett beszédfejlődésűek, a beszédhibások, a specifikus nyelvi zavarral küzdők, az olvasási, illetve tanulási nehézséggel küzdők, a diszlexiások, a figyelemzavarosok, a mentálisan érintettek (és a felsorolás még folytatható). A továbbiakban ezen gyermekek néhány csoportjának GMP-vizsgálatával kapott eredményeit ismertetjük.

## 7.1. Hallássérültek

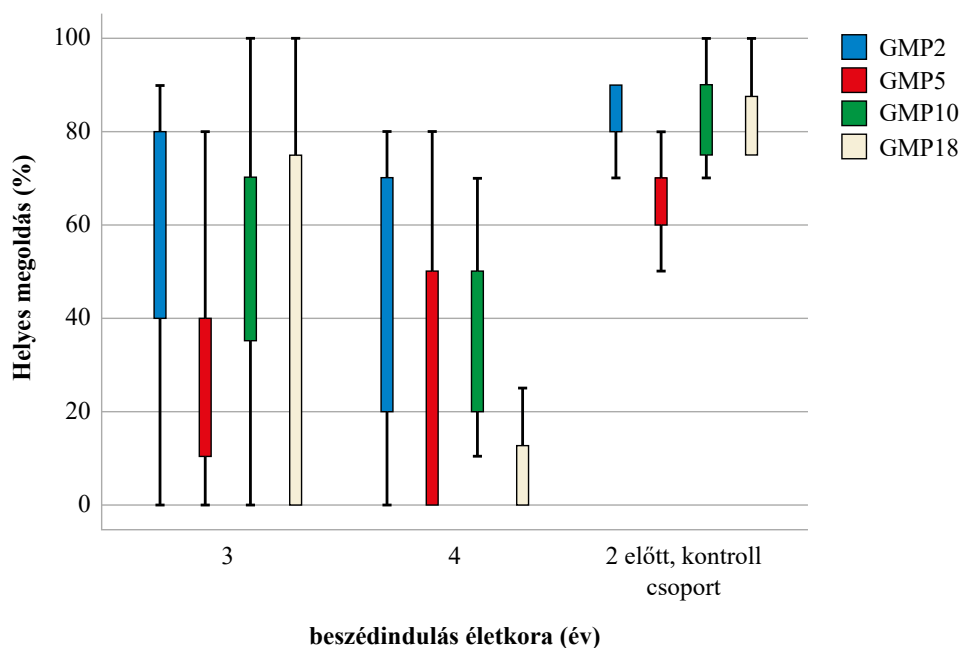
A hallás csökkent volta vagy hiánya miatt a nyelvi folyamatok módosulnak, a fejlődés sajátos, soktényezős és nagymértékben egyénfüggő. A tisztahang-audiometriás és a szabad hangterés hallókészülékes mérések hozzávetőleges információt adnak arról, hogy milyen lehet a hallássérült gyermek beszédfeldolgozása. A pontos információ a célzott beszédpercepció tesztelés révén kapható meg. A beszédfeldolgozási mechanizmus már az első szinten, az akusztikus paraméterek azonosításában zavart szenved. Ennek következtében hibás a fonetikai észlelés, és nem megfelelően működik a fonológiai észlelés. Az alapvető beszédpercepció folyamatok tökéletlen működése az összes többi folyamatot negatívan befolyásolja. A vizsgálatok azt igazolták, hogy a hallássérültek minden beszédfeldolgozási folyamatában jelentős az eltérés az ép hallókhöz képest (BOMBOLYA 2007).

52, hallókészüléket viselő 6-7 éves gyermeket vizsgáltak a GMP-diagnosztikával (a halláscsökkenésük mértéke 50 dB és 80 dB között volt mindkét fülön). A természetes ejtésű (nem torzított) mondatok azonosítása során (GMP6) a gyermekek meg sem közelítették a 3–3,5 éves korban elvárt szintet. Az életkoruknak megfelelő 90%-os teljesítmény helyett a GMP2-tesztben (akusztikai-fonetikai észlelés mondat szinten) 40%-os, a GMP3-tesztben (akusztikai-fonetikai észlelés szószerint) mintegy 70%-os, a GMP4-ben (beszédhangok felismerése) átlagosan 20%-os eredményt értek el. A GMP5-teszt (fonológiai észlelés) adatai 10% és 30% közöttinek adódtak. A szeriális észlelésben (GMP10) a 2-3 évvel fiatalabb gyermekek teljesítményét mutatták. Nem meglepő ugyanakkor, hogy ezek a gyermekek igen jó eredményt hoztak a vizuális észlelési tesztben, vagyis a szájról olvasásban (GMP7). Némegyszer az életkorukban elvártnál is jobb volt a vizuális észlelésük. A 6-7 éves nagyothalló gyermekek a mondatértési tesztben (GMP16) a négyéveseknél tapasztalt teljesítményt nyújtották. Mindehhez hozzá kell tennünk, hogy a gyermekek egy részével a tesztek egyáltalán nem voltak felvehetőek; a szövegértési teszt esetében még a feladatot sem értették meg (BOMBOLYA 2007).

## 7.2. Megkésett beszédindulás

A gyermekkori beszédzavarok egyik gyakori jelensége az anyanyelv-elsajátítás megindulásának késése. Sok kutatás foglalkozik ezeknek a gyermekeknek a nyelvi teljesítményével későbbi életkorukban is (RICE et al. 2008; POLL–MILLER 2013). Vizsgálták, hogy a 3, illetve 4 évesen

beszélni kezdő gyermekeknek milyenek a beszédészlelési és beszédmegértési működéseik hatéves korukban (GÓSY–GYARMATHY 2018). Az eredmények azt mutatták, hogy a beszédindulás ideje nagymértékben meghatározza a vizsgált receptív folyamatokat. Mind a 3, mind a 4 évesen beszélni kezdők szignifikánsan elmaradtak a kontrollcsoport teljesítményétől valamennyi tesztben. A 4 évesen beszélni kezdőknél nagyobb mértékű folyamatzavarokat és nagyobb egyéni különbségeket tapasztaltak a 3 évesen beszélni kezdőkhöz képest. A 8. ábra négy teszt eredményeit mutatja a megkésve és az időben beszélni kezdők csoportjaiban, az akusztikai-fonetikai (GMP2), a fonológiai (GMP5), a szeriális (GMP10) és a transzformációs (GMP18) észlelést.



8. ábra

*Megkésve beszédfejlődésű és tipikusan fejlődő gyermekek beszédpercepciósi teljesítménye*

### 7.3. Beszédhibás gyermekek

Az ép halló, időben beszélni kezdő gyermekek egy részénél előfordul, hogy a fiziológiás beszédhiba időben elhúzódik, a gyermek nem tud korrigálni. Ez az artikulációs gesztusok kivitelezésének zavara, amely érinthet egyetlen beszédhangot, beszédhangok csoportját, avagy több magán- és mássalhangzót. A beszédhibás gyermekek kevesebb mint egyharmadánál nem találtak eltérést a beszédfeldolgozásban, az elmaradások és zavarok azonban a kétharmaduknál

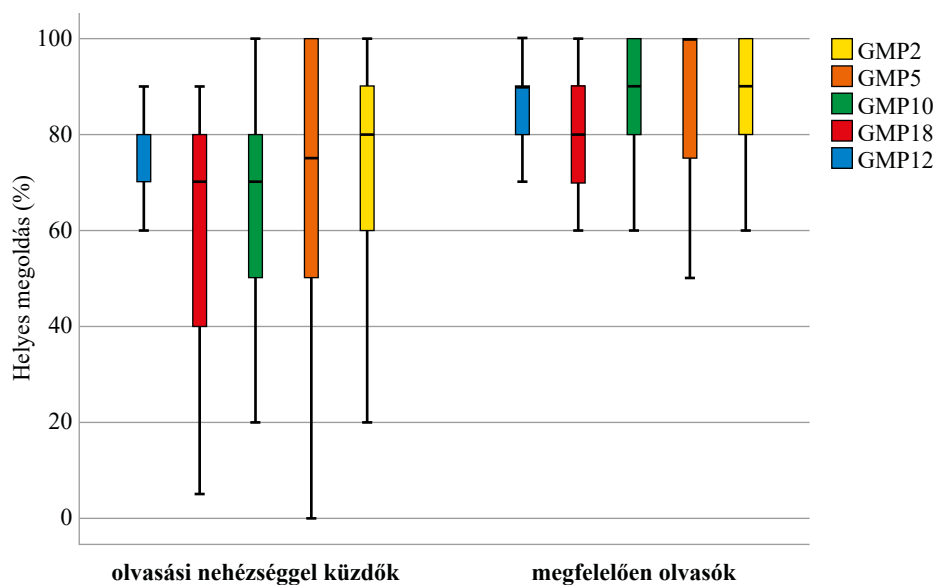
tapasztalhatóak voltak, mind a beszédészlelés, mind a beszédmegértés folyamataiban (pl. BÓNA 2007; ROSTA–SUCHNÉ RUMPLI 2007; GÓSY–HORVÁTH 2008). A beszédhibás tanulók a percepció nehézségeik és a társuló beszédmegértési problémáik miatt gyakran küszködnek iskolai kudarccokkal.

BÓNA (2007) súlyosan beszédhibás 4 és 5 éves gyermekekkel végzett kutatása azt mutatta, hogy a beszédhibások percepció mutatója (a vizsgált percepció folyamatokban nyújtott átlagos teljesítmény) 54%, ami jóval elmarad az ép beszédűekétől, akiknél ez az érték 73%. A 4 éves beszédhibás és az ép beszédű gyermekek között gyakorlatilag nem volt különbség a percepció mutatóban, a beszédhibás 5 évesek mutatója azonban jelentősen elmaradt a kontrollgyermekétől, a különbség közel 40%.

Különbő okoknál fogva a beszédhibás gyermekek egy része úgy kerül iskolába, hogy a beszédhibájuk nem rendeződött. GÓSY és HORVÁTH (2008) súlyos beszédhibás iskolások (első, második és harmadik osztályosok) beszédfeldolgozási teljesítményét vizsgálta, és azt tapasztalták, hogy az osztályfoktól függetlenül nem volt különbség a gyermekek beszédészlelési működéseiben (GMP2, 3, 5, 10). Pozitívan változott ugyanakkor a szavak előhívhatósága, azaz a mentális lexikon aktiválása (GMP11), továbbá biztosabbá vált a beszédhangok megkülönböztetése (GMP17) és a transzformációs észlelés is (GMP18). A mondat- és a szövegértés (GMP16 és 12) eredményei stagnálást mutattak. A beszédhibás tanulók és az ép beszédű csoportok között csaknem valamennyi tesztben jelentős (szignifikáns) volt a különbség.

## 7.4. Olvasási nehézséggel küzdő tanulók

A nyelvi képesség bármely összetevője oka lehet az olvasási nehézségnek, meghatározó azonban a beszédfeldolgozási teljesítmény. Jelenlegi ismereteink alapján a következő alapösszefüggések állapíthatók meg. Beszédészlelési zavar, illetőleg elmaradás és ép értés esetén az olvasás és írás technikájának kisebb-nagyobb mértékű zavara látható. Az időviszonyok bizonytalan észlelése jelentkezhet a beszédhangok megkülönböztetésének hibáiban. Beszédmegértési elmaradás és ép észlelés esetén olvasásértési és értelmezési nehézségeket tapasztalunk, továbbá megmutatkoznak a gondolkodási korlátok. A beszédpercepció folyamat nem megfelelő működését súlyosbíthatja a szóismeret hiánya, a lassú vagy bizonytalan lexikális hozzáférés, a memóriaprobléma, a kézdominancia (illetőleg az agyfélteke-dominancia) kialakulatlansága s az ezekkel gyakran együtt járó irányfelismerési zavar. Az olvasástanulás nehezítettsége már óvodáskorban előre jósolható, ha a fonológiai, a szeriális, a transzformációs észlelés és a szövegértés elmaradott (pl. JÓZSA 2006; IMRE 2007; JÓZSA et al. 2015; GÓSY et al. 2018). A 9. ábra olvasási nehézséggel küzdő kilencéves gyermekek és megfelelően olvasók beszédészlelési és szövegértési eredményeit összegzi.



9. ábra

*Olvasási nehézséggel küzdő és megfelelően olvasó kilencéves gyermekek beszédfeldolgozási teljesítménye*

## 7.5. Ikergyermekek beszédfeldolgozása

Az ikergyermekek nyelvfejlődésének tanulmányozása lehetőséget nyújt a genetikai és a környezeti hatások elemzésére. 198 egynyelvű, öt és kilenc év közötti, tipikus fejlődésű ikergyermek és 198 nem ikergyermek beszédészlelését és beszédmegértését vizsgálták a GMP-diagnosztika négy tesztjével (GÓSY–PREGITZER 2019). (A két csoport adatközlői minden lehetséges tekintetben kiegyenlítették voltak.) A kutatási kérdés arra irányult, hogy vajon a környezet hatást gyakorol-e az adott beszédpercepciósi folyamatokra, vagyis az a tény, hogy a család egyidejűleg két azonos korú gyermekhez intézi a közléseit. A kutatásban a szeriális észlelést (GMP10), a beszédhang-differenciálást (GMP17), a mondatértést (GMP16) és a szövegértést (GMP12) vizsgálták. Az eredmények szerint csupán a beszédhang-differenciálásban, illetve a 7 és a 8 évesek szeriális észlelésében nincs különbség az ikrek és a nem ikrek között. A mondatmegértés csak az 5 éveseknél mutatott különbséget az ikrek és a nem ikrek között. A szövegértésben (az 5 évesek kivételével) minden életkorban szignifikáns különbséget adtak az ikrek és a nem ikrek között, az utóbbiak teljesítettek jobban. Az eredmények azt igazolták, hogy a környezetnek meghatározó szerepe van a beszédfeldolgozási folyamatok egy részének fejlődésében. Az azonos korú testvérek relatíve kevesebb beszédingert kapnak, mint a nem ikrek, és ez egyes folyamatok esetében valamilyen mértékben lassítja a fejlődést. A szülők, a család a közléseit rendszerint az

ikrekhez együtt, azaz mindkét gyermekhez intézi, az egy(ik) (iker)gyermeknek szóló üzenetek ritkábbak, amihez az élethelyzet, maga az ikerlét is hozzájárul (és ami teljesen más a nem azonos korú testvérek esetében). Az ikrek tehát rizikógyermeknek tekintendők, ez pedig felhívja a figyelmet a nyelvi teljesítményük rendszeres ellenőrzésére.

## 8. Kitekintés

A magyar beszédkutatók az elsők között voltak a kísérletes kutatások indulásának hajnalán, a 19. században, hogy minél jobban megismerjék a hangzó nyelvet, és objektíven tudják tanulmányozni. Az egyik fő céljuk az volt, hogy a „hibás beszédűeken” segítsenek, és a kapott vizsgálati eredményeiket hasznosítsák (VÉRTES 1980). A ma beszédkutatója is igyekszik az alap kutatások mellett, azokra ráépítve olyan alkalmazott kutatásokat végezni, amelyek közvetve vagy közvetlenül felhasználhatók a gyakorlatban. A beszédfeldolgozás korszerű kutatása alig néhány évtizedes. A beszédészlelés és a beszédmegértés kísérletes megközelítésének eredményei (Gósy 1989), illetve a beszéd akusztikai jellemzőinek elemzési adatai (OLASZY 1985) tették lehetővé az e tanulmányban bemutatott GOH-hallásszűrő és a GMP-diagnosztika fejlesztését. Mindkettőt több mint 30 éve használják az egész országban és a határon túli magyarok területeken is.

Érdeemes újra és újra hangsúlyozni, hogy a beszédfeldolgozás elmaradásai, zavarai éven át észrevétlenül maradhatnak, és okoznak következményes problémákat a tanulásban, a viselkedésben, sőt felnőttkorban a munkában is. Ezeket a nehézségeket nem lehet „kinőni”. Minél hamarabb felismerjük a problémát, a veszélyeztetettséget, pedagógusként és szülőként figyelünk a mindennapokban tapasztalható jelekre, annál nagyobb lesz az esély a megfelelő terápia, fejlesztés alkalmazására. A probléma azonosításához elengedhetetlenek a megfelelő diagnosztikai eljárások, amelyek rendelkezésre állnak, és amelyeket a pedagógusok már óvodáskortól használhatnak. Nagy a felelőssége ebben a pedagógusoknak (pl. óvodapedagógusok, tanítók, tanárok, logopédusok, gyógypedagógusok), a fejlesztőknek és mindazoknak, akik gyermekekkel foglalkoznak (pl. bölcsődei gondozók, védőnők, pszichológusok). A beszéd észlelése és megértése szükség esetén fejlesztendő, a következményes gondok, kellemetlenségek megelőzhetők, a problémák korrigálhatók (pl. SZÁNTÓ 2017).

### FOGALMAK

*beszédfeldolgozás; hallás; beszédészlelés; akusztikai észlelés; fonetikai észlelés; fonológiai észlelés; szeriális észlelés; transzformációs észlelés; vizuális észlelés; ritmusészlelés; beszédhang-differenciálás; beszédmegértés; mentális lexikon; vezetési típusú halláscsökkenés; idegi eredetű halláscsökkenés; kevert típusú halláscsökkenés; tisztahang-audiometria;*

*tympanometria; értelmezés; hierarchikus beszédmegértési modell; GOH-eljárás; mesterséges (szintetizált) szavak; GMP-diagnosztika; rizikógyermekek*

## IRODALOM

- BALATONI Teréz 1999. A megszólalás esélyei az iskolában. In V. RAISZ Rózsa – H. VARGA Gyula (szerk.): *Nyelvi és kommunikációs kultúra az iskolában*. A Magyar Nyelvtudományi Társaság Kiadványai 212. MNYT, Budapest. 602–615.
- BAVIN, Edith L. – NAIGLES, Letitia R. (eds) 2015. *The Cambridge Handbook of Child Language*. Cambridge University Press, Cambridge.
- BALÁZS Boglárka – GÓSY Mária 1994. Hangazonosítás és -differenciálás adenotómia előtt és után. *Beszédgyógyítás* 3/2. 1–7.
- BALÁZS Boglárka – GÓSY Mária – SZABÓ Iván 1997. A G-O-H, a küszöbaudiogram és a tympanogram összefüggéseiről. *Fül-orr-gégegyógyászat* 43/3. 126–132.
- BISHARA, Laila – BEN-DAVID, Jonathan – PODOSHIN, Ludwig et al. 1999. Correlations between audiogram and objective hearing tests in sensorineural hearing loss. *International Tinnitus Journal* 5/2. 107–112.
- BOMBOLYA Mónika 2007. Hallássérült gyermekek beszédfeldolgozási folyamatai. In GÓSY Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kiadó, Budapest. 70–82.
- BÓNA Judit 2007. A beszédképzés és a beszédhallás összefüggése beszédhibásoknál (4–5 éves korban). In GÓSY Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kiadó, Budapest. 113–124.
- BOOTHROYD, Arthur 1997. Auditory development of the hearing child. *Scandinavian Audiology. Supplementum* 46/46. 9–16.
- BUNTA, Ferenc – BÓNA, Judit – GÓSY, Mária 2016. HU-LARSP: Assessing children's language skills in Hungarian. In FLETCHER, Paul – BALL, Martin J. – CRYSTAL, David (eds): *Profiling Grammar. More Languages of LARSP*. Multilingual Matters, Bristol–Buffalo–Toronto. 80–98.
- CSEPE Valéria 2007. A beszédészlelés kritikus kérdései és a beszédészlelés fejlődése „neuro”-nézetből. In GÓSY Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kiadó, Budapest. 18–42.
- FENT Zoltán 2007. A hallószerv, a hallás folyamata, zavarok. In GÓSY Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kiadó, Budapest. 43–56.
- FÓNAGY Iván – MAGDICS Klára 1960. Beszédsebesség, szólam, ritmusérzék. *Magyar Nyelv* 56. 450–458.
- GÓSY Mária 1989. *Beszédészlelés*. MTA Nyelvtudományi Intézet, Budapest.
- GÓSY Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- GÓSY Mária 1995/2006. *GMP-diagnosztika*. Nikol Kiadó, Budapest.



- GÓSY, Mária 2008. Synthesized speech used for the evaluation of children's hearing and speech perception. In GARDNER-BONNEAU, Daryle – BLANCHARD, Harry E. (eds): *Human Factors and Voice Interactive Systems*. Springer, New York. 127–139.
- GÓSY, Mária 2012. Patterns of verbal speech processing skills in 4 to 9-year-old children. In ILLÉS, Éva – EITLER, Tamás (eds): *Studies in applied linguistics in honor of Edit H. Kontra*. ELTE BTK, Budapest. 155–164.
- GÓSY, Mária – OLASZY, Gábor – HIRSCHBERG, Jenő – FARKAS, Zsolt. 1987. Phonetically based new method for audiometry: The G-O-H measuring system using synthetic speech. In *Proceedings of the XIth ICPHS 4*. University of Tallinn, Tallinn. 185–189.
- GÓSY Mária – HORVÁTH Viktória 2006. A percepció folyamatok összefüggései hatéveseknél. *Alkalmazott Nyelvtudomány* 6/1–2. 25–42.
- GÓSY Mária – HORVÁTH Viktória 2008. Kettős deficit: beszédhibás iskolások beszédfeldolgozása. *Gyógypedagógiai Szemle* 36/3. 177–193.
- GÓSY Mária – HORVÁTH Viktória 2011. Eltérő beszédpercepció működések az iskoláskor kezdetén. *Magyar Nyelvőr* 135/2. 161–180.
- GÓSY Mária – IMRE Angéla 2007. *Beszédpercepció fejlesztő modulok*. Nikol Kiadó, Budapest.
- GÓSY, Mária – KREPSZ, Valéria 2018. Evaluation of cognitive processes using synthesized words: Screening of hearing and global speech perception. *Acta Polytechnica Hungarica* 15/5. 31–45.
- GÓSY, Mária – BAHR, Ruth – GYARMATHY, Dorottya – BEKE, András 2018. Dichotic listening and sentence repetition performance in children with reading difficulties. *Clinical Linguistics and Phonetics* 32/9. 787–803.
- GÓSY Mária – GYARMATHY Dorottya 2018. A beszédindulás késésének hatása a beszédfeldolgozásra. *Beszédkutatás* 26/1. 186–219.
- GÓSY Mária – PREGITZER Márta 2019. Beszédészlelés és beszédmegértés ikreknél és nem ikreknél. *Anyanyelv-pedagógia* 12/3. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=795> (A letöltés ideje: 2020. április 18.)
- GYARMATHY Dorottya 2017. *Megakadásjelenségek a magyar spontán beszédben*. MTA Nyelvtudományi Intézet, Budapest.
- HAGOORT, Peter 2019. The meaning making mechanism(s) behind the eyes and between the ears. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 375/1791. 2019–2031.
- HAGOORT, Peter – BECKMAN, Christian F. 2019. Key issues and future directions: The neural architecture for language. In HAGOORT, Peter (ed.): *Human Language: From Genes and Brains to Behavior*. MIT Press, Cambridge, MA. 527–532.
- HAZAN, Valerie – FOURCIN, Adrian J. 1983. Interactive synthetic speech tests in the assessment of the perceptive abilities of hearing impaired children. *Speech, Hearing, and Language* 3. 41–57.

- HEALD, Shannon L. – NUSBAUM, Howard C. 2014. Speech perception as an active cognitive process. *Frontiers in Systems Neuroscience* 8/35. <https://doi.org/10.3389/fnsys.2014.00035>
- IMRE Angéla 2007. A beszédmegértés és az olvasás összefüggése. In GÓSY Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kiadó, Budapest. 186–204.
- JÓZSA Krisztián (szerk.) 2006. *Az olvasási képesség fejlődése és fejlesztése*. Dinasztia Tankönyvkiadó, Budapest.
- JÓZSA Krisztián – FAZEKASNÉ FENYVESI Margit – SZENCZI Beáta – SZABÓ Ákosné 2015. Tanulásban akadályozott és tipikusan fejlődő gyermekek szóolvasási készségének, szövegértésének és olvasási motivációjának fejlődése. In LÁNYINÉ ENGELMAYER Ágnes – GYÓRI Miklós (szerk.): *Gyógypedagógiai lélektan*. Akadémiai Kiadó, Budapest. 181–204.
- JUHÁSZ Ágnes (szerk.) 1999. *Logopédiai vizsgálatok kézikönyve*. Új Múza Kiadó, Piliscsaba.
- KEMALOĞLU, Yusuf K. – GÜNDÜZ, Bülent – GÖKMEN, Selda – YILMAZ, Metin 2005. Pure tone audiometry in children. *International Journal of Pediatric Otorhino-laryngology* 69/2. 209–214.
- KORHONEN, Johan – LINNANMÄKI, Karin – AUNIO, Pirjo 2014. Learning difficulties, academic well-being and educational dropout: A person-centred approach. *Learning and Individual Differences* 31. 1–10.
- LIBÁRDI Péter 2017. *A mentális lexikon többszemontú vizsgálata*. ELTE Eötvös Kiadó, Budapest.
- LILJENCRAITS, Johan C. 1968. The OVE III speech synthesizer. *IEEE Transactions on Audio and Electroacoustics* 16/1. 137–140.
- MARKÓ Alexandra 2007. A mondat- és szövegértés jellemzői és összefüggése 6–9 éves korban. In GÓSY Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kiadó, Budapest. 290–306.
- MARTIN, Randi C. – TAN, Yingying – NEWSOME, Mary R. – VU, Hoang 2017. Language and lexical processing. In DUNBAR, Robin (ed.): *Neuroscience and Biobehavioral Psychology*. Elsevier, Amsterdam. 427–452.
- OLASZY Gábor 1985. *A magyar beszéd leggyakoribb hangsorépítő elemeinek szerkezete és szintézise*. Nyelvtudományi Értekezések 121. Akadémiai Kiadó, Budapest.
- PISONI, David B. – REMEZ, Robert E. (eds) 2005. *The Handbook of Speech Perception*. Wiley–Blackwell, Hoboken.
- PLÉH Csaba 2013. A magyarázat és a megértés a szellemtudományos pszichológiában – egy évszázad elteltével. *Magyar Filozófiai Szemle* 57/2. 29–39.
- POLL, Gerard H. – MILLER, Carol A. 2013. Late talking, typical talking, and weak language skills at middle childhood. *Learning and Individual Differences* 26. 177–184.
- PYTEL József 1996. *Audiológia*. Victoria Kiadó, Budapest.
- RICE, Mabel L. – TAYLOR, Catherine L. – ZUBRICK, Stephen R. 2008. Language outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *Journal of Speech, Language, and Hearing Research* 51/2. 394–407.

ROSTA Katalin – SUCHNÉ RUMPLI Henriette 2007. Beszédfogyatékos gyermekek beszédpercepció vizsgálat. In GÓSY Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kiadó, Budapest. 137–148.

SZÁNTÓ Anna 2017. *Beszédfeldolgozási folyamatok változásai a fejlesztés függvényében*. ELTE Eötvös Kiadó, Budapest.

VÉRTEŠ O. András 1980. *A magyar leíró hangtan története az újgrammatikusokig*. Akadémiai Kiadó, Budapest.

### **KÖSZÖNETNYILVÁNÍTÁS**

A tanulmány megírását a Tématerületi Kiválósági Program támogatta.

# A diszfónia és automatikus felismerése

Tulics Miklós Gábor

BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM

## 1. A beszéd mint objektív biomarker

A **biomarkerek** vagy biológiai markerek olyan mérhető jelek, amelyek valamely biológiai állapotot jeleznek – jelezhetik többek között egy betegség jelenlétét. A biomarkerek nagyon változatosak lehetnek, ilyen például a gyulladást jelző CRP (C-reaktív fehérje) a vérben vagy specifikus antitest, amely fertőzés jelenlétét jelzi. A hangunk is olyan biomarker, amely (egyebek között) számos betegségre utalhat (STRIMBU–TAVEL 2010).

A beszéd az emberi kommunikáció legfontosabb módja, de sokkal több információt tartalmaz, mint a kimondott szavak jelentése (NÉMETH–OLASZY 2010). A beszéd nagyon sokat elárul a mentális és fizikai állapotunkról is. Ha egy ember szomorú, fáradt, megfázik stb., azonnal hallani lehet a hangján. A beszédprodukciónak összetett folyamat, amelynek eredményeként mechanikus hullám terjed egy rugalmas közegben, amely hangérzetet kelt az élőlényekben.

Amikor beszélünk, sok szervünk részt vesz a fonációs folyamatban (a fonáció a zöng kialakításának folyamata), tehát ha bármilyen változás lép fel ezekben a szervekben, akkor a létrehozott beszéd tulajdonságai is megváltoznak. Ha a változás hallható és látható (például a hullámalakban vagy a spektrogramban), akkor azt mérni lehet. Ez lehetséges a hang akusztikai tulajdonságainak vizsgálatával.

## 2. Beszédprodukción

A beszéd nyelvi jelentéssel bíró akusztikai termék, de olyan információkat is magában hordoz, amelyek nem kapcsolódnak ehhez a nyelvi jelentéshez (NÉMETH–OLASZY 2010). A fizikai jelet nagyon pontosan ellenőrzött élettani folyamatok hozzák létre. Nemcsak a beszédgenerálás összetett folyamat, de az előállított jel fizikai megjelenése és maga a beszéd rezgése is komplexitást és dinamikus változásokat mutat frekvenciában, időben és intenzitásban.

A beszéd nyelvi és nem nyelvi jellemzői egyszerre vannak jelen a beszédprodukción során keletkező komplex akusztikus jelben. A nyelvi jellemzők hordozzák a fogalmi tartalmat, míg a nem nyelvi jellemzők olyan információkat hordoznak, mint a beszélő neme, fizikai állapota (például fáradtság vagy betegség) vagy a beszélő érzelmi állapota (öröm, düh, ámulat, izgalom,

szomorúság, undor stb.). Ez a két komponens a beszédet kifejezőbbé és élénkebbé teszi, mint amilyen az írott szöveg.

A beszédprodukciónban részt vevő alapvető szerveket három területre lehet felosztani. Az első a szubglottális (a hangszalagok/gége alatti) tér, ahol a tüdő és a légcső található, majd következik a glottális tér a gégevel, végül a szupraglottális tér (más néven toldalékcső vagy artikulációs csatorna), ahol a garat-, a száj- és az orrüreg található. Minden szerv működését az agy irányítja. A levegő felfelé áramlik a tüdőből a légcsövön át a gégeig, amely a beszéd forrása, a hangprodukciónak elsődleges területe. Ha zöngés hangot ejtünk, akkor a hangszalagok kváziperiodikus nyitása és zárása kváziperiodikus hang kialakulását eredményezi. Zöngétlen hang esetén a levegő akadálytalanul halad át a gégen (a hangszalagok szűk lélegző vagy más néven fúvó állásban vannak). A levegő továbbhalad a garat-, majd a szájüregen (és az orrüregen), ahol a passzív artikulációs szervek (fogak, kemény szájpad) és az aktív artikulációs szervek (nyelv, nyelvcsap, ajkak, állkapocs) módosítják a hullámformát. Minden egyes beszédhang kiejtéséhez sajátos szervekkel kivitelezett mozgások összessége tartozik, amelyeket artikulációs konfigurációnak nevezünk.

Az emberek beszéd szervei kisebb-nagyobb mértékben különböznek egymástól, ezért az emberek által produkált beszéd akusztikailag eltérő (pl. egyéni artikulációs sajátosságok, hangszínezet). A beszéd akusztikai szerkezetét befolyásolja az is, hogy az artikuláció nem univerzális a beszélők között. Ezt hívjuk a beszélők közötti varianciának (vagy változatosságnak). A beszédnek a beszélőn belüli variabilitása pedig azt jelenti, hogy egy adott beszélő beszédprodukciónja sem állandó. A beszédproduktum függ a beszédhelyzettől és az adott személy fizikai, egészségi és érzelmi állapotától. A beszédet beszédhangok sorozataként írjuk le, holott a beszéd folyamatos és folytonosan változó jelenség, így valójában különböző artikulációs konfigurációk közötti átmenetek sorozata. Ez az úgynevezett koartikuláció, amelynek révén az egyes (absztrakt) beszédhangok kiejtését befolyásolják az őket megelőző és követő beszédhangok. Ez a hatás is hozzájárul a beszédhangok akusztikai jellemzőinek beszélőn belüli változékonyságához.

A kiejtett beszédhangok akusztikai jellemzőit objektív módon lehet mérni. A fent bemutatott variabilitás miatt az egyén beszédének akusztikai jellemzői nem állandóak, de ennek ellenére az egészséges emberek beszédének akusztikai jellemzői valamilyen mértékig konstansnak mondhatók (például ezért lehet felismerni egyes személyek hangját). Az emberi beszédészlelés az akusztikai variabilitás ellenére is biztosítja, hogy a fonológiai döntések ugyanazok maradnak, érteni tudjuk az elhangzott beszédet.

### 3. A diszfónia és kapcsolata a hang rekedtségével

Életének egy bizonyos pontján a lakosság csaknem egyharmadát érinti egy nagyon gyakori egészségügyi panasz, amit diszfóniának vagy kóros hangprodukciónak nevezünk (STACHLER et al. 2018; COHEN 2010; COHEN et al. 2012b; REITER et al. 2015). A diszfónia fogalmát gyakran

a rekedtséggel cserélik fel, viszont ez terminológiai pontatlan: a **rekedtség** a beteg megnevezése az általa érzékelt hangminőség-változásra, míg a **diszfónia** a hangképzési zavar meghatározására használatos klinikai terminus (JOHNS III et al. 2010).

A jelenség előfordulása széles körű változatosságot mutat, mivel bármilyen korú és nemű beteget érinthet, de magasabb számban van jelen az idősebb felnőttek és a kivételesen nagy hangigénnyel rendelkező emberek körében, mint az énekesek, színészek, tanárok stb. (DAVIDS et al. 2012; JONES et al. 1998; LONG et al. 1998; SMITH et al. 1998).

A statisztikai adatok szerint évente 13 felnőttből egyet valamilyen módon érintenek hangképzési problémák (BHATTACHARYYA 2014). A betegek jelentős része nem keres orvosi segítséget, annak ellenére, hogy felmerülő hangképzési problémáikat felismerik (BENNINGER et al. 1998; BHATTACHARYYA 2014; COHEN–DUPONT–COUREY 2006). Ugyanakkor egyes források szerint a diszfónia óriási negatív hatással van a termelékenységre és a gazdaságra azáltal, hogy a diszfóniában szenvedő betegek sűrűn veszik igénybe az egészségügyi szolgáltatásokat, és ezáltal gyakran hiányoznak a munkából (RAMIG–VERDOLINI 1998). A diszfóniát gyakran enyhe állapotok jellemzik, de néha a fennálló tünet ennél súlyosabb vagy progresszív (előrehaladó) állapotra utal, amely azonnali diagnózist és kezelést igényel.

STACHLER és munkatársai (2018) klinikai gyakorlati útmutatót dolgoztak ki a rekedtségre és a diszfóniára vonatkozóan, amelyben így határozzák meg a diszfóniát: megváltozott hangminőség, hangmagasság, hangerő vagy hangképzési nehézség, amely rontja a páciens kommunikációját a klinikus megállapítása szerint, és/vagy befolyásolja a páciens életminőségét. Ezzel szemben a rekedtséget a következőképpen határozzák meg: a páciensek beszámolóiban leírt, megváltozott hangminőségre utaló tünet (i. m. 3). A klinikai gyakorlati útmutatót megfogalmazó testület hipotézise, hogy a diszfónia sokféleképpen érinti az embereket. Az útmutató rendkívül hasznos minden orvos számára, aki diszfóniában szenvedő betegeket diagnosztizál és kezel. Az útmutató minden olyan esetben alkalmazható, amikor a diszfónia azonosítása, kezelése és monitorozása a cél.

STACHLER és munkatársai (2018) útmutatója sok fontos dologra hívja fel a figyelmet, többek között arra, hogy a diszfóniával és annak kialakulásának okaival kapcsolatos ismeretek hiánya megakadályozza a megfelelő ellátást. Például míg az öregedés természetes hangváltozásokat okoz az idősebb korcsoportban, néhány diszfóniás hangváltozás súlyos betegségek tünete lehet. Ezenkívül a szülők félreértelmezhetik vagy normálisnak tarthatják a rekedtséget a gyermekeknél. A tünetek ilyen téves értelmezése hátráltatja a megfelelő feltárást, diagnózist és néhány súlyos mögöttes egészségi állapot kezelését. A helytelen diagnosztizálás minimalizálása és az ellátás minőségének javítása érdekében elengedhetetlen az egészségre vonatkozó ismeretek frissítése a gondozási szakemberek körében.

Az útmutató segít a megelőzésben azzal, hogy megkönnyíti a felismerést és a diagnózist, ugyanakkor a diszfónia célzott kezelését is tárgyalja. Ezenkívül korszerűsíti a diszfónia kezelési lehetőségeit, a diszfóniában szenvedő különféle populációk igényei szerint. Dióhéjban az iránymutatás célja, hogy gazdagítsa a szakemberek ismereteit a diszfónia kezelését és megelőzését illetően, kiegészítve a diszfónia által érintett egyének megfelelő kezelési lehetőségeivel.

### 3.1. A diszfónia típusai

A hangképzési rendellenesség és a diszfónia fogalma nem azonos, sok forrás mégsem választja el élesen a kettőt. Míg hangképzési rendellenesség akkor fordul elő, amikor az egyén hangminősége, a hang magassága, illetve a hangossága nem felel meg az egyén életkorának, nemének vagy kulturális háttérének, addig a diszfónia összesíti a hangképzési rendellenességek hallási-percepció tüneteit (ASHA 2020). A hangképzési rendellenességgel küzdő emberek hangjának észlelési minősége a betegség típusától és súlyosságától függően változik. A hangképzési rendellenesség súlyosságát nem mindig lehet egyedül a hallási-perceptuális hangminőség alapján meghatározni (ASHA 2020; BEHLAU–MADAZIO–OLIVEIRA 2015; HIRSCHBERG–HACKI–MÉSZÁROS 2013).

Az American Speech-Language-Hearing Association szerint a hangképzési rendellenességek két csoportra oszthatók: **organikus** (vagy szervi) **hangképzési rendellenességek** és **funkcionális hang rendellenességek** (ASHA 2020). Az organikus rendellenességek két csoportba sorolhatók: neurogén és strukturális rendellenességekre. A **neurogén hang rendellenesség** az idegrendszer problémájának lehet a következménye. Ez magában foglal olyan hangproblémákat, amelyeket a gége izmainak rendellenes kontrollja, koordinációja vagy erejének hiánya okozhat egy neurológiai betegség miatt. Ilyen betegség például a stroke, a Parkinson-kór, a szklerózis multiplex, a myasthenia gravis, az amiotróf laterális szklerózis és a hangszalagbénulás (rekurrens parézis, angolul: *recurrent paresis*, RP). A **strukturális rendellenesség** magában foglal valamilyen fiziológiai rendellenességet, például a hangszalagok szöveteinek elváltozásait, amilyen az ödéma vagy a hangszalagon lévő csomók, polipok. További strukturális rendellenességekhez tartozó betegség lehet a gastroözofoageális reflux betegség (GERD) és a ciszta (HIRSCHBERG–HACKI–MÉSZÁROS 2013; JANI et al. 2008).

Ezekből adódóan kijelenthetjük, hogy a diszfónia etiológiailag, azaz oki hátterét tekintve is két alapvető kategóriára osztható: funkcionális diszfóniára (FD) és organikus diszfóniára (OD) (ASHA 2020; BEHLAU–MADAZIO–OLIVEIRA 2015; HIRSCHBERG–HACKI–MÉSZÁROS 2013). Az **organikus diszfóniák** a hanghasználat nem összefüggő tényezők következményei. A **funkcionális diszfónia** nem fiziológiai elváltozásból eredő hangprobléma, hanem a rossz hangképzési technika és/vagy izomegyensúly-probléma, pszichoemotikus problémák következménye.

Az organikus diszfóniát számos tényező okozhatja. Például egy gégegyulladásos egyén diszfóniás hangot produkálhat. Az akut gégegyulladást okozhatja vírusos vagy bakteriális fertőzés, a krónikus gégegyulladás lehet erős dohányzás vagy GERD következménye. A hangképző szervek traumát szenvedhetnek műtét vagy intubáció következményeként, ami ugyancsak diszfónikus hangot eredményezhet. A funkcionális diszfónia pedig a hangképzési mechanizmus helytelen vagy nem hatékony használatának eredménye.

A hangképzési rendellenességek osztályozásának ezt a hagyományos felosztását sokan bírálják, mivel bizonytalannak gondolják. A foniatér szakemberek és a logopédusok nem rekedt hangokkal foglalkoznak, hanem kóros hangképzéssel. Ezért a rekedt hang kialakulásának tanulmányozása inkább fizikai-akusztikai feladat (TITZE 1994). Az organikus elváltozások diszfunkcionális működést is okoznak, tehát vannak organikus eredetű és nem organikus eredetű diszfunkciók.

Egyes szakemberek kritikával illetik, ha a nem organikus diszfunkciókat funkcionális problémáknak nevezik. A nem organikus diszfónia megfelelője nem a funkcionális, hanem a regulatív (rendellenes szabályozásból származó) diszfónia (HIRSCHBERG–HACKI–MÉSZÁROS 2013).

Egy friss tanulmány javasolja a *funkcionális diszfónia* kifejezés helyettesítését a *malregulatív diszfónia* kifejezéssel (HACKI–MOERMAN–RUBIN 2020). Azzal érvelnek, hogy a *funkcionális* kifejezés zavaros, és nem egyértelmű a mindennapi klinikai használat során.

A diszfóniában szenvedő betegek tünetei összességében megváltozott hangminőséget mutathatnak, ami rekedtséggel és torokfájással vagy a torok kiszáradásával jár. Egyéb kapcsolódó tünetek lehetnek például az étvágytalanság, súlyvesztés, véres köhögés, savas reflux vagy gyomorégés, nem anginális (szívkoszorút érintő) mellkasi fájdalom és nyelési nehézség (HIRSCHBERG–HACKI–MÉSZÁROS 2013; STACHLER et al. 2018).

Bárkinek, akinek rekedt hangja van négy hétig vagy ennél hosszabb ideig, javasolt orvoshoz fordulnia. Minden diszfóniás hangot produkáló páciens állapotának megvan a maga specifikus kezelési módja, a kezelő fül-orr-gégésznek vagy logopédus szakembernek egyénre szabott kezelést kell biztosítania (HIRSCHBERG–HACKI–MÉSZÁROS 2013; RAMIG–VERDOLINI 1998; STACHLER et al. 2018).

### 3.2. A diszfónia gyakorisága

Egy 2001-es amerikai elemzés alapján a diszfónia előfordulása 0,98% (536 943 azonosított diszfóniával rendelkező beteg 55 000 000 emberből). A korábbi vizsgálatok eredményei szerint magasabb arányt állapítottak meg a nők körében (1,2% a nők, 0,7% a férfiak esetében), valamint a 70 évnél idősebbek körében (2,5%-os arány a 70 év fölötti korcsoportban, 0,6%–1,8% az összes többi korcsoportban) (COYLE–WEINRICH–STEMPLE 2001; ROY et al. 2005; ROY et al. 2016; TITZE–LEMKE–MONTEQUIN 1997). A *Betegségek Nemzetközi Osztályozása* (BNO; *International Classification of Diseases, Ninth Revision*) szerint az egészségügyi szakemberek leggyakoribb diszfóniával kapcsolatos diagnózisai az akut gégegyulladás, a nem specifikus diszfónia, a hangszalag-elváltozások (pl. ciszták, polipok, csomók) és a krónikus gégegyulladás. A diszfóniával kapcsolatos állapotok reális előfordulásának aránya valószínűleg magasabb, mert a hangelváltozásoktól szenvedő betegek többsége nem fordul szakemberhez, különösen, ha a tünetek átmenetiek, és felső légúti fertőzéshez kapcsolódnak (ROY et al. 2005). Egy korábbi, Utah és Iowa államokban végzett vizsgálat 29,9%-os hangrendellenességi kumulatív kockázatot állapított meg 65 éves kor előtt olyan felnőttek körében, akik nem igényeltek orvosi kezelést. A kumulatív kockázat egy bizonyos betegség kialakulásának esélye, ismert kockázati tényezőkkel nem rendelkező személyek esetében.



### 3.3. Költségek

A diszfóniás betegek kezelési költségei jelentősek. Az Egyesült Államokban széles körben végzett adminisztratívadatbázis-tanulmány szerint a diszfóniás kezelések közvetlen költségeit évente betegenként átlagosan 577–953 dollárra becsülik (COHEN et al. 2012b). Ha a becslések szerinti 5,2 millió diszfóniában szenvedő beteg évente kezelést igényelne, a teljes közvetlen egészségügyi ellátás költsége kb. 13,5 milliárd dollár lenne. Összehasonlításképpen, ezek a kiadások hasonló mértékűek a következő betegségekre fordított összegekhez: krónikus obstruktív tüdőbetegség (COPD), asztma, cukorbetegség és allergiás nátha.

### 3.4. Az életminőségre gyakorolt hatás

A diszfónia hatásai lényegesen befolyásolhatják az életminőséget. A diszfónia okozta korlátozások gyakran mellékhatásokkal járnak, mint a depresszió, társadalmi elszigeteltség, szorongás, elmulasztott munka stb. (FRANCIS 2014; MIRZA et al. 2003; ROY et al. 2005). Több kutatás is megállapította, hogy a hangrendellenességgel küzdő betegek életminőségbeli és termelékenységi veszteségei hasonlóak más légzőszervi betegségekben szenvedőkéhez, mint amilyen az asztma, a krónikus obstruktív tüdőbetegség, az akut koszorúér szindróma vagy a depresszió (BENNINGER et al. 1998; COHEN–DUPONT–COUREY 2006). Azokban az esetekben, amikor a betegek a betegség súlyosabb típusaiban szenvednek (például egyoldali hangszalagbénulásban), megfigyelhető az életminőség és a termelékenység még jelentősebb romlása is.

### 3.5. A diszfónia mint valamely alapbetegség tünete

A diszfónia számos betegség gyakori tünete. Fej- és nyakrák esetén a diszfónia kísérőtünet lehet, ezért elengedhetetlen a diszfóniának és korai jeleinek felismerése. A gége pontatlan vizsgálata késleltetheti a pontos diagnózist, így ez a későbbiekben a betegség egy súlyosabb stádiumát, majd invazívabb kezelést és magasabb halálozási arányt eredményezhet (CHEN et al. 2007). Neurológiai elváltozások is idézhetnek elő diszfóniát. A teljesség igénye nélkül néhány példa ilyen esetekre a hangszalagbénulás, a spazmodikus diszfónia (SD), a Parkinson-kór, az amiotróf laterális szklerózis, a szklerózis multiplex, egyes reumatológiai autoimmun betegségek (pl. reumás ízületi gyulladás, Sjögren-szindróma, szarkoidózis, amiloidózis), bizonyos mozgásszervi betegségek (pl. izomfeszültség-diszfónia, tisztázatlan eredetű reumás kötőszöveti és izomfájdalom vagy *fibromyalgia*, axiális nyaki fájdalom vagy *cervicalgia*), illetve pszichológiai betegségek (funkcionális hangproblémák).

A diszfónia előfordulási aránya ilyen körülmények között természetesen változhat, mert például szinte minden spazmodikus diszfóniában vagy gégedisztóniában (a gége izmainak túlzott akaratlan összehúzódásának rendellenességében) szenvedő beteg diszfóniában is szenved, de nem minden refluxban szenvedő betegnek van diszfóniája.

### 3.6. A diszfónia és az életkor

A hangképzési rendellenességek minden korosztályt érintenek, de nagyobb a kockázat a gyermekek és az idősebbek körében (a 65 évnél idősebbek között). Becslések szerint a gyermekek 23,4%-ának van diszfóniája gyermekkoruk egy bizonyos szakaszában, továbbá általában a fiúk esetében gyakoribb, mint a lányoknál, illetve a fiúk csoportján belül a 8 és 14 év közötti fiúk körében magasabb az előfordulási arány (BHATTYCHARYA 2015; CARDING et al. 2006; DUFF–PROCTOR–YAIRI 2004; SILVERMAN–ZIMMER 1975). A diszfónia előfordulása szintén szignifikánsan magasabb korfüggő gége mutációban szenvedő idősebb felnőttek között (*presbylarynx*) (COYLE–WEINRICH–STEMPLE 2001).

Egy biztosítási adatbázist elemző tanulmány szerint a diszfónia kezelését igénylő idősök aránya 1,3% volt a 60–69 éves korosztály körében, és 2,5% volt a 70 évnél idősebb betegek körében (COHEN et al. 2012b; ROY et al. 2016). Ebben a csoportban a leggyakoribb diagnózisok az akut krónikus gégegyulladás, a nem specifikus diszfónia és a gégeelváltozások voltak. Egy korábbi felmérés, amely kezelést nem kereső idősebb önkénteseket vizsgált, kimutatta, hogy 47%-uk szenvedett életében hangképzési rendellenességben, 29%-uknál pedig aktív diszfóniát tapasztaltak (ROY et al. 2007).

### 3.7. A diszfónia és a foglalkozás típusának kapcsolata

A hang fokozott terhelését követelő szakmák körében a diszfónia sokkal nagyobb eséllyel jelenik meg. A legveszélyeztetettebb helyzetben az énekesek és a szórakoztatóiparban dolgozók, a jogászok, a tanárok, az edzők és a papok vannak, illetve bárki, akinek a szakmájában kiemelkedő hangkapacitásra van szükség (GUSS et al. 2014; FORTES et al. 2007; ROSA–BEHLAU 2017; ROY et al. 2004; SMITH et al. 1998; STACHLER et al. 2018; VERDOLINI–RAMIG 2001).

Ebből a szempontból természetes, hogy a hangbetegségek, különösen a diszfónia nagyban befolyásolják az ember munkaképességét (ISETTI–MEYER 2014). Az Egyesült Államokban körülbelül 28 millió munkavállaló küzd napi szinten hangproblémákkal (VERDOLINI–RAMIG 2001). A megkérdezettek körében a válaszadóknak átlagosan a 7,2%-a hiányzott hangprobléma miatt a munkából egy vagy több napig (ROY et al. 2005). A tanárok 20%-a hiányzik a munkából diszfónia miatt, ami csak ezen a szakterületen 2,5 milliárd dollár veszteséget okoz évente az Egyesült Államoknak (VERDOLINI–RAMIG 2001).

### 3.8. Környezeti tényezők

A környezeti feltételek is befolyásolhatják a beszédhangot, például a háttérzaj, a rossz levegőminőség és a száraz környezet (HAZLETT–DUFFY–MOORHEAD 2011; JANI et al. 2008; MARTINS et al. 2014). Magas szintű környezeti vagy munkahelyi irritációk szintén növelhetik a diszfónia kialakulásának valószínűségét, mint például a vegyi anyagok, füst, porszemcsék és szennyezőanyagok.

Egy tanulmány tíz munkást vizsgált, akik a World Trade Center katasztrófája után részt vettek a mentésben. A mentőmunkások megváltozott és rekedtes hangja összefüggött a helyszínen jelen lévő irritáló anyagok nagy mennyiségével (DE LA HOZ et al. 2008).

Ugyanakkor egyéb tanulmányok azt is kimutatták, hogy a környezeti párasítás hatékony a gége felületi kiszáradása ellen, és ez segíthet a negatív hangváltozások megelőzésében vagy csökkentésében (LEVENDOSKI–SUNDARRAJA–SIVASANKAR 2014).

### 3.9. A rekedtség és a hang percepciók alapú minősítése

Hagyományosan a hang minőségének érzékeléséhez a hallgató számtalan tulajdonságot társít, amelyek a helyi nyelvi hagyományok hatására alakultak ki (HIRSCHBERG–HACKI–MÉSZÁROS 2013). A hallgatók gyakran más érzékek által közvetített benyomásokra támaszkodnak, és olyan érzeteket társítanak a hangminőséghez, mint sötét, világos, fedett, tompa, éles stb. A tudomány igyekszik hangkategóriákat bevezetni, amelyekhez hozzárendelhető a hangképzés fiziológiai és akusztikai valósága. Azonban még egy fejlett akusztikai elemzés sem képes meghatározni és osztályozni a hallásunk által érzékelt hangjellemzőket teljes részletességgel, ezért (még) szükségünk van a fül megítélésére. Az egyik legfontosabb területe a hallási megítélésnek a rekedtség, ami a hangminőség eltérését jelzi a megszokottól, amint azt a korábbi fejezetekben definiáltuk.

A hang megítélése szempontjából döntő fontosságú, hogy a kategóriák száma ne legyen nagy, azonban különüljenek el világosan.

A rekedtség, mint a hangminőség patológiás kategóriája, élettanilag visszavezethető az érdességre (a hangszalagok rezgésének szabálytalansága) és a levegősségre (az áramlási turbulenciák), amelyek a hangszalagok nem megfelelő lezárásából erednek. A hangminőség percepciók értékelésében, minősítésében többféle mérőszámot alkalmaznak, amelyek alapvetően ezeknek a fiziológiai jellemzőknek a hallható következményeire építenek.

A rekedtség percepciók „mérés”-ére használatos egyik skála a HIRANO 1981-es tanulmányában bevezetett GRBAS, amely az angolszász területeken és Japánban terjedt el. A GRBAS skálában a G (*grade/severity*) a súlyosság fokát jelenti, ami a rekedtség általános megítélésére vonatkozik. Az R (*roughness*) az érdességet, a B (*breathiness*) a hang levegősségét, az A (*asthenia*) a hang gyengességét, az S (*strain*) pedig a hang megterheltségét jelenti (HIRANO 1981; OMORI 2011). A gyengesség és a megterhelés kategóriái a beteg általános tónusát írják le. A kategóriákat 0-tól 3-ig kell osztályozni, a G kategóriában a 0 érték a rekedtség hiányát jelenti (normális beszédhang), az 1-es érték enyhe fokú, a 2-es közepes fokú és a 3-as súlyos fokú rekedtségre utal.

A klinikai tapasztalatok alapján a GRBAS skála nagyon jól használható, annak ellenére, hogy az utolsó két kategóriát hallás szerint néha nehéz megkülönböztetni.

A német nyelvterületen az RBH-skálát alkalmazzák széles körben a hangminőség megítélésében. Az RBH-skála (R = *Rauchigkeit*, H = *Heiserkeit*, B = *Behauchtheit*) megadja

a diszfónia súlyosságát, ahol az R jelentése érdesség, a B jelentése levegősség, a H pedig az általános rekedtség. A H kategória mértéke nem lehet kisebb, mint a másik két kategória legmagasabb értéke. Például, ha  $B = 3$  és  $R = 2$ , akkor  $H = 3$ , és nem lehet 2 vagy 1. Az egészséges hang kódja R0B0H0; a maximális H, illetve az RBH értéke 3, tehát egy minden szempontból súlyos diszfóniával rendelkező hang kódja R3B3H3. ПТОК és munkatársai (2006) kimutatták, hogy az RBH-skála alkalmazása megfelel klinikai célokra. Ha a H értéke 0, akkor azt mondhatjuk, hogy a beteg hangja nem rekedt. Ha a  $H = 1$ , akkor enyhe rekedtségként, ha a  $H = 2$ , az mérsékeltként, míg ha a  $H = 3$ , az súlyos rekedtségként értelmezhető. A rekedtes hang súlyos alapbetegségre utalhat.

### 3.10. A diszfónia: összegzés

A fenti alfejezetekből arra lehet következtetni, hogy a diszfónia a népesség egyharmadát érinti életük során. A diszfónia nemcsak az időseket érinti, hanem a gyerekeket is. Olyan emberek, akik munkájuk során sokat használják a hangjukat (tanárok, énekesek stb.), hajlamosabbak a diszfóniára, és ez az állapot befolyásolhatja a munkaképességüket. A diszfóniát nemcsak az ember hangjának kiemelt használata okozhatja, hanem sok más ok is előidézheti, mint például a folyamatosan növekvő környezeti zaj, a környezeti szennyeződés vagy a száraz levegő. A diszfónia kezelési költsége jelentős terhet jelent az egészségügyi ellátás számára.

A diszfónia számos súlyos és kevésbé súlyos betegség indikátora lehet. Például a rekedtség (így a diszfónia is) életveszélyes betegség jele is lehet, mint amilyen a gégerák. Ezenkívül a rekedtség a hangszalag bénulását is jelezheti, amelyet gyakran tüdőrák okoz. A diszfónia általában rontja az emberek életminőségét is, magányhoz és depresszióhoz vezethet. Összefoglalva, a korai diagnózis és a terápia életmentő lehet.

A diszfóniával és annak okaival kapcsolatos ismeretek és a tudatosság hiánya szintén potenciálisan akadályozhatja a megfelelő egészségügyi ellátáshoz való hozzáférést. Az idősebb felnőttek tapasztalhatnak hangváltozást az öregedés részeként, de egyes tünetek súlyosabb alapbetegségre utalhatnak. Ezenkívül a szülők félreérthetik gyermekük rekedtségét, és azt természetesnek, normálisnak tekinthetik. Ez a feltételezés lelassíthatja a gyermek egészségügyi állapotának feltárását, diagnosztizálását és kezelését. Ezért van szükség a téma megismertetésére a pedagógusok és az egészségügyi szakemberek körében.

Ezek a tények mind arra utalnak, hogy az orvosoknak eszközöket kell biztosítani a diagnózis felállításához annak érdekében, hogy a diszfónia könnyebben és gyorsabban diagnosztizálható legyen a korai szakaszban. Célszerű volna elkülöníteni a diszfónia típusait a lehető leghamarabb, ezáltal felgyorsítva az orvosi eljárást. Továbbá, a diszfónia súlyosságának automatikus meghatározása elősegítené a beteg állapotának jobb feltérképezését.

A gyerekekkel dolgozó pedagógusoknak eszközökre lenne szükségük a rekedt hangú gyerekek szűrésére. Ezáltal a gyerekek minél hamarabb hozzáférhetnének a szakemberek, például logopédusok segítségéhez. Egy folyamatos beszédet használó, hangalapú diagnosztikai

támogatórendszer alkalmas lenne a diszfónia kimutatására, és bizonyos esetekben akár életeket is menthetne.

## 4. Szelíd bevezető a gépi tanulásba

A mesterséges intelligencia észrevétlenül beopta magát a hétköznapjainkba. 1997-ben sokként hatott a világra, amikor az IBM Deep Blue-ra keresztelt szuperszámítógépe egy játszmaiban legyőzte az akkori világbajnok sakkozót, Garri Kaszparovot (PANDOLFINI 1997). Azóta nem ér minket meglepetésként, ha az okostelefonunk a kameraképünk alapján prediktálja a korunkat, ha cicás maszkot tehetünk a fényképeinkre, ha bediktálhatjuk az útvonaltervező applikációnknak az úti célunkat, és az megtervezi a leghatékonyabb útvonalat. Nem lepődünk meg a raktárakban, gyártósorokon dolgozó robotokon, az okosporszívók népszerűségén, ahogyan az is egyre kevésbé sokkoló, hogy a Facebook a megadott adataink alapján célzott reklámokat küld a hírfolyamunkra (TUTTLE 2018). A Boston Dynamics 2020 végén bulizó fiatalokat megszegeyénítő táncudású robotokról töltött fel videót, jelezvén, hogy technológiájuk egyre kiforrottabb (BOSTON DYNAMICS 2020). Manapság a telefonunkon sokkal jobb sakkprogram fut, mint ami anno legyőzte Garri Kaszparovot, sőt a Google DeepMind által fejlesztett AlphaGo 2016-ban legyőzte a tizennyolcszoros Go-világbajnokot, Lee Sedolt (GIBNEY 2017). Ez korábban elképzelhetetlennek tűnt. Ugyanezt a technológiát felhasználva a DeepMind AlphaFold nevű algoritmus 2020-ban pontosan előre tudta jelezni a fehérjék térbeli konformációját a Critical Assessment of Structure Prediction (CASP) verseny keretében (WEI 2019). Ez a rendkívüli eredmény felgyorsíthatja a folyamatot, amelynek során megérthetjük a sejtek építőelemeit, és felgyorsíthatja a gyógyszerek fejlesztését is. Egyértelmű, hogy a mesterséges intelligencia szebbé, jobbá és élhetőbbé változtathatja a világunkat, ha azt jól használjuk.

A mesterséges intelligencia legismertebb ága a **gépi tanulás**, amelynek célja, hogy a rendszerek minimális emberi beavatkozással tanuljanak az adatokból, azonosítsanak mintákat, valamint hozzanak döntéseket (ALPAYDIN 2020). Gépi tanulás során az adatokat tanító- és tesztelőadatokra osztjuk, majd modellt építünk. A modellt a tanítóadatok alapján tanítjuk, majd a tesztelőadatokon értékeljük ki.

A gépi tanulásnak három legelterjedtebb módszere a felügyelt, a felügyelet nélküli és a megerősítéses tanulás.

A **felügyelt tanulás** során a tanító és tesztelő adathalmaz mintáihoz címkék is tartoznak. A címkék lehetnek kategorikusak (pl. cica vs. kutya, beteg vs. egészséges), ordinálisak (pl. enyhe, mérsékelt, súlyos), vagy akár folytonos értékűek is (a személy kora, magassága, fizetése). Tipikusan felügyelt tanulási feladatok az osztályozási és regressziós problémák. **Osztályozás** során kategorikus címkék predikciója a cél, míg **regresszió** során folytonos értékeket akarunk prediktálni az adataink segítségével. Ordinális címkék esetén választhatunk, hogy osztályozással vagy

regresszióval kezeljük-e a problémát. A tanulás során olyan adatokkal rendelkezünk, amelyeknek ismerjük a helyes címkéit. Ha személyek fizetésének predikciója a cél, a minták a különböző személyek, akikről információink vannak. A jellemzők lehetnek például a személyek kora, neme, végzettsége, családi állapota, a címke pedig a személy fizetése. Azt feltételezzük, hogy a minta jellemzői leírják a címkét; a gépi tanulási algoritmus feladata megtalálni a kapcsolatot a minták és a címkék között a jellemzőik segítségével. Ezt nevezzük modellépítésnek, amelynek a segítségével új mintákon pontos predikciót végezhetünk. Egy mintához a legtöbb esetben egy címke tartozik, de tartozhat egyszerre több is, amit a gépi tanulási feladat során prediktálni szeretnénk. Tehát a tanítóadat-halmazban lévő (jellemzőkkel leírt) mintákhoz helyes címkéket rendelünk, modellt építünk, majd az ugyancsak helyes címkékkel ellátott tesztalmazon predikciót végzünk, és összehasonlítjuk a modell által jóslt címkéket a valós címkéinkkel, ezáltal megtudhatjuk, mennyire jó a modellünk. A modellt utána új minták során használhatjuk fel predikcióra.

**Felügyelet nélküli tanulás** esetében az adatbázisban lévő mintákhoz nincsenek címkék hozzárendelve. Ilyenkor a gépi tanulási algoritmus mintázatokat keres az adatokban, amelyek az adatok jobb megértéséhez vezethetnek. Az algoritmus az adatokban lévő hasonlóságok alapján címkéket rendel a mintákhoz, amit majd a felhasználónak kell értelmezni. Ezeket a generált címkéket ilyenkor klasztereknek nevezzük, így a felügyelet nélküli tanulási problémákat leginkább klaszterezési problémának nevezzük. A pénzügyi, valamint a szociális médiával foglalkozó cégek tipikusan klaszterezési problémaként tekintenek a felhasználóik profilozására.

**Megerősítéssel tanulás** esetén az algoritmus a kijelölt feladatban nyújtott teljesítményével kapcsolatos visszacsatolások által tanul. A visszacsatolás segít eldönteni, hogy az előző döntés helyes vagy helytelen volt-e, esetleg semleges. Így a gépi tanulási algoritmus a hibáiból tanul. Ezt a módszert abban az esetben használjuk, amikor több kisebb, emberi beavatkozás nélküli döntést kell hozni. Tipikusan ilyen feladat lehet egy önvezető autó algoritmusának fejlesztése. Ez a megközelítés leginkább ahhoz hasonlít, amikor például egy kutya egy általunk kívánt magatartást sajátít el jutalomfalatért cserébe.

A technológia fejlődésével a beszéddel foglalkozó kutatóknak lehetőségük lett a beszédből olyan akusztikai jellemzőket kinyerniük, amelyek hangképzési rendellenességek leírásával hozhatók kapcsolatba. Ezeket az akusztikai jellemzőket fel lehet használni gépi tanuló algoritmusok tanítására. Az ötlet az, hogy ugyanazokat az akusztikai jellemzőket egészséges emberek, valamint hangképzési rendellenességekkel rendelkező emberek hangmintáiból nyerik ki, majd egy gépi tanulási algoritmus megtalálhatja a kapcsolatot a minták és a címkék között. Ebben az esetben a minták a beszédminták, amikből akusztikai jellemzők lettek kinyerve, a címke, hogy az adott beszédminta egészséges vagy diszfóniás személytől származik-e. Ugyanezzel az ötlettel megállapítható a rekedtség súlyossága, ez esetben regressziós módszerrel címkézzük a rekedtség súlyosságát. Osztályozásra és regresszióra használható gépi tanuló eljárások lehetnek például a szupportvektorgépek (*Support Vector Machines, SVM*), a döntési fák (*Decision Trees*), a véletlen erdők (*Random Forest*) vagy a nagy népszerűségnek örvendő mély neurális hálózatok (*Deep Neural Networks, DNN*). A következő fejezetek a témában született kutatásokat foglalják össze.

## 5. A diszfónia automatikus osztályozása

A bináris osztályozásnak számos megközelítése van az egészséges és a valamilyen rendellenességgel rendelkező hangok automatikus szétválasztására. Az első kérdés az, hogy melyik, a hanganyagból számolt akusztikai jellemzőket használjuk fel a felülyelt osztályozási modelünk tanítására. Továbbá, kitartott (önmagukban hosszan ejtett) magánhangzókat vagy folyamatos beszédet használunk az osztályozásra?

A kutatók ígéretes eredményeket értek el az egészséges beszéd és a valamilyen rendellenesség által érintett beszéd szétválasztása során, kitartott magánhangzókat felhasználva (AL-NASHERI 2017; OROZCO-ARROYAVE et al. 2015; TEIXEIRA-FERNANDES-ALVES 2017; ZHANG-JIANG 2008). A kitartott magánhangzók használatának megvannak az előnyei, mivel könnyen használhatók, hiszen nincs szükség erőforrás-igényes és nyelvfüggő szegmentálásra. A kutatók jelentős része azonban folyamatos beszédet használ kutatásai során (CORDEIRO-MENESES-FONSECA 2015; GUEDES et al. 2019; VICSÍ-IMRE-MÉSZÁROS 2011). VICSÍ, IMRE és MÉSZÁROS (2011) is rámutattak, hogy magasabb osztályozási pontosságokat lehet elérni folyamatos beszédet használva. A folyamatos beszédnek számos előnye van a kitartott magánhangzók elemzéséhez képest, mivel alapvető frekvenciaváltozásokat, szüneteket tartalmaz, és lehetővé teszi, hogy több fonémát és fonémakapcsolatot lehessen elemezni. A kutatási eredmények várhatóan jobban alkalmazhatóak a gyakorlati munkában, mivel folyamatos beszédet használunk a mindennapokban is. Így a hangprodukciónak a természetesebb módja vizsgálható.

Számos hangképzési rendellenességet tartalmazó beszédatadtbázis létezik kutatási célokra, amelyek kitartott magánhangzókat tartalmaznak, ilyen például az arab nyelvű Arabic Voice Pathology Database (AVPD) (MESSALAM et al. 2017), a német nyelvű Saarbrücken Voice Database (SVD) (MARTÍNEZ et al. 2012) vagy az angol Massachusetts Eye and Ear Infirmary (MEEI) beszédatadtbázis (DIBAZAR-NARAYANAN-BERGER 2002). Folyamatos beszédet tartalmazó beszédatadtbázisokat azonban nem könnyű találni. Olyan gondosan tervezett magyar nyelvű diszfóniás beszédatadtbázisra is szükség van még, amely folyamatos beszédet tartalmaz. Érdeemes megjegyezni, hogy a Parkinson-kór diszfónikus hanghoz vezethet (de nem feltétlenül rekedt hanghoz), ezért vannak olyan tanulmányok, amelyekben a diszfóniás kutatásokban használatos akusztikai jellemzőket hasznosítják a Parkinson-kór felismerésében. Ilyen akusztikai jellemzők például a *jitter* /dʒit:er/ (a zöngé periódusidejének ingadozása), a *shimmer* /ʃim:er/ (a zöngé periódusonkénti amplitúdóingadozása), a harmonikus-zaj arány (*Harmonics-to-Noise Ratio*, HNR), a beszédjelenben lévő harmonikus komponensek és zajkomponensek arányának jellemzője) vagy a zaj-harmonikus arány (*Noise-to-Harmonics Ratio*, NHR). Ezeket az akusztikai jellemzőket a diszfónia mértékeként is számontartják, az egyéb akusztikai jellemzők mellett (LAHMIRI-DAWSON-SHMUEL 2018).

További kérdés lehet, hogy milyen jellemzőket számoljunk a hanganyagokból, amelyeket aztán az osztályozó algoritmus bemenetére illeszthetünk egy vektor formájában (bemeneti

vektor). A bemeneti vektor felépülhet az előbb említett akusztikai jellemzőkből vagy más előfeldolgozási eljárással kinyert jellemzők alapján, például  $x$ -vektorokból és  $i$ -vektorokból (BOTELHO et al. 2020), glottális jellemzőkből (GIDAYE et al. 2020), autokorrelációs értékekből és különböző frekvenciasávokban mért entrópiajellemzőkből (AL-NASHERI et al. 2017), automatikus beszédfelismerők (*Automatic Speech Recognition, ASR*) kimeneteként kapott fonéma-valószínűségekből (LIU et al. 2017) vagy ezek kombinációjából.

A legmegfelelőbb osztályozási algoritmus kiválasztása újabb kihívást jelent. A mély neurális hálózatok tanítására általában nagyok sok adatra van szükség, viszont találunk a szakirodalomban példát kisebb adathalmazokon sikeresen használt mély neurális hálózatra is. Tipikusan ilyenek az orvosi kutatásokban, esetünkben az egészséges és a hangképzési rendellenességgel rendelkező személyek hangjának szétválasztására, alkalmazott neurális hálók (FANG et al. 2019; GUEDES et al. 2019; RUEDA–KRISHNAN 2019; WU et al. 2018). Kisebb adatbázisokban azonban érdemes lehet kipróbálni más osztályozókat, amelyek jó általánosítási képességgel rendelkeznek kisebb mennyiségű adathalmazokon, ilyen például a szupportvektorgép (*Support Vector Machine, SVM*) (AL-NASHERI et al. 2017; BOTHELO et al. 2020; CORDEIRO–MENESES–FONSECA 2015; VICSI–IMRE–MÉSZÁROS 2011).

A kutatások főként az egészséges és a valamely rendellenességgel jellemezhető beszéd közötti bináris osztályozásra összpontosítanak (GIDAYE et al. 2018). Az osztályozási eredmények nagymértékben függenek a kutatásokban vizsgált diszfónia típusától, a beszédatadabázis méretétől, a felhasznált hanganyag típusától (kitartott magánhangzók vagy folyamatos beszéd), a választott előfeldolgozási módszerektől és az alkalmazott osztályozótól. A kitartott magánhangzók használata általában nagyobb pontossághoz vezet. ALI, TALHA és ALSULAIMAN (2017) tanulmányában a kutatók a hangképzési rendellenességgel rendelkező személyek hangjait is tartalmazó MEEI-hangadatbázist használták fel egészségügyi szoftverük tesztelésére, melynek célja hangképzési rendellenességek felismerése volt nem periodikus beszédjelben. Az osztályozást SVM-mel végezték, és a maximálisan elért pontosság (a helyesen prediktált minták száma osztva a kísérletben felhasznált összes minta számával) 96,21% volt. A kutatásban diszfóniás betegek és egészséges emberek kitartott magánhangzókat tartalmazó hanganyagát használták fel. AL-NASHERI és munkatársai (2017) tanulmányában a MEEI-, az SVD és az AVPD beszédatadabázisokon végeztek osztályozásokat, és a pontosságok 99,54%, 99,53% és 96,02% voltak az egyes adatbázisokon. Az osztályozásokat SVM-mel végezték, és /a/ kitartott magánhangzókat tartalmazó hangfájlokkal dolgoztak. A MEEI-adatbázison végzett tanulmányok egy része nagyon magas osztályozási pontosságot mutat, emiatt a kutatók megkérdőjelezték az adatbázis hasznosságát. MUHAMMAD és munkatársai (2017) azt állították, hogy az egészséges és kóros hangfelvételeket két különböző környezetben rögzítették ebben az adatbázisban, ezért nehéz kikövetkeztetni, hogy az osztályozó a hangmintákat a hangképzési rendellenesség szerint (egészséges vagy beteg személy hangfelvétele) vagy a rögzítési környezetek szerint osztályozza-e. Ebből kifolyólag a MEEI-adatbázison kapott eredményeket fenntartásokkal kell kezelni. Ez a munka rámutat a megfelelő beszédatadabázis létrehozásának fontosságára is, hogy ilyen osztályozási kísérleteket egyáltalán elvégezhessünk. Huiyi WU és



munkatársai (2018) a Saarbrücken-hangadatbázison vizsgálta 482 egészséges és 482 organikus diszfóniával rendelkező személy kitartott /a/ magánhangzóit. Konvolúciós neurális hálózatot (*Convolutional Neural Network*, CNN, a mély neurális hálózatok egy fajtája, amelyet elsősorban képek esetén alkalmaznak) használtak a jellemzők kinyerésére, valamint az egészséges és az organikus diszfónia tüneteit mutató beszéd elkülönítésére. 88,5%, 66,2% és 77,0%-os osztályozási pontosságot értek el a tanító, a validáló és a tesztelő adathalmazokon.

A folyamatos beszéd használatával az osztályozási feladat sokkal nehezebb, de ezek az eredmények várhatóan jobban alkalmazhatóak a gyakorlatban. A GUEDES és munkatársai (2019) által bemutatott munkában a kutatók a német Saarbrücken-hangadatbázisban szereplő *Guten Morgen, wie geht es Ihnen?* című rész hanganyagát használták a diszfóniás és az ép beszéd-minták osztályozásához. 66%-os értéket értek el az LSTM (*Long-Short-Term-Memory*) és CNN mély neurális hálózatokkal. A kutatók rámutattak arra, hogy a mély tanulás (*Deep Learning*) használata a kóros beszéd felismeréséhez hatalmas adatmennyiséget igényel. VICSI, IMRE és MÉSZÁROS (2011) munkájukban 26 egészséges és 33 diszfóniás személy magyar nyelvű beszédét osztályozták folyamatos beszédükből kinyert akusztikai jellemzők alapján. Az osztályozási pontosságuk 86% és 88% között volt.

TULICS és munkatársai (2020) tovább gyűjtötték ezeket a hangfelvételeket, létrehozva a már nemzetközi viszonylatban is nagy mintaszámúnak tekinthető „Magyar Diszfóniás és Egészséges Felnőtt Beszédatadtbázis”-t. A hangadatbázis felvételei közeltéri mikrofonnal (Monacor ECM-100), alacsony zajszintű külső hangkártyával (Creative Soundblaster Audigy 2 NX), jó minőségű A/D konverterrel (kódolás: PCM, mintavételezési frekvencia: 16 kHz, kvantálás: 16 bit) kerültek rögzítésre, csendes irodai környezetben (orvosi szobában). Minden páciens Aiszóposz *Az északi szél és a nap* című meséjét olvasta fel, amely szöveg gyakran használt a foniátriai kutatásokban. Az adatbázis fonémaszintű szegmentálása automatikus fonémaszegmentáló programmal történt, amit a Beszédakusztikai Laboratórium munkatársai fejlesztettek ki (KISS–VICSI 2017), majd szükség esetén kézzel történt ezek javítása. A hangadatbázis felvételei között számos hangképzési rendellenességgel rendelkező beteg hangja fordul elő: funkcionális diszfónia, hangszalagbénulás (rekurrens parézis), a hangképző szervrendszer különböző pontjain előforduló tumorok, gasztroözofoageális reflux (GERD), krónikus gégegyulladás, agyideggyulladás (*bulbar paresis*), amiotrófiás laterálszklerózis (ALS), leukoplakia, spazmodikus diszfónia stb. A leggyakoribb betegségek a funkcionális diszfónia (FD) és a hangszalagbénulás (rekurrens parézis, RP). Az adatbázis összesen 450 felvételt tartalmaz, 257 diszfóniában szenvedő páciens hangját (156 nő és 101 férfi) és 193 ép beszéddel rendelkező adatközlő felvételét (108 nő és 85 férfi). TULICS és munkatársai kimutatták, hogy a diszfóniás és egészséges hangok bináris osztályozása a magyar beszédre vonatkozóan is lehetséges. 88%-os osztályozási pontosság érhető el akusztikai jellemzőkből álló bemeneti vektorral, mély neurális hálózat alkalmazásával, a „Magyar Diszfóniás és Egészséges Felnőtt Beszédatadtbázis”-t használva (TULICS 2020; TULICS et al. 2019; TULICS et al. 2020).

A legtöbb esetben a kutatások az egészséges beszéd és néhány specifikus betegség elkülönítésére összpontosítanak, de néhány tanulmány a többosztályos osztályozásra összpontosít

(GUEDES et al. 2019). Például KAZINCZI és munkatársai (2015) munkájukban először azt mutatták be, hogy egy SVM-alapú osztályozó képes automatikusan elkülöníteni az egészséges és kóros hangokat, viszonylag nagy, 87%-os pontossággal. Ezek után egy többsztályos osztályozást végeztek az egészséges, a funkcionális diszfóniával (FD) diagnosztizált és a hangszalag-bénulásos (RP) csoportok között, 60%-os pontossággal. Ezekon kívül egy kétsztályos osztályozás során 85%-os pontosságot értek el az egészséges és az RP-csoport között, 78%-os pontosságot az FD- és az RP-csoport között, az egészséges és az FD-csoport között pedig 66%-os pontosságot. Egy másik, hasonló jellegű kutatást végeztek MUHAMMAD és munkatársai (2011). A kutatók csak formánsfrekvencia-értékeket (energiacsúcsok a hangspektrumban) alkalmaztak bemeneti jellemzőként az osztályozó számára, hogy több hangképzési rendellenességet különítsenek el automatikusan egymástól, egy többsztályú forgatókönyvben (tehát kettőnél több célcímként használtak, és nem csak egészséges vs. diszfónia címkéket). Ezek az organikus eredetű betegségek a következők voltak: ciszta, GERD, bénulás, polip és sulcus. Az osztályozási pontosság 67,86% volt a nők és 52,50% a férfiak esetében.

Érdekes kutatási kérdés, hogy szét lehet-e automatikusan választani egymástól a funkcionális és az organikus diszfóniát. Ha a funkcionális diszfóniát nagy valószínűséggel meg lehetne határozni egy diagnosztikát segítő rendszerrel, akkor a pácienseket gyorsan foniáterhez vagy logopédushoz lehetne irányítani. Ha viszont a rendszer organikus diszfóniát észlel, a beteget otolaringológushoz vagy onkológushoz lehetne irányítani. Egy ilyen rendszer sok időt takaríthatna meg.

A funkcionális diszfónia (FD) és az organikus diszfónia (OD) meghatározása és szétválasztása körül vita van, a két kategória nem mindig zárja ki egymást. Természetes, hogy a két csoport jobban osztályozható egy olyan beszédatadabázison, ahol a rekedtség súlyosságának eloszlása statisztikailag eltérő a két csoport között, például, ha az OD-csoport hangfelvételei statisztikailag szignifikánsan magasabb fokú rekedtségi súlyosságot mutatnak, mint az FD-csoportban lévő hangfelvételek. Ilyen módon az osztályozó a hangfelvételeket a betegség típusa helyett a rekedtség súlyossága szerint osztályozhatja alacsony és magas rekedtségi osztályba. Azonban a célunk valójában az, hogy a diszfóniában szenvedő emberek hangfelvételeit funkcionális és organikus diszfóniacsoportokba soroljuk. Ennek érdekében TULICS és munkatársai (2020) létrehozták a „Kiválasztott Diszfóniás Beszédatadabázis”-t (ami a „Magyar Diszfóniás és Egészséges Felnőtt Beszédatadabázis” egy szűrt változata), amelyben a rekedtség súlyosságának megoszlásában az OD- és az FD-csoportok között nem volt szignifikáns különbség. Munkájuk során bemutatták, hogy az organikus és a funkcionális diszfónia automatikus elválasztása akusztikus jellemzők felhasználásával lehetséges 71%-os osztályozási pontossággal, lineáris kernelű SVM használatával.

## 6. A diszfónia súlyosságának automatikus megítélése

A hangminőség objektív értékelése az értékelési folyamat alapvető része. Korábban említettünk néhány olyan szubjektív skálát, amelyet az orvosi területen használnak a diszfónia súlyosságának számszerűsítésére. Nehézségekkel szembesülünk azonban, amikor megbízható értékelést akarunk kapni a diszfónia súlyosságáról, mivel ezek az értékelő belső normáitól függenek.

A betegekkel foglalkozó szakemberek eltérő érzékenységet mutathatnak bizonyos auditív dimenziókra (például érdesség vagy levegősség), továbbá a fáradtság, a figyelem, a hangképzési rendellenességekben való jártasság vagy az ezzel kapcsolatos korábbi képzés is tényező lehet, és mindez következtelen ítéletekhez vezethet (EADIE–BAYLOR 2006). Ilyen pontatlanságok miatt a kutatások összefüggést próbáltak kimutatni az akusztikai jellemzők és a diszfónia súlyossági értékelései között. A mai technológiával megvannak a lehetőségek arra, hogy akusztikai jellemzőket nyerjünk ki a hangfelvételekből, és segítségükkel megkísérélhetjük a diszfónia súlyosságának automatikus becslését. Ez az eljárás nem invazív becslést biztosít, és független az orvos szakértelmétől.

Az ilyen típusú probléma a gépi tanulás szempontjából tipikusan többszintű osztályozási vagy regressziós feladat. Felmerülhet a kérdés, hogy hogyan válasszuk ki a célváltozót egy regressziós eljáráshoz. A kutatásokban általában egy (jobb esetben több) szakember értékelése alapján hajtunk végre regressziós elemzést (CHIEN–BORSKY–GUDNASON 2017; LIU et al. 2017; MARYN et al. 2010), azonban a szakemberek értékelésének konzisztenciáját csak néhány esetben vizsgálták. A fent említett tényezők miatt ez problémát jelenthet, ezért érdemes megvizsgálni a szakemberek értékelését a következetesség szempontjából.

Továbbá, ahogyan az osztályozás esetében is, itt is kérdés, hogy folyamatos beszédet vagy kitarított magánhangzókat érdemes-e inkább használni a vizsgálatok során, hiszen ezek befolyásolhatják az emberi minősítést. Tanulmányok azt mutatták ki, hogy az értékelők a diszfóniát, különösen a hang levegősségét sokkal súlyosabbnak ítélik kitarított hangok esetében, mint folyamatos beszédben (DE BODT 1997; MARYN–ROY 2012; WOLFE–CORNELL–FITCH 1995; ZRAICK–WENDEL–SMITH–OLINDE 2005).

MARYN és munkatársai (2010) a hangból kinyerhető akusztikai jellemzők és a hangminőség értékelése közötti kapcsolatot vizsgálták. Kísérletük során 22 egészséges személy és 228 hangképzési rendellenességgel küzdő beteg folyamatos beszédét és kitarított magánhangzóit használták fel. Egy több akusztikai jellemzőt is használó modellt fejlesztettek ki, melynek célja az volt, hogy meghatározza a kísérletben részt vevők általános hangminőségét, öt tapasztalt szakember értékelése alapján. A modell predikciója és az öt szakember értékelésének átlaga közötti korreláció 0,78 lett. Bár ez a kutatás ígéretes, a javasolt módszer nem használható a gyakorlatban a hangminőség gyors klinikai felmérésére, ha a pácienseknek kitarított hangot és folyamatos beszédet is produkálnia kell, mivel ez meghosszabbítja az orvosi vizsgálat időtartamát.

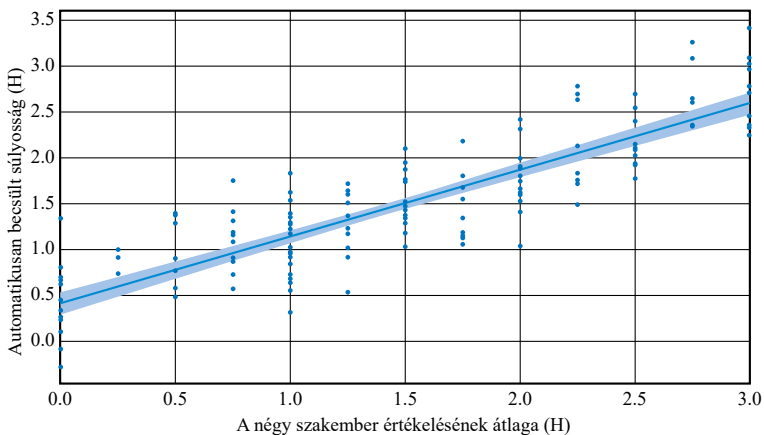
A kutatások között további különbségek még, hogy a diszfónia súlyosságát súlyossági csoportokba való besorolással (többszintű osztályozással) (FUJIMURA et al. 2020; GIDAYE et al. 2020; LIU et al. 2017) vagy regressziós eljárás alkalmazásával határozzák-e meg (LATOSZEK

et al. 2019; TULICS–VICSI 2019; WUYTS et al. 2000). GIDAYE és munkatársai (2020) az AVPD adatbázist használták, ahol a beszédmintákat három súlyossági fokozatba sorolták: enyhe, közepes és súlyos. Különböző hangképzési rendellenességekre vonatkozóan külön elvégezték a többosztályos osztályozást. Ciszta esetén az osztályozás pontossága az enyhe, közepes és súlyos kategóriák között 73,8% volt, hangszalagbénulás esetén 95,2%, polip esetében 67,5%. A kísérlet azt is mutatja, hogy a betegségek különböző hatékonysággal sorolhatók be súlyossági csoportokba. A tanított modellek osztályozási pontossága nagyban függ a beszédatadatbázisokban található betegségek típusától. A FUJIMURA és munkatársai (2020) által közölt kutatás célja az volt, hogy megkülönböztessék a GRBAS skála által leírt perceptuális dimenziókat egydimenziós konvolúciós neurális hálózat segítségével (*One-Dimensional Convolutional Neural Network, 1D-CNN*). Kísérletük során 1377 kitartott /a/ magánhangzót tartalmazó adatbázist használtak. Minden hangmintát három szakértő értékelt a GRBAS skála szerint, és helyes címkeként a három értékelés mediánját választották. A G-skála pontossága 77,1%, az R-skála modelljének pontossága pedig 76% volt.

Mivel az RBH-skálát a GRBAS skála alapján fejlesztették ki, találunk mind a GRBAS (FUJIMURA et al. 2020; SÁENZ-LECHÓN et al. 2006; WUYTS et al. 2000), mind az RBH (SZKLANNY–WRZECIONO 2019; TULICS–VICSI 2019) szubjektív skálák felhasználásával végzett kutatásokat. Mindkét skálát széles körben alkalmazzák a terület szakértői, és együttesen a hangproblémák minden típusára kiterjeszthetők, etiológiájuktól függetlenül.

TULICS és VICSI a diszfónia súlyosságának automatikus becslési lehetőségeit vizsgálta magyar nyelvű adatbázison, folyamatos beszédet felhasználva (TULICS 2020; TULICS–VICSI 2019). A kísérletben használt adatbázis összesen 148 felvételt tartalmazott, négy szakembert kértek fel a hangfelvételek RBH-skála szerinti értékelésére. A négy szakember egyike állította fel a diagnózist, és értékelte a beteg beszédének minőségét a konzultációk során, míg a másik három szakember nem került személyes kapcsolatba a páciensekkel, csupán a hangfelvételek alapján határozták meg a diszfónia súlyosságát. Mind a négy szakembernek nagy tapasztalata van a hangképzési rendellenességekkel rendelkező páciensek kezelésében. A szakemberek átlagos H perceptuális értékelését (RBH-skála H értéke) használták fel célváltozóként a regressziós modellhez, ezért érdemes volt megvizsgálni a négy szakember értékeléseinek konzisztenciáját. A szakemberek belső konzisztenciájának értéke képet ad arról, hogy mekkora a legnagyobb korrelációs érték, amelyet a regressziós modellünktől elvárhatunk, hiszen nem várjuk el, hogy a regressziós modell jobb eredményt érjen el, mint egy jól képzett szakember. A szakemberek belső konzisztenciáját (megbízhatóságát) a Cronbach alfa és az Intra Class Correlation Coefficient (ICC) metrikákkal számolták ki. A szakemberek belső konzisztenciája magas fokú volt a metrikák szerint (Cronbach Alpha = 0,89; ICC = 0,89). A kísérlet azt is kimutatta, hogy a diszfónia súlyosságának automatikus becslése lehetséges csupán nyolc akusztikai jellemző felhasználásával, lineáris kernelű SVR-rel. Ezzel a módszerrel a modell által prediktált súlyosság és a szakemberek értékelésének átlaga között 0,85 Pearson-korrelációs együtthatót kaptak, és 0,46 volt az átlagos négyzetes hiba gyöke (*Root Mean Square Error, RMSE*).

Az 1. ábra mutatja a diszfónia automatikusan prediktált súlyosságát az eredeti perceptuális referenciaértékekhez képest.



1. ábra

*Automatikusan becsült diszfóniasúlyosság a perceptuálisan megítélt H-érték alapján, lineáris kernelű SVR-rel és 8 jellemzővel*

Az ábra a lineáris kernelű SVR regressziós modelljét mutatja be, melynek jellemzőhalmazát egy jellemzőkiválasztó algoritmus eredményezte. Az ábra bemutatja a javasolt megközelítés becslési hatékonyságát, függetlenül a páciens patológiás háttérétől és a diszfóniájának súlyosságától. Látható, hogy a modell jó predikciót ad a H1 súlyossági osztályra, súlyosabb esetekben pedig alulról becsli azokat. Mivel az ICC-érték a 4 szakember között 0,89-t eredményezett, ez tekinthető egy olyan elméleti célnak, amelyet el akarunk érni. Ennek fényében a regressziós modellel kapott 0,85 korrelációs értéket szinte tökéletesnek tekinthetjük.

## 7. Automatikus diszfóniaosztályozás gyermekek esetében

A gyermekkori diszfónia akusztikai elemzéseinek vizsgálata ritka, annak ellenére, hogy nagyon fontos terület. Körülbelül a gyermekek 23,4%-át érinti diszfónia (STACHLER et al. 2018). A diszfónia negatív hatással van a gyermek testi és lelki egészségére, mivel a diszfóniás gyermek nem képes megfelelően kommunikálni. Tanulási nehézségek alakulhatnak ki a diszfóniás gyerekeknél, valamint az önkép- és önbecsülési zavar sem ritka esetükben (SWAIN et al. 2019). Egy másik nagy probléma, hogy a szülők nem társítják a gyermekük diszfóniás hangját hangképzési rendellenességhez, és ez késleltetheti a diagnózist és a megfelelő orvosi ellátást.

Mivel gyermekeknél a hangszalagsomók a diszfónia leggyakoribb okai (KILIC et al. 2004), a hangszalagsomók és a beszédváltozások kapcsolata áll a kutatások középpontjában. Egy 254 beteget vizsgáló tanulmány kapcsolatot mutatott ki a hangszalagsomó mérete és a beszéd rekedtsége, levegőssége, megterheltsége és az afónia között (SHAH et al. 2005).

Mivel a diszfóniás gyermekbeszédet tartalmazó beszédadatbázisok mintamérete kicsi, a kutatók elsősorban statisztikai elemzésekre koncentrálnak, nem pedig osztályozásra (DOHAR–SHAFFER–WHITE 2019; MEREDITH et al. 2008; TULICS–KAZINCZI–VICSÍ 2016).

Janete COELHO és munkatársai (2016) hangszalagsomóval rendelkező iskoláskorú gyermekek percepció és akusztikai jellemzőit vizsgálták, és eredményeiket összevetették egy ép beszédű kontrollcsoportéval. A kutatás során tíz 7 és 12 év közötti gyermeket vizsgáltak, öt fiút és öt lányt. Statisztikailag szignifikáns különbségeket találtak a hangszalagsomóval rendelkező csoport és az ép beszédű kontrollcsoport között a következő paramétereket felhasználva: alapfrekvencia, shimmer, HNR, maximális fonációs idő kitartott hangokra.

TULICS és VICSÍ (2018) egészséges és diszfóniás gyermekek hangjának automatikus elkülönítését tűzte ki célul. Ehhez előbb egy jól felépített hangadatbázisra volt szükség, és ehhez több óvodában gyűjtöttek hanganyagot. A felvételek során Bartos Erika *Mókus* című versét mondta el az összes gyerek. A vers választását a sok szó eleji /m/ nazális hangzó indokolta, amely a hangterápiában a hang előrehozásánál jelentős szerepet játszik. Egy felvétel átlagosan 20 másodperc hosszú. A vers leggyakoribb magánhangzója az /o/ fonéma 16 előfordulással, amelyet 14 előfordulással az /v/ fonéma és 9 darab előfordulással az /ε/ követ. A felvételek közeltéri mikrofonnal készültek (Monacor ECM-100), alacsony zajszintű külső hangkártyával (Creative Soundblaster Audigy 2 NX), jó minőségű A/D konverterrel (kódolás: PCM, mintavételezési frekvencia: 16 kHz, kvantálás: 16 bit). Az adatbázis összesen 59 felvételt tartalmaz: 25 felvétel diszfóniában szenvedő gyermek hangját (átlagos életkor: 6,52±1,94; a gyerekek többségének diagnózisa funkcionális diszfónia, míg 3 gyermeknek hangszalagsomói voltak), valamint 34 felvétel egészséges gyermekektől származik (átlagéletkor: 5,35±0,54). A „Diszfóniás és Egészséges Gyermek Beszédadatbázis”-ban összegyűjtött adatmennyiség már elegendő ahhoz, hogy osztályozási kísérleteket lehessen végezni egészséges és diszfóniás gyerekek hangja között.

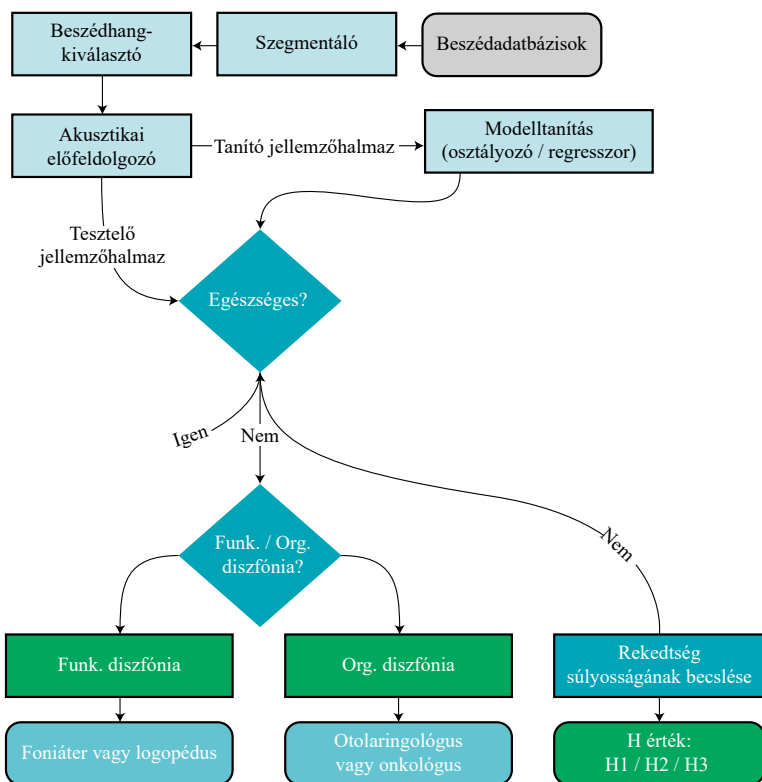
Bináris osztályozás során a bemeneti vektor akusztikai jellemzőkből lett összeállítva. A kísérlet során megmutatták, hogy az egészséges és a diszfóniás gyermekek hangjainak automatikus osztályozása lehetséges 93%-os osztályozási pontossággal, SVM-mel a kialakított, gyermekhangokat tartalmazó beszédadatbázison. Ez a kutatás jó alapként szolgálhat a további kutatásokhoz, továbbá úgy tűnik, hogy a diszfónia korai szakaszban hatékonyabban szűrhető, de ilyen állítások megfogalmazásához sokkal több adatra van szükség.

## 8. Alkalmazhatóság

Ahogy a fentiekben összefoglalt eredmények is mutatják, gyakorlatilag már megvalósítható egy olyan diagnosztikát segítő rendszer kifejlesztése, amely képes a diszfóniás és az egészséges beszéd megkülönböztetésére. Fontos azonban megjegyezni, hogy míg a rendszer jól használható előszűrésre, a pontos diagnózis felállítása továbbra is az orvos feladata.

A felnőttek számára javasolt rendszer több lépésből áll: a betegek beszédfelvételeit beszédadatbázisokba rendezzük (például „Magyar Diszفóniás és Egészséges Felnőtt Beszédadatbázis”). A felvételeket normalizáljuk és szegmentáljuk, fonémaszinten. Az elemzendő fonémák kiválasztása után akusztikai jellemzőket nyerünk ki a hangfelvételekből, és vektorba rendezzük őket. A jellemzővektort a rendszer az előzetes ismeretek alapján binárisan osztályokba sorolja (egészséges vagy diszفóniás). Az előzetes ismeretekbe tartozik egy gondosan felépített beszédadatbázis, valamint az optimális osztályozási és regressziós modellek. Ha a hangfelvétel alapján a beszélő hangja egészségesnek minősül, a folyamatnak itt vége. Ha diszفóniásnak minősül, a diagnosztikát segítő rendszer felismeri a diszفónia típusát (funkcionális vagy organikus diszفónia), és a betegség súlyosságát is megbecsüli regressziós modell alapján.

Az új beszédminták esetében az osztály (egészséges vagy diszفóniás) és a diszفónia súlyossága ismeretlen. A hangfelvétel előfeldolgozási módszere ugyanaz, mint korábban: akusztikai jellemzőket mérünk a fonémák szintjén, majd egy tesztelő jellemzőhalmazt alakítunk ki. Ez a tesztelő jellemzőhalmaz egy összehasonlító egységbe kerül, és elvégezzük rajta az osztályozási vagy regressziós műveleteket. Ez a folyamat a 2. ábrán látható.



2. ábra

A felnőttek számára javasolt diagnosztikát segítő rendszer kerete

Az összehasonlító egység prediktálja, hogy a tesztminta az egészséges vagy a diszfóniás csoportba tartozik-e. Ha a tesztminta egészséges, akkor a folyamat véget ér, ha viszont az összehasonlító egység diszfóniát prediktált, akkor megbecsüli a rekedtség súlyosságát, továbbá egy újabb összehasonlító egység aktiválódhat, ahol a rendszer a diszfónia típusát osztályozza. Ha a diagnosztikát segítő rendszer funkcionális diszfóniát prediktál nagy valószínűséggel, akkor a beteget gyorsan foniatérhez vagy logopédushoz lehetne irányítani. Ha viszont a rendszer organikus diszfóniát észlel, a beteget otolaringológushoz vagy onkológushoz lehetne irányítani. Egy ilyen diagnosztikát segítő rendszer működtetésével rengeteg időt nyerhetünk, a pácienseket gyorsan el lehetne irányítani a megfelelő szakrendelésre. A javasolt rendszer fiatal orvosoknak vagy házi orvosoknak is segítene hatékonyan kiszűrni a diszfóniás betegeket és automatikusan megállapítani a betegségük súlyosságát.

A diszfónia korai felismerését szolgáló diagnosztikát segítő rendszer a gyermekek esetében is a fent leírt logikát követné, azzal a különbséggel, hogy az organikus és funkcionális okok még nem megkülönböztethetőek. Mivel az osztályozási eredmények a gyermekek esetében ígéretesek, további beszédminták gyűjtése javasolt, hogy általánosítani lehessen az osztályozó modellt nagyobb adatkészleten. Hosszú távon érdemes kifejleszteni egy eszközt, amely automatikusan észleli a diszfóniás hangokat a gyerekek körében. A mobiltelefonok alkalmasak lennének ennek a módszernek a bevezetésére és a gyakorlatban való alkalmazására. Az egészségügyi alkalmazásokat általában okostelefonokhoz vagy táblagépekhez, esetenként okosórákhoz tervezik. Ezek lehetővé teszik a felhasználók számára, hogy akkor és ott férjenek hozzá az információhoz, amikor és ahol szükségük van rá, csökkentve az információ keresésére pazarolt időt. Ezek az eszközök olcsók, könnyen használhatóak és hordozhatóak.

A hangmintákat, a metaadatokat, a mért akusztikai jellemzők értékeit és az osztályozó kimenetét össze lehet gyűjteni, és fel lehet tölteni egy felhőszerverre. Ily módon hosszú távon monitorozhatjuk a gyermekek hangminőségét. A cél egy olyan szűrőrendszer felépítése, amelyet óvodai dolgozók használhatnak. Ha egy diszfóniás gyermeket időben kiszűrünk, akkor nagyobb esélyünk van arra, hogy a gyermek szakszerű segítséget kapjon egy fül-orr-gégésztől vagy egy logopédustól.

## FOGALMAK

*biomarker; diszfónia; organikus hangképzési rendellenességek; funkcionális hangirregularitások; neurogén hangirregularitás; strukturális rendellenesség; gépi tanulás; felügyelt gépi tanulás; felügyelet nélküli gépi tanulás; megerősítéssel gépi tanulás; osztályozás; regresszió*

## IRODALOM

AL-NASHERI, Ahmed – MUHAMMAD, Ghulam – ALSULAIMAN, Mansour – ALI, Zulfiqar – MALKI, Khalid H. – MESALLAM, Tamer A. – IBRAHIM, Mohamed Farahat 2017. Voice pathology detection and classification using auto-correlation and entropy features in different frequency regions. *IEEE Access* 6. 6961–6974.



- ALI, Zulfiqar – TALHA, Muhammad – ALSULAIMAN, Mansour 2017. A practical approach: Design and implementation of a healthcare software for screening of dysphonic patients. *IEEE Access* 5. 5844–5857.
- ALPAYDIN, Ethem 2020. *Introduction to machine learning*. MIT Press, Cambridge–London.
- ASHA 2020 = American Speech-Language-Hearing Association: *Voice Disorders*. <https://bit.ly/2OEIPZm> (A letöltés ideje: 2020. április 3.)
- BEHLAU, Mara – MADAZIO, Glaucya – OLIVEIRA, Gisele 2015. Functional dysphonia: Strategies to improve patient outcomes. *Patient Related Outcome Measures* 6. 243–253.
- BENNINGER, Michael S. – AHUJA, Anoop S. – GARDNER, Glendon – GRYWALSKI, Cindy 1998. Assessing outcomes for dysphonic patients. *Journal of Voice* 12/4. 540–550.
- BHATTACHARYYA, Neil 2014. The prevalence of voice problems among adults in the united states. *The Laryngoscope* 124/10. 2359–2362.
- BOSTON DYNAMICS 2020. *Do You Love Me?* <https://bit.ly/360bdtb> (A letöltés ideje: 2021. január 19.)
- BOTELHO, Catarina – TEIXEIRA, Francisco – ROLLAND, Thomas – ABAD, Alberto – TRANCOSO, Isabel 2020. Pathological speech detection using x-vector embeddings. arXiv preprint arXiv:2003.00864.
- CARDING, Paul N. – ROULSTONE, Sue – NORTHSTONE, Kate – ALSPAC Study Team 2006. The prevalence of childhood dysphonia: a cross-sectional study. *Journal of Voice* 20/4. 623–630.
- CHEN, Amy Y. – SCHRAG, Nicole M. – HALPERN, Michael – STEWART, Andrew – WARD, Elizabeth M. 2007. Health insurance and stage at diagnosis of laryngeal cancer: Does insurance type predict stage at diagnosis? *Archives of Otolaryngology – Head & Neck Surgery* 133/8. 784–790.
- CHIEN, Yu-Ren – BORSKÝ, Michal – GUÐNASON, Jón 2017. Objective severity assessment from disordered voice using estimated glottal airflow. In *Interspeech 2017*. 304–308. [https://www.isca-speech.org/archive/Interspeech\\_2017/pdfs/0138.PDF](https://www.isca-speech.org/archive/Interspeech_2017/pdfs/0138.PDF) (A letöltés ideje: 2020. április 3.)
- COELHO, Janete – RAMOS, Daniela – MONTEIRO, Isabel – PAIVA, António Diogo 2016. Vocal nodules in school age children. *Revista de Logopedia, Foniatria y Audiología* 36/3. 103–108.
- COHEN, Seth M. 2010. Self-reported impact of dysphonia in a primary care population: An epidemiological study. *The Laryngoscope* 120/10. 2022–2032.
- COHEN, Seth M. – DUPONT, William D. – COUREY, Mark S. 2006. Quality-of-life impact of non-neoplastic voice disorders: A meta-analysis. *Annals of Otolaryngology & Rhinology* 115/2. 128–134.
- COHEN, Seth M. – KIM, Jaewhan – ROY, Nelson – ASCHE, Carl – COUREY, Mark S. 2012a. Direct health care costs of laryngeal diseases and disorders. *The Laryngoscope* 122/7. 1582–1588.
- COHEN, Seth M. – KIM, Jaewhan – ROY, Nelson – ASCHE, Carl – COUREY, Mark S. 2012b. Prevalence and causes of dysphonia in a large treatment-seeking population. *The Laryngoscope* 122/2. 343–348.

- CORDEIRO, Hugo – MENESES, Carlos – FONSECA, José 2015. Continuous speech classification systems for voice pathologies identification. In CAMARINHA-MATOS, Luis M. – BALDISSERA, Thais Andrea – DI ORIO, Giovanni – MARQUES, Francisco (eds): *Technological Innovation for Cloud-Based Engineering Systems. DoCEIS 2015. IFIP Advances in Information and Communication Technology*. Vol 450. Springer, Cham. [https://doi.org/10.1007/978-3-319-16766-4\\_23](https://doi.org/10.1007/978-3-319-16766-4_23)
- COYLE, Suzanne M. – WEINRICH, Barbara D. – STEMPLER, Joseph C. 2001. Shifts in relative prevalence of laryngeal pathology in a treatment-seeking population. *Journal of Voice* 15/3. 424–440.
- DAVIDS, Taryn – KLEIN, Adams, M. – JOHNS III., Michael M. 2012. Current dysphonia trends in patients over the age of 65: Is vocal atrophy becoming more prevalent? *The Laryngoscope* 122/2. 332–335.
- DE BODT, Marc 1997. *A Framework for Voice Assessment: The relation between subjective and objective parameters in the judgment of normal and pathological voice*. University of Antwerp, Antwerpen.
- DE LA HOZ, Rafael E. – SHOHEET, Michael R. – BIENENFELD, Laura A. – AFILAKA, Aboaba A. – LEVIN, Stephen M. – HERBERT, Robin 2008. Vocal cord dysfunction in former World Trade Center (WTC) rescue and recovery workers and volunteers. *American Journal of Industrial Medicine* 51/3. 161–165.
- DIBAZAR, Alireza A. – NARAYANAN, Shrikanth – BERGER, Theodore W. 2002. Feature analysis for automatic detection of pathological speech. In *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society Engineering in Medicine and Biology*. Vol. 1. 182–183.
- DOHAR, Joseph E. – SHAFFER, Amber D. – WHITE, Katherine E. 2019. Pediatric dysphonia: It's not about the nodules. *International Journal of Pediatric Otorhinolaryngology* 125. 147–152.
- DUFF, Melissa C. – PROCTOR, Adele – YAIRI, Ehud 2004. Prevalence of voice disorders in African American and European American preschoolers. *Journal of Voice* 18/3. 348–353.
- EADIE, Tanya L. – BAYLOR, Carolyn R. 2006. The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. *Journal of Voice* 20/4. 527–544.
- FANG, Shih-Hau – TSAO, Yu – HSIAO, Min-Jing – CHEN, Ji-Ying – LAI, Ying-Hui – LIN, Feng-Chuan – WANG, Chi-Te 2019. Detection of pathological voice using cepstrum vectors: A deep learning approach. *Journal of Voice* 33/5. 634–641.
- FORTES, Felipe S. G. – IMAMURA, Rui – TSUJI, Domingos H. – SENNES, Luiz U. 2007. Profile of voice professionals seen in a tertiary health center. *Brazilian Journal of Otorhinolaryngology* 73/1. 27–31.
- FRANCIS, David O. – MCKIEVER, Monique E. – GARRETT, Gaelyn – JACOBSON, Barbara – PENSON, David F. 2014. Assessment of patient experience with unilateral vocal fold immobility: A preliminary study. *Journal of Voice* 28/5. 636–643.

- FUJIMURA, Shintaro – KOJIMA, Tsuyoshi – OKANOUE, Yusuke – SHOJI, Kazuhiko – INOUE, Masato – OMORI, Koichi – HORI, Ryusuke 2020. Classification of voice disorders using a one-dimensional convolutional neural network. *Journal of Voice* <https://doi.org/10.1016/j.jvoice.2020.02.009>
- GIBNEY, Elizabeth 2017. Google secretly tested AI bot. *Nature* 541. 142. <https://doi.org/10.1038/nature.2017.21253>
- GIDAYE, Girish – NIRMAL, Jagannath – EZZINE, Kadria – SHRIVAS, Avinash – FRIKHA, Mondher 2020. Application of glottal flow descriptors for pathological voice diagnosis. *International Journal of Speech Technology* 23/1. 205–222.
- GUEDESA, Victor – TEIXEIRA, Felipe – OLIVEIRA, Alessa – FERNANDES, Joana – SILVA, Leticia – JUNIOR, Arnaldo – TEIXEIRA, João P. 2019. Transfer learning with audio set to voice pathologies identification in continuous speech. *Procedia Computer Science* 164. 662–669.
- GUSS, Joel – SADOUGHI, Babak – BENSON, Brian – SULICA, Lucian 2014. Dysphonia in performers: Toward a clinical definition of laryngology of the performing voice. *Journal of Voice* 28/3. 349–355.
- HACKI, Tamás – MOERMAN, Mieke – RUBIN, John S. 2020. ‘Malregulative’ rather than ‘functional’ dysphonia: A new etiological terminology framework for phonation disorders — A position paper by the Union of European Phoniaticians (UEP). *Journal of Voice* <https://doi.org/10.1016/j.jvoice.2020.04.032>
- HAZLETT, Diane E. – DUFFY, Orla M. – MOORHEAD, Anne S. 2011. Review of the impact of voice training on the vocal quality of professional voice users: Implications for vocal health and recommendations for further research. *Journal of Voice* 25/2. 181–191.
- HIRANO, Michio 1981. GRBAS scale for evaluating the hoarse voice & frequency range of phonation. *Clinical Examination of Voice* 5. 83–84.
- HIRSCHBERG Jenő – HACKI Tamás – MÉSZÁROS Krisztina (szerk.) 2013. *Foniatória és társtudományok*. ELTE Eötvös Kiadó, Budapest.
- ISSETTI, Derek – MEYER, Tanya 2014. Workplace productivity and voice disorders: A cognitive interviewing study on presenteeism in individuals with spasmodic dysphonia. *Journal of Voice* 28/6. 700–710.
- JOHNS III, Michael M. – SATALOFF, Robert T. – MERATI, Albert L. – ROSEN, Clark A. 2010. Shortfalls of the American Academy of Otolaryngology–Head and Neck Surgery’s clinical practice guideline: Hoarseness (dysphonia). *Otolaryngology – Head and Neck Surgery* 143/2. 175–177.
- JONES, Katherine – SIGMON, Jason – HOCK, Lynette – NELSON, Eric – SULLIVAN, Marsha – OGREN, Frederic 2002. Prevalence and risk factors for voice problems among telemarketers. *Archives of Otolaryngology – Head and Neck Surgery* 128/5. 571–577.
- KAZINCZI, Ferenc – MÉSZÁROS, Krisztina – VICSI, Klára 2015. Automatic detection of voice disorders. In *International Conference on Statistical Language and Speech Processing*. 143–152. [https://doi.org/10.1007/978-3-319-25789-1\\_14](https://doi.org/10.1007/978-3-319-25789-1_14)

- KILIÇ, Mehmet A. – OKUR, Erdoğan – YILDIRIM, İlhami – GÜZELSOY, Saime 2004. The prevalence of vocal fold nodules in school age children. *International Journal of Pediatric Otorhinolaryngology* 68/4. 409–412.
- KISS, Gábor – VICSI, Klára 2017. Mono- and multi-lingual depression prediction based on speech processing. *International Journal of Speech Technology* 20/4. 919–935.
- LAHMIRI, Salim – DAWSON, Debra A. – SHMUEL, Amir 2018. Performance of machine learning methods in diagnosing parkinson's disease based on dysphonia measures. *Biomedical Engineering Letters* 8/1. 29–39.
- LATOSZEK, Ben B. – ULOZAITĖ-STANIENĖ, Nora – MARYN, Youri – PETRAUSKAS, Tadas – ULOZA, Virgilijus 2019. The influence of gender and age on the Acoustic Voice Quality Index and Dysphonia Severity Index: A normative study. *Journal of Voice* 33/3. 340–345.
- LEVENDOSKI, Elizabeth E. – SUNDARRAJAN, Anusha – SIVASANKAR, Preeti M. 2014. Reducing the negative vocal effects of superficial laryngeal dehydration with humidification. *Annals of Otolaryngology, Rhinology & Laryngology* 123/7. 475–481.
- LIU, Yuanyuan – LEE, Tan – CHING, Pak-Chung – LAW, Thomas K. T. – LEE, Kathy Y. 2017. Acoustic assessment of disordered voice with continuous speech based on utterance-level ASR posterior features. In *Interspeech 2017*. 2680–2684. [https://www.isca-speech.org/archive/Interspeech\\_2017/pdfs/0280.PDF](https://www.isca-speech.org/archive/Interspeech_2017/pdfs/0280.PDF) (A letöltés ideje: 2021. január 7.)
- LONG, Joanne – WILLIFORD, Henry N. – OLSON, Michele S. – WOLFE, Virginia 1998. Voice problems and risk factors among aerobics instructors. *Journal of Voice* 12/2. 197–207.
- MARTINEZ, David – LLEIDA, Eduardo – ORTEGA, Alfonso – MIGUEL, Antonio – VILLALBA, Jesús 2012. Voice pathology detection on the Saarbrücken Voice Database with calibration and fusion of scores using multifocal toolkit. In TORRE TOLEDANO, Doroteo – ORTEGA GIMÉNEZ, Alfonso – TEIXEIRA, António – RODRÍGUEZ, Joaquín González – HERNÁNDEZ GÓMEZ, Luís – SAN SEGUNDO HERNÁNDEZ, Rubén – CASTRO, Daniel Ramos (eds): *Advances in Speech and Language Technologies for Iberian Languages*. Communications in Computer and Information Science vol. 328. Springer, Berlin–Heidelberg. 99–109.
- MARTINS, Regina H. G. – PEREIRA, Eny R. B. N. – HIDALGO, Caio B. – TAVARES, Elaine L. M. 2014. Voice disorders in teachers. A review. *Journal of Voice* 28/6. 716–724.
- MARYN, Youri – CORTHALS, Paul – VAN CAUWENBERGE, Paul – ROY, Nelson – DE BODT, Marc 2010. Toward improved ecological validity in the acoustic measurement of overall voice quality: Combining continuous speech and sustained vowels. *Journal of Voice* 24/5. 540–555.
- MARYN, Youri – ROY, Nelson 2012. Sustained vowels and continuous speech in the auditory-perceptual evaluation of dysphonia severity. *Jornal da Sociedade Brasileira de Fonoaudiologia* 24/2. 107–112.
- MEREDITH, Morgan L. – THEIS, Shannon M. – McMURRAY, Scott – ZHANG, Yu – JIANG, Jack J. 2008. Describing pediatric dysphonia with nonlinear dynamic parameters. *International Journal of Pediatric Otorhinolaryngology* 72/12. 1829–1836.

- MESALLAM, Tamer A. – FARAHAT, Mohamed – MALKI, Khalid H. – ALSULAIMAN, Mansour – ALI, Zulfiqar – AL-NASHERI, Ahmed – MUHAMMAD, Ghulam 2017. Development of the arabic voice pathology database and its evaluation by using speech features and machine learning algorithms. *Journal of Healthcare Engineering* 8. <https://doi.org/10.1155/2017/8783751>
- MIRZA, Natasha – RUIZ, Cesar – BAUM, Eric D. – STAAB, Jeffrey P. 2003. The prevalence of major psychiatric pathologies in patients with voice disorders. *Ear, Nose & Throat Journal* 82/10. 808–814.
- MUHAMMAD, Ghulam – ALSULAIMAN, Mansour – ALI, Zulfiqar – MESALLAM, Tamer A. – FARAHAT, Mohamed – MALKI, Khalid H. – AL-NASHERI, Ahmed – BENCHERIF, Mohamed A. 2017. Voice pathology detection using interlaced derivative pattern on glottal source excitation. *Biomedical Signal Processing and Control* 31. 156–164.
- MUHAMMAD, Ghulam – ALSULAIMAN, Mansour – MAHMOOD, Awais – ALI, Zulfiqar 2011. Automatic voice disorder classification using vowel formants. In *IEEE International Conference on Multimedia and Expo (ICME)*. 1–6.
- NÉMETH Géza – OLASZY Gábor (szerk.) 2010. *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest.
- OMORI, Koichi 2011. Diagnosis of voice disorders. *Japan Medical Association Journal* 54/4. 248–253.
- OROZCO-ARROYAVE, Juan R. – HÖNIG, Florian – ARIAS-LONDOÑO, Julián D. – VARGAS-BONILLA, Francisco J. – SKODDA, Sabine – RUSZ, Jan – NÖTH, Elmar 2015. Voiced/unvoiced transitions in speech as a potential bio-marker to detect Parkinson's disease. In *Sixteenth Annual Conference of the International Speech Communication Association*. 95–99.
- PANDOLFINI, Bruce 1997. *Kasparov and Deep Blue: The historic chess match between man and machine*. Simon and Schuster, New York.
- PTOK, Martin – SCHWEMMLE, Cornelia – IVEN, Cari – JESSEN, Michael – NAWKA, Tadeus 2006. On the auditory evaluation of voice quality. *HNO* 54/10. 793–802.
- RAMIG, Lorraine O. – VERDOLINI, Katherine 1998. Treatment efficacy: voice disorders. *Journal of Speech, Language, and Hearing Research* 41/1. S101–S116.
- REITER, Rudolf – HOFFMANN, Thomas K. – PICKHARD, Anja – BROSCHE, Sibylle 2015. Hoarseness – causes and treatments. *Deutsches Ärzteblatt International* 112/19. 329.
- ROSA, Milka – BEHLAU, Mara 2017. Mapping of vocal risk in amateur choir. *Journal of Voice* 31/1. 118.e1–118.e11. <https://doi.org/10.1016/j.jvoice.2016.03.002>
- ROY, Nelson – KIM, Jaewhan – COUREY, Mark S. – COHEN, Seth M. 2016. Voice disorders in the elderly: A national database study. *The Laryngoscope* 126/2. 421–428.
- ROY, Nelson – MERRILL, Ray M. – GRAY, Steven D. – SMITH, Elaine M. 2005. Voice disorders in the general population: Prevalence, risk factors, and occupational impact. *The Laryngoscope* 115/11. 1988–1995.

- ROY, Nelson – MERRILL, Ray M. – THIBEAULT, Susan – PARSA, Rahul A. – GRAY, Steven D. – SMITH, Elaine M. 2004. Prevalence of voice disorders in teachers and the general population. *Journal of Speech, Language, and Hearing Research* 47/2. 281–293.
- ROY, Nelson – STEMPLE, Joseph – MERRILL, Ray M. – THOMAS, Lisa 2007. Epidemiology of voice disorders in the elderly: Preliminary findings. *The Laryngoscope* 117/4. 628–633.
- RUEDA, Alice – KRISHNAN, Sridhar 2019. Augmenting dysphonia voice using fourier-based synchrosqueezing transform for a CNN classifier. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 6415–6419.
- RUOTSALAINEN, Jani – SELLMAN, Jaana – LEHTO, Laura – VERBEEK, Jos 2008. Systematic review of the treatment of functional dysphonia and prevention of voice disorders. *Otolaryngology – Head and Neck Surgery* 138/5. 557–565.
- SÁENZ-LECHÓN, Nicolas – GODINO-LLORENTE, Juan I. – OSMA-RUIZ, Víctor – BLANCO-VELASCO, Manuel – CRUZ-ROLDAN, Fernando 2006. Automatic assessment of voice quality according to the GRBAS scale. In *International Conference of the IEEE Engineering in Medicine and Biology Society*. 2478–2481.
- SEIFERT, Eberhard 2005. Stress and distress in non-organic voice disorder. *Swiss Medical Weekly* 135/27–28. 387–397.
- SHAH, Rahul – WOODNORTH, GERALYN H. – GLYNN, Amy – NUSS, Roger 2005. Pediatric vocal nodules: Correlation with perceptual voice analysis. *International Journal of Pediatric Otorhinolaryngology* 69/7. 903–909.
- SILVERMAN, Ellen-Marie – ZIMMER, Catherine H. 1975. Incidence of chronic hoarseness among school-age children. *Journal of Speech and Hearing Disorders* 40/2. 211–215.
- SMITH, Elaine – KIRCHNER, Lester H. – TAYLOR, Margaretta – HOFFMAN, Henry – LEMKE, Jon H. 1998. Voice problems among teachers: Differences by gender and teaching characteristics. *Journal of Voice* 12/3. 328–334.
- SMITH, Elaine – LEMKE, Jon H. – TAYLOR, Margaretta – KIRCHNER, Lester H. – HOFFMAN, Henry 1998. Frequency of voice problems among teachers and other occupations. *Journal of Voice* 12/4. 480–488.
- STACHLER, Robert J. – FRANCIS, David O. – SCHWARTZ, Seth R. – DAMASK, Cecelia C. – DIGOY, German P. – KROUSE, Helene J. – MCCOY, Scott J. – OUELLETTE, Daniel R. – PATEL, Rita R. – REAVIS, Charles C. W. – SMITH, Libby J. – SMITH, Marshall – STRODE, Steven W. – WOO, Peak – NNACHETA, Lorraine C. 2018. Clinical practice guideline: Hoarseness (dysphonia) (Update). *Otolaryngology – Head and Neck Surgery* 158/1. S1–S42.
- STRIMBU, Kyle – TAVEL, Jorge A. 2010. What are biomarkers? *Current Opinion in HIV and AIDS* 5/6. 463–466.
- SWAIN, Santosh K. – NAHAK, Bulu – SAHOO, Loknath – MUNJAL, Sampada – SAHU, Mahesh C. – SANTOSH, Kumar 2019. Pediatric dysphonia – a review. *Indian Journal of Child Health* 6. 1–5.

- SZKLANNY, Krzysztof – WRZECIONO, Piotr 2019. Relation of RBH auditory-perceptual scale to acoustic and electroglottographic voice analysis in children with vocal nodules. *IEEE Access* 7/41. 647–658.
- TEIXEIRA, João P. – FERNANDES, Paula O. – ALVES, Nuno 2017. Vocal acoustic analysis–classification of dysphonic voices with artificial neural networks. *Procedia Computer Science* 121. 19–26.
- TITZE, Ingo R. 1994. *Principles of voice production*. Prentice Hall, Englewood Cliffs, N. J.
- TITZE, Ingo R. – LEMKE, Julie – MONTEQUIN, Doug 1997. Populations in the US workforce who rely on voice as a primary tool of trade: A preliminary report. *Journal of Voice* 11/3. 254–259.
- TULICS, Miklós G. – KAZINCZI, Ferenc – VICSI, Klára 2016. Statistical analysis of acoustical parameters in the voice of children with juvenile dysphonia. In *International Conference on Speech and Computer*. 667–674.
- TULICS, Miklós G. – LAVATI, Livia J. – MÉSZÁROS, Krisztina – VICSI, Klára 2019. Possibilities for the automatic classification of functional and organic dysphonia. In *International Conference on Speech Technology and Human-Computer Dialogue (SpeD)*. 1–6.
- TULICS, Miklós G. – SZASZÁK, György – MÉSZÁROS, Krisztina – VICSI, Klára 2019. Artificial Neural Network and SVM based Voice Disorder Classification. In *10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. 307–312.
- TULICS, Miklós G. – SZASZÁK, György – MÉSZÁROS, Krisztina – VICSI, Klára 2020. Using ASR Posterior Probability and Acoustic Features for Voice Disorder Classification. In *11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. 155–160.
- TULICS, Miklós G. – VICSI, Klára 2018. Automatic classification possibilities of the voices of children with dysphonia. *Infocommunications* 10/3. 30–36.
- TULICS, Miklós G. – VICSI, Klára 2019. The automatic assessment of the severity of dysphonia. *International Journal of Speech Technology* 22/2. 341–350.
- TULICS, Miklós G. 2020. *Automatic classification of dysphonia*. PhD-értekezés. BME, Budapest.
- TUTTLE, Hilary 2018. Facebook scandal raises data privacy concerns. *Risk Management* 65/5. 6–9.
- VERDOLINI, Katherine – RAMIG, Lorraine O. 2001. Occupational risks for voice problems. *Logopedics Phoniatics Vocology* 26/1. 37–46.
- VICSI, Klára – IMRE, Viktor – MÉSZÁROS, Krisztina 2011. Voice disorder detection on the basis of continuous speech. In *5th European Conference of the International Federation for Medical and Biological Engineering*. 86–89.
- WEI, Guo-Wei 2019. Protein structure prediction beyond AlphaFold. *Nature Machine Intelligence* 1/8. 336–337.
- WOLFE, Virginia – CORNELL, Richard – FITCH, James 1995. Sentence/vowel correlation in the evaluation of dysphonia. *Journal of Voice* 9/3. 297–303.

- WU, Huiyi – SORAGHAN, John – LOWIT, Anja – DI CATERINA, Gaetano 2018. Convolutional neural networks for pathological voice detection. In *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 1–4.
- WUYTS, Floris L. – BODT, Mark S. D. – MOLENBERGHS, Geert – REMACLE, Marc – HEYLEN, Louis – MILLET, Benoit – LIERDE, Kristiane V. – RAES, Jan – VAN DE HEYNING, Paul H. 2000. The Dysphonia Severity Index: An objective measure of vocal quality based on a multiparameter approach. *Journal of Speech, Language, and Hearing Research* 43/3. 796–809.
- ZHANG, Yu – JIANG, Jack J. 2008. Acoustic analyses of sustained and running voices from patients with laryngeal pathologies. *Journal of Voice* 22/1. 1–9.
- ZRAICK, Richard I. – WENDEL, Kristen – SMITH-OLINDE, Laura 2005. The effect of speaking task on perceptual judgment of the severity of dysphonic voice. *Journal of Voice* 19/4. 574–581.

## **KÖSZÖNETNYILVÁNÍTÁS**

A bemutatott kutatás a K128568 számú projektben, a Nemzeti Kutatási Fejlesztési és Innovációs Alapból biztosított támogatással, a K pályázati program finanszírozásában valósult meg. A kutatást részben a CELSA (CELSA /18/027) projekt finanszírozta.





# A depresszió automatikus becslése a beszéd akusztikai-fonetikai jellemzői alapján

Kiss Gábor

BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM

## 1. Bevezetés

Az emberi beszéd összetett, sokrétű információ továbbítására alkalmas kommunikációs eszköz. Beszédjel feldolgozása során, a verbális információ (nyelvi tartalom) kinyerése mellett számos nemverbális információ kinyerése is lehetséges, mint például a beszélő neme, a beszélő által használt nyelv, becsülhető a beszélő kora, érzelmi állapota (VETRÁB–GOSZTOLYA 2019), fáradtsági szintje (LI et al. 2014), alkoholos befolyásoltsága (ZHANG–WENINGER–SCHULLER 2017), az, hogy dohányzik-e (SATORI et al. 2017), meg van-e fázva (GOSZTOLYA és mtsai 2017). A beszélő általános állapotának becslése mellett egyes neurológiai és/vagy pszichés megbetegedéseire is lehet következtetni, mint például Parkinson-kór (SZTAHÓ és mtsai 2017), amyotrófiás laterálszklerózis (ALS) (VASHKEVICH et al. 2018) vagy depresszió (KISS 2019) meglétére. Általánosságban elmondható, hogy minden olyan nemverbális információ kinyerése lehetséges az emberi beszédjelből, ami hatással bír a beszédproduktumra, ezáltal mérhető változást gyakorol az akusztikai-fonetikai jellemzőkre, és megfelelő statisztikai módszerekkel az adott nemverbális információ kimutatható, becsülhető.

A depresszió feltehetőleg egyidős az emberiséggel. Már az ókorban megfigyelték, hogy vannak depressziótól (melankóliától) szenvedő emberek, és lejegyezték ennek a feltételezett okait és lehetséges hatásait (JACKSON 1986). Arra a tényre, hogy a depresszió és a beszéd között kapcsolat állhat fenn, már 1921-ben rámutatott KRAEPELIN (1921), a modern pszichológia egyik megalapozója. Azonban az érzeti jellemzők és a beszédproduktum mérhető paraméterei közötti kapcsolatnak a pontosabb megértésére csupán az 1980-as évektől volt lehetőség az informatika fejlődésének és így a gépi feldolgozásnak köszönhetően. A korai megfigyelések elsősorban az átlagos alapfrekvencia értékének csökkenéséről, az alaphang dinamikájának beszűküléséről és a beszédtempó lassulásáról számolnak be (ALPERT–POUGET–SILVA 2001). Azóta a depresszió beszédre gyakorolt hatását számos esetben kimutatták, és ez alapján kijelenthető, hogy érzeti oldalon a depressziós ember beszédére jellemző a beszűkült beszéddinamika, a monotonitás, az élettelen, „fémés” hangszínezet, a lassú beszédtempó (CUMMINS et al. 2015; KISS 2019; LOW–BENTLEY–GHOSH 2020). Napjainkra számos akusztikai-fonetikai jellemzőről mutatták ki statisztikai módszerekkel, hogy az értékük megváltozik a depressziós állapot hatására, bár még mindig nincs teljes konszenzus, hogy pontosan mely akusztikai-fonetikai jellemzők

tükrözik a legjobban a depressziós állapotot (Kiss 2019). Azonban a mai modern statisztikai módszereknek köszönhetően, elsősorban a gépi tanulóeljárások megjelenésével, lehetőség nyílt a depressziós állapot automatikus detektálására, az állapot súlyosságának becslésére beszédjelfeldolgozás alapján. Egy ilyen rendszer alkalmas lehet a depresszió esetleges meglétének jelzésére, a diagnózis támogatására, ezzel elősegítve, hogy a depressziótól szenvedő egyén minél hamarabb megfelelő szakorvosi (pszichiátriai) ellátásban részesüljön. Emellett egy ilyen eljárás alkalmas lehet a depressziós állapotnak a súlyosság szerinti követésének támogatására is.

A jelen tanulmány célja, hogy bemutassa: a depressziós állapot hatására milyen változások figyelhetők meg a beszéd akusztikai-fonetikai jellemzőiben. Hogyan lehet megtervezni és megvalósítani a depresszió beszédalapú automatikus felismerését? Illetve vázlatosan bemutatja a Budapesti Műszaki és Gazdaságtudományi Egyetem (BME) Beszédakusztikai Laboratóriuma (LSA) által fejlesztett rendszer működését (Kiss 2019) és annak kiértékelését.

## 2. Az alkalmazott akusztikai-fonetikai jellemzők

Az emberi beszéd gépi elemzéséhez szükséges a beszédminta digitalizálása, vagyis a beszéd amplitúdó–idő függvényének (hullámforma) meghatározása. A nyers amplitúdó–idő függvény azonban önmagában nem vagy nem megfelelően alkalmazható a gépi tanulóeljárások bemeneteként, elsősorban azért, mert az így kapott jellemzővektor hossza függ a felvétel hosszától, másrészt túl nagy lenne a bemeneti jellemzővektor, emiatt célravezetőbb olyan származtatott akusztikai-fonetikai jellemzők kiszámolása és megadása, amelyek egyetlen adott értékekkel jellemzik a teljes vizsgált beszédmintát. Természetesen a beszéd összetettsége, változatossága miatt számos ilyen akusztikai-fonetikai jellemző nyerhető ki, amelyek mind releváns információt hordoznak a beszédproduktról. Ebben a fejezetben az akusztikai-fonetikai jellemzők olyan részhalmozát mutatjuk be, amelyeknek az értékei igazolhatóan megváltoznak a depressziós állapot hatására, és könnyebben összekapcsolhatók a beszéd érzeti leíró jellemzőivel.

### 2.1. Alapfrekvencia

Az emberi beszédalkeltéskor, az egyes beszédhangok kiejtése során, a hangszalagok kvázi-periodikus szétnyílása és összezáródása következtében, az adott hang amplitúdó–idő függvénye is kváziperiodikus lesz. Az ilyen beszédhangokat nevezzük zöngés hangoknak. Az **alapfrekvencia** ( $f_0$ ) alatt a zöngés hang periódusidejének a reciprokát értjük, mértékegysége a Hz. Az alapfrekvencia értéke a beszédalkeltés során folyamatosan változik, ez a variabilitás (és annak mértéke) adja a beszéd dallamoságának érzetét. Az alapfrekvencia értékét széles tartományban képes változtatni az ember, azonban az alapfrekvencia átlagos értéke jellemző a gyerekek (200–300 Hz), a nők (180 Hz) és a férfiak (100 Hz) esetében is (OLASZY 2010a).

## 2.2. Jitter

A beszédjel-feldolgozás során **jitter** /dʒit:ɛr/ alatt a zöngés szakaszok periódusidejének kis elemzési ablakában (néhány periódus) vett átlagos ingadozását értjük, általában százalékban kifejezve (HORII 1982). Olvasott beszéd esetében 2–6% közötti értékek a legtipikusabbak.

## 2.3. Shimmer

A beszédjel-feldolgozás során **shimmer** /ʃim:ɛr/ alatt a zöngés szakaszok periódusonkénti amplitúdóértékeinek kis elemzési ablakában (néhány periódus) vett átlagos ingadozását értjük, általában százalékban kifejezve (HORII 1982). Olvasott beszéd esetében 10–20% közötti értékek a legtipikusabbak.

## 2.4. Formánsfrekvenciák

A zöngés hangok felharmonikusszerkezettel (az alapfrekvencia egész számú többszöröse) is rendelkeznek. Ez a felharmonikusszerkezet jut az artikulációs csatornába, ahol az artikulációs csatorna által kialakított rezonanciafrekvenciákon és környékükön erősítés figyelhető meg. Az ilyen felerősödött felhangsoportokat nevezzük formánsoknak. A formánsfrekvencián pedig a rezonanciafrekvencia-tartomány burkológörbéjének maximum amplitúdóhelyét értjük, mértékegysége a Hz. Zöngés hangok esetében több **formánsfrekvencia** is megfigyelhető, azokat a frekvenciaértékük szerinti növekvő sorrendben nevezzük első ( $F_1$ ), második ( $F_2$ ) stb. formánsfrekvenciáknak (VICSÍ 2010a).

## 2.5. A formánsfrekvenciák sáv szélessége

A beszédjel-feldolgozás során a **formánsfrekvencia sáv szélessége** alatt a jelfeldolgozásban értendő általános sáv szélességet értjük, vagyis az alsó és felső határfrekvencia tartományát, ahol a határfrekvencia alatt azt az alsó és felső frekvenciaértéket értjük, ahol 3 dB-es csökkenés figyelhető meg a formánsfrekvenciánál mért erősítéshez viszonyítva. Mértékegysége a Hz. Az adott formánsfrekvencia sáv szélességét a formánsfrekvenciákhoz hasonlóan  $B_1$ ,  $B_2$  stb. rövidítéssel jelöljük.

## 2.6. Intenzitás szint

A hangintenzitás az egységnyi felületen merőlegesen átáramlott hangenergia [ $W/m^2$ ] mennyiségét jelenti adott időegység alatt. Az **intenzitás szint** alatt az aktuális hangintenzitás és egy adott referencia-hangintenzitás decibelskálán történő ábrázolását értjük. A beszédjel-feldolgozás

során a referenciaértéknek a hallásküszöböt használják, ami  $10^{-12}$  [W/m<sup>2</sup>]. Az intenzitásszint nagysága a hangerősség érzeti jellemzőt határozza meg. A beszéd során a magánhangzók intenzitásszintje a legnagyobb, így azok bírnak meghatározó jelleggel egy hosszabb beszédszakasz esetében (Vicsi 2010a).

## 2.7. Mel-sávós energiaértékek és MFCC

Az emberi hallásnak jól meghatározható frekvenciafelbontó-képessége van. A mel-skála az emberi fül felbontóképességét követi. A skála mértékegysége a mel, amely értékének a duplázódása a hangmagasságérzetet is duplázza. A mel-sávós energiaértékek a beszédjel sávszűrésével számíthatók ki, amelyet adott frekvenciaértéktől (például 60 mel), adott frekvenciatartományokkal (például 100 mel-enként) szokás elvégezni, és dB-ben kifejezni. Érdeemes megjegyezni, hogy a mel mértékegységben megadott fix tartományok frekvenciaértékben vizsgálva eltérő méretű tartományokat határoznak meg. Alacsony frekvenciákon ezek közel 100 Hz szélesek, míg a magasabb frekvenciaértékek esetében ennél jelentősen nagyobbak. Beszédjel-feldolgozás során általában elégséges 8 kHz-ig vizsgálni a beszéd frekvenciatartományát (spektrumát). 8 kHz-ig 27 darab mel-sávós energiaértéket lehet a fent bemutatott paraméterek mellett kiszámítani. Ha a kiszámított mel-sávós energiaértékeken diszkrét koszinusz transzformációt hajtunk végre, akkor kapjuk az MFCC együtthatókat (Vicsi 2010b), aminek tömörítő és lényegkiemelő hatása van, így ennek használatával jobb eredmények érhetők el kis méretű beszédatbázisok esetében.

## 2.8. Artikulációs sebesség

Az artikulációs sebesség (AS) alatt az egy másodperc alatt kiejtett beszédhangok számát értjük, mértékegysége pl. hang/s. Az artikulációs sebesség változik a beszéd során, és átlagos értéke nagyban függ a beszélő stílusától, illetve a beszéd típusától (olvasott vagy spontán beszéd). A túl magas artikulációs sebesség érthetetlenül gyors beszédhez vezethet, míg ha túl lassú az artikuláció, a beszéd folytonossága sérülhet (OLASZY 2010b).

## 2.9. Relatív szünethossz

A beszédproduktum természetes részét képezik a bennük található szünetek (csendes szakaszok). A szünetek adnak lehetőséget a levegővételre, gondolkodásra és beszédtervezésre, továbbá különböző információkat is közvetíthetnek, így segítve például a párbeszédben a beszélőváltás gördülékeny megvalósítását. A ritka és rövid szünetek tagolatlan, összefolyó beszédérzetet keltenek, míg a hosszú és gyakori szünetek akadozó beszédet eredményeznek. Relatív szünethossz

(RSZH) alatt a vizsgált beszédszakaszban található szünetek teljes hosszának arányát értjük a vizsgált beszédszakaszhoz viszonyítva, százalékban kifejezve (KISS 2019).

## 2.10. Tranziensarány

Normál beszéd esetében a beszédjel felbontható változó (tranziens) és közel állandó (kvázistacioner) részek egymásutánjaira. A közel állandó részeket általában a beszédhangok közepén találhatjuk, míg a változó részek a beszédhangok határainál jellemzőek (összetett beszédhangok esetében azok belső akusztikai határainál is megjelenhetnek tranziens szakaszok). Ebből kifolyólag, ha a hangok rövidek, és emellett gyors és nagy változások figyelhetők meg a beszédproduktumban, akkor kevés lesz az adott szakaszon belül a közel állandó szakasz, míg a lassú és monoton beszéd esetében több. A tranziensarány alatt (*ratio of transients* – RoT) a vizsgált beszédszakaszban található tranziens szakaszok teljes hosszának arányát értjük a vizsgált beszédszakasz teljes hosszához viszonyítva, százalékban kifejezve (KOVÁCS et al. 2015; KISS 2019).

## 3. Depresszió

A depresszió igen gyakori pszichiátriai betegség, amely a súlyosságának függvényében jelentősen befolyásolja az ettől szenvedő egyén életminőségét. FRIEDRICH (2017) tanulmánya rámutat, hogy a depresszió a jelenleg ismert egyik leggyakoribb gyógyítható betegség, ám gyakoriságának pontos meghatározását nehezíti, hogy diagnosztizálása szaktudást igényel. A World Health Organization 2002-es jelentése alapján a depresszió a negyedik leggyakoribb betegség világszerte (MATHERS–LONCAR 2006), a világszervezet 2012-es felmérése alapján már a harmadik leggyakoribb betegség (PREISS–LEAH–DAVID 2013), és a szervezet előrejelzése alapján 2030-ra a második leggyakoribb betegség lehet (MATHERS–LONCAR 2006).

Depresszióról akkor beszélhetünk, ha a vizsgált személy legalább két héten át szenved levertségtől, mély szomorúságtól, szeretetképtelenségtől. Ezek mellett a következő tünetek közül legalább négy jellemző rá, úgymint a pszichomotoros gátlás vagy remegés, csökkent kognitív állapot vagy megnövekedett határozatlanság, ájulás vagy energiavesztés, kóros álmatlanság vagy kóros álmoság, jelentős súlyvesztés vagy súlygyarapodás, jelentős negatív önértékelés vagy alaptalan túlzott bűntudat, gyakori gondolatok a halálról vagy gyakori öngyilkossági gondolatok (KLEIN 2008).

Jelenleg nem ismertek a depresszió pontos kiváltó okai (DWIVEDI et al. 2003), de a betegség meghatározható úgy, mint a kortikális limbikus rendszer egyfajta működési zavara (diszfunkciója). A depressziós állapot kialakulásában a genetikai hajlam mellett szerepet játszanak még a környezeti tényezők, mint például a stressz, traumatikus életesemény, bezártság, elszigeteltség.

A depressziós állapot megjelenésének következtében, illetve az állapot súlyosbodásával lecsökken az ettől szenvedő egyének kezdeményezőképesége, önbizalma, motiváltsága, ami akár a mindennapos tevékenységük elvégzésében is komoly akadályt jelenthet. Emiatt a depresszió jelentős terhet jelent a gazdaság számára. OLESEN és munkatársai (2012) tanulmánya rámutat, hogy Európa 30 országát vizsgálva 2010-ben a depresszió által okozott károk elérték a 92 milliárd eurós veszteséget, melyből 54 milliárd euró a depressziótól szenvedő egyének munkájukból való kiesése következtében keletkezett. Azonban talán még ennél is fontosabb, hogy a depressziós állapot következtében és súlyosbodásával jelentősen megnő az öngyilkosság kockázata. HAWTON és munkatársai (2013) tanulmánya rámutat, hogy az öngyilkosságok felében komoly szerepet játszik a depressziós állapot megléte.

Mivel nem ismert a depressziós állapotot meghatározó megfelelő érzékenységgel bíró objektív biomarker (biológiai jelző), így a betegség felismerése a pszichiáterekekre hárul. Figyelembe véve a depresszió társadalomban előforduló gyakoriságát, súlyosságát, a depressziós emberek diagnosztizálása komoly kihívás elé állítja az egészségügyi ellátórendszert. Emiatt napjaink fontos kutatási területének számít a depresszió diagnosztizálásának automatikus támogatása, a betegség esetleges jelenlétének automatikus jelzését megvalósító rendszerek kutatása, olyan nem invazív biomarkerek (mérésük nem igényel behatolást az emberi testbe) keresése, amelyek alkalmasak a depresszió gyors, lehetőség szerint költséghatékony felismerésére. A beszéd az egyik ilyen ígéretes biomarker (CUMMINS et al. 2015; KISS 2019; LOW–BENTLEY–GHOSH 2020).

### 3.1. A depressziós beszéd jellemzése

Az emberi beszédkeltés igen bonyolult, összetett folyamat, a beszéd előállításánál számos izom összehangolt, precíz együttes működtetésére, irányítására van szükség. A depressziós állapot hatással bír az izmok működtetésére, ezáltal megváltoznak a beszéd egyes akusztikai-fonetikai jellemzőinek átlagos értékei, így megfigyelhető az alaphérfvencia átlagos értékének és variabilitásának csökkenése, megnő a beszéd érdekssége, levegőssége (enyhe növekedés figyelhető meg a *jitter*- és *shimmer*értékekben). Emellett mivel a depresszió hatással bír az ettől szenvedő egyén kognitív képeségeire, így a depressziós emberek beszéde meglassul, több és hosszabb szünetet tartanak, a beszédükben gyakoribbak a különböző artikulációs és fonációs hibák, és beszédük monotonabbá válik (KISS 2019).

KISS (2019) tanulmánya magyar anyanyelvű beszélők olvasott beszédét vizsgálva számos akusztikai-fonetikai jellemző átlagos értéke esetében mutatott ki szignifikáns (99%-os szignifikanciaszint mellett) eltérést egészséges és depressziós személyek beszédmintáit összehasonlítva. Ilyen jellemzők voltak: a *shimmer* megnövekedése; az első és második formáns hérfvencia-értékének csökkenése és azok sáv szélességeinek emelkedése; illetve a beszéd szinképének megváltozása (a spektrális meredekség csökkenése). A tanulmány nemcsak azt mutatja be, hogy a depressziós személyek és az egészséges személyek beszéde között szignifikáns eltérés mutatható ki, hanem arra is rávilágít, hogy a depresszió súlyossága és az adott akusztikai-fonetikai jellemzők értéke között szignifikáns korreláció is kimutatható.

### 3.2. A depresszió súlyosságának mérése

A depressziót súlyossága alapján leggyakrabban 4 kategóriába szokás sorolni, úgymint minimális depresszió (nem depressziós állapot), enyhe depresszió, közepes depresszió és súlyos depresszió. A depressziós állapot súlyosságának automatikus becsléséhez azonban ennél részletesebb skálára, vagyis a súlyosság számszerű leírására van szükség.

A depressziós állapot súlyosságának számszerű jellemzésére a két leggyakrabban használt skála a Beck Depression Inventory-II (BDI) (BECK et al. 1996) és a Hamilton Rating Scale for Depression (HAM-D) (HAMILTON 1986). Mindkét skáláról elmondható, hogy széles körben elfogadottak, és a klinikai gyakorlatban is gyakran alkalmazzák ezeket. Mindkét skála (HAM-D és BDI) alkalmas a depresszió súlyosságának pontos, számszerű leírására, ugyanakkor néhány fontos különbség van köztük. Míg a HAM-D-értékek a megállapításához pszichiáter részvétele szükséges, addig a BDI-értéket egy önkitöltős kérdőív segítségével bárki megállapíthatja, ebből kifolyólag a HAM-D-érték tekinthető az objektívebb értéknek a kettő közül. Emellett a HAM-D-érték megállapításához több időre van szükség, szemben a BDI-kérdőív kitöltésével, ami 5–10 perc alatt elvégezhető. Az itt bemutatott eljárás esetében a BDI-értékek álltak rendelkezésre a depresszió súlyosságának jellemzéséhez.

A BDI-érték meghatározása 21 kérdés megválaszolásával lehetséges, ahol az egyes kérdésekre minden esetben 4 válaszlehetőség közül azt az egyet kell kiválasztania a vizsgált személynek, ami a legjobban jellemzi az aktuális állapotát. Minden válasznak 0-tól 3-ig adott az értéke. Az összes válasz értékét összegezve, a vizsgált személy depresszió szerinti súlyosságának jellemzése 0 és 63 között valósul meg, ahol a 0 érték a teljesen egészséges állapot, a 63 pedig a skála által mérhető legsúlyosabb depressziót jelenti. A BDI-skála és a fentebb említett 4 kategória összekapcsolhatósága az 1. táblázatban látható.

#### 1. táblázat

*A BDI-skála értékei és a depresszió súlyosságának kategóriái közti kapcsolat*

Kategória	BDI-érték
Minimális depresszió (nem depressziós)	0–13
Enyhe depresszió	14–19
Közepes depresszió	20–28
Súlyos depresszió	29–63

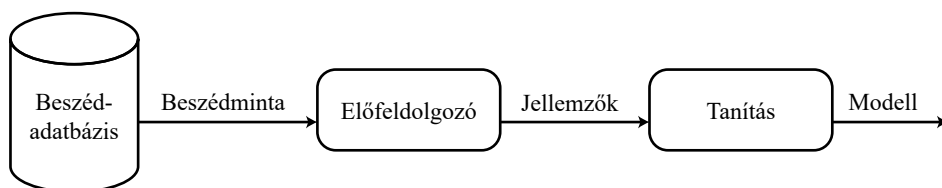
### 4. A depresszió súlyosságának automatikus becslése

Az a tény, hogy a depressziós állapot hatására, illetve a depresszió súlyosságának mértékével megváltozik az emberi beszédproduktum, lehetővé teszi ezen változások mérését, és a megváltozott akusztikai-fonetikai jellemzők alapján, megfelelő statisztikai módszerek segítségével



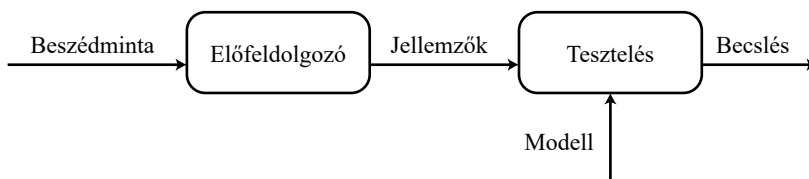
létrehozható olyan automatikus eljárás, amely képes becsléni a vizsgált személy depressziójának a súlyosságát. Ehhez szükséges egy beszédadatbázis, amelynek a segítségével megvalósítható a gépi tanulás, valamint egy előfeldolgozó rendszer, amely megteremti a kapcsolatot a beszédminta és a gépi tanulóeljárás között. Végül szükséges egy gépi tanulóeljárás választani, amely a kinyert jellemzők alapján előállítja a feladat megoldására alkalmas modellt.

Az eljárás megvalósítása önmagában nem bonyolult, azonban egy pontosan becselő rendszer létrehozásához több szempontot is érdemes figyelembe venni a tervezés és megvalósítás során. A BME LSA laboratóriumában egy ilyen eljárás készült (Kiss 2019), illetve áll folyamatos fejlesztés alatt. A modell tanításának az egyszerűsített folyamatábrája az 1. ábrán, míg a létrehozott modell tesztelésének egyszerűsített folyamatábrája a 2. ábrán látható.



1. ábra

*A BME LSA által fejlesztett, a depressziós állapot súlyosság szerinti becslését megvalósító modell tanításának egyszerűsített folyamatábrája*



2. ábra

*A BME LSA által fejlesztett, a depressziós állapot súlyosság szerinti becslését megvalósító eljárás működésének egyszerűsített folyamatábrája*

Megfigyelhető, hogy a modell tanítása és tesztelése nagyban hasonlítanak egymásra. Bár elsőre szimmetrikusnak tűnhet a két lépés, mégis lényeges különbségek vannak köztük. Egyrészt egy modell tanítása nagyságrendekkel több ideig tart, nagyobb erőforrás-igényű, mint egy új minta kiértékelése a már létrehozott modell alapján. Másrészt a tanítás célja valójában egy olyan modell kiválasztása, amely alkalmas lehet a feladat elvégzésére. A modell létrehozásakor a gyakorlatban több száz, akár több ezer modellt is betanítanak, majd azok kiértékelésével kiválasztják a legjobban teljesítő modellt, ami az eljárás kimenetele. Emiatt míg a modell létrehozása erőforrás- és időigényes, addig a modell tesztelése, az új minta kiértékelése gyorsabb és kisebb erőforrás-igényű. Az egyes lépések bővebb részletezését és a megvalósításukkor megfontolandó tervezési döntések tárgyalását a következő fejezetek tartalmazzák.

## 4.1. A modell létrehozása

A modell létrehozása statisztikai alapon történik, vagyis a modell arra keresi a választ, hogy a tanítóadat-bázis mintáiban rejlő információk és a célváltozó között milyen összefüggés mutatható ki statisztikai úton, és az összefüggés hogyan terjeszthető ki új minták esetében.

Az előfeldolgozó egység teremti meg a kapcsolatot a tanítóadat-bázis és a gépi tanulóljárás között, vagyis ez készíti el a tanítóadat-bázis mintáinak egy olyan reprezentációját, ami megadható a gépi tanulóljárás bemeneteként.

Értelemszerűen a tanítóadat-bázis minősége determinálja a létrehozott modell teljesítő-képességét, hiszen a tanítóadat-bázis információtartalma a problémáról nem lehet kisebb, mint a létrehozott modellté. Az előfeldolgozó egységnek emiatt is jelentős a szerepe, mivel a minták reprezentációja szükségszerűen információvesztéssel jár, így annak pontos megtervezése szintén nagyban hozzájárul a létrehozott modell teljesítő-képességéhez. Azonban magának a tanításnak is számos optimalizálandó paramétere van, mint például a gépi tanulóljárás kiválasztása, a tanító jellemzővektor kialakítása, a gépi tanulóljárás hiperparamétereinek optimalizálása.

### 4.1.1. A beszédatbázis

A beszédatbázis minősége – akár a legtöbb beszédtechnológiai alkalmazás esetében – nagymértékben meghatározza a létrehozott modell használhatóságát. A beszédatbázis létrehozása időigényes és költséges, emiatt a pontos megtervezése az egyik legfontosabb része a rendszer kialakításának. A tervezés során a következő főbb szempontokat érdemes figyelembe venni. A beszédatbázisba kiválasztandó személyek kiválasztása: hány személytől gyűjtünk mintákat, és ennek a csoportnak milyen legyen a nem, életkor, iskolai végzettség, egészségi állapot szerinti megoszlása. A gyűjteni kívánt beszédminták típusa lehet spontán beszéd (miről beszéljenek és mennyit) vagy felolvasás (milyen szöveget olvassanak fel). Idetartozik még a beszédminták rögzítési körülményeinek megtervezése, a rögzítőeszközök kiválasztása is.

Általánosságban elmondható, hogy beszédatbázisban a beszélők számát minden esetben érdemes maximalizálni úgy, hogy a kiválasztott személyek tulajdonságainak eloszlása jól illeszkedjen a tervezett eszköz célközönségének eloszlásához, vagy pedig minél szélesebb skálát felölelve a teljes populáció eloszlását közelítse.

A rögzíteni kívánt beszédfeladatok 4 leggyakoribb típusa: a kitarított hangok (egy beszédhang hosszas hangoztatása), a felolvasott beszéd, a spontán monológ és a párbeszéd. Azt, hogy milyen típusú beszédmintákat gyűjtünk, a tervezett alkalmazás és a feladat függvényében kell megválasztani. Például depresszió esetében a kitarított hangok nem hordoznak elégséges információt a depressziós állapotról, így azok gyűjtése nem szükséges. Egy diagnosztikát támogató rendszer esetében az olvasott beszéd kézenfekvőnek tűnik, állandó a beszéd nyelvi tartalma, így a beszédproduktum milyensége lényegében kizárólag a vizsgált személytől függ. Ezzel szemben hátránya lehet, hogy kevésbé tekinthető spontánnak, ezáltal lehetséges, hogy kevésbé alkalmas

a depressziós állapot becslésére. A monológ és a párbeszéd általában spontánnak tekinthető, ezáltal jobban tükrözhetik a beszélő depressziós állapotát. Hátrányuk lehet, hogy nagyobb lesz a nyelvi tartalom és a beszédminta hosszának variabilitása, ami viszont ronthatja a gépi felismerő teljesítőképességét, illetve – éppen a spontaneitásuk következtében – megnehezíthetik az előfeldolgozást.

Az itt bemutatott eljárás esetében a depressziós személyek beszédmintáinak nehézkes gyűjtése miatt komoly kompromisszumok meghozására volt szükség a beszédatadtbázis megtervezése során, és elsősorban a beszédminták számának a maximalizálása volt a cél. A modell tanítása a „Magyar Depressziós Beszédatadtbázis” olvasott beszédmintáinak felhasználásával zajlott. Maga az adatbázis jelenleg is bővítés alatt áll, a beszédminták gyűjtését a BME LSA és a Semmelweis Egyetem Pszichiátriai és Pszichoterápiás Klinikája közösen végzik.

A minták gyűjtésénél fontos szempont volt, hogy azok a depressziós állapot súlyosságának minél szélesebb tartományát lefedjék. A depresszió súlyosságának leírásához a BDI-skálát használták. További tervezési szempont volt még, hogy az adatbázisba csak olyan személyek beszédmintái kerüljenek rögzítésre, akik esetében a depresszió mellett nem áll fenn más olyan tényező (megfázás, gyógyszerhatás, egyéb beszédet érintő betegség), amely lényegesen befolyásolná a beszédtermék minőségét.

A felvételek csendes helyiségben kerültek rögzítésre, ATR3350 típusú csipetős mikrofonnal. A mikrofon a vizsgált személy mellkasához volt rögzítve. A beszédminták 44,1 kHz-es mintavétellel és 16 bites kvantálással kerültek rögzítésre, tömörítetlen wav formátumba elmentve. A felhasznált adatbázis főbb jellemzői a 2. táblázatban láthatók.

## 2. táblázat

*A BME LSA által fejlesztett, a depressziós állapot súlyosság szerinti becslését megvalósító eljárás beszédatadtbázisának fontosabb leíró jellemzői*

	Rész- vevők (fő)	BDI (pontszám)				Életkor (év)			
		Átlag	Szórás	Mini- mum	Maxi- mum	Átlag	Szórás	Mini- mum	Maxi- mum
<b>Női beszélők</b>	136	15,7	13,7	0	50	42,9	15,4	18	75
<b>Férfi beszélők</b>	73	14,9	12,0	0	43	42,7	17,4	21	85
<b>Összesítve</b>	209	15,4	13,1	0	50	42,8	16,1	18	85

### 4.1.2. Az előfeldolgozó

Az előfeldolgozó egység végzi a beszédjellemezők kiszámolását, és a kiszámolt értékek alapján létrehozza a jellemzővektort, amely a gépi tanulóeljárás bemenete. A jellemzők kiszámolása előtt célravezető a beszédminta elejéről és végéről levágni a felesleges részeket, amik nem tartoznak a beszédfeladathoz, mint például a felvétel elején és végén lévő szüneteket, illetve

a vizsgált személy oda nem tartozó megjegyzéseit (pl. *Kezdhetem?, Végeztem*). Ezután a beszédmintákat érdemes normalizálni intenzitásuk alapján, ezzel kiszűrni a felvétel körülményeiből adódó különbségeket, mint például az eltérő erősítést, eltérő mikrofon és száj közti távolságot. Normalizálásra kétféle megoldást szoktak alkalmazni, az egyik esetben csúcsra normalizálnak a teljes beszédminta alapján, a másik lehetséges megoldás, hogy gördülő ablakot alkalmazva (vagyis annak kijelölésével, hogy a beszédminta mely részéből számítsa a gép a jellemzőt) az adott ablak által kijelölt szakasz teljesítményét normalizálják.

Ezután érdemes a jellemzők kiszámolását elvégezni. Először az alacsony szintű jellemzők kiszámolása történik meg, fix ablakmérettel és eltolással. Az ablakméret megfelelő megválasztása is fontos lépés, amelyhez két főbb szempontot kell figyelembe venni. Egyfelől minél rövidebb ablakméretet választunk, annál pontosabban az adott időpillanatra jellemző értékeket kapjuk meg, másrésztől viszont annál rosszabb lesz a frekvenciartománybeli felbontásunk. A beszédjelfeldolgozásban a szokásos ablakméretek a 20–60 ms-os tartományban mozognak. Az alacsony szintű jellemzőkből statisztikai függvények segítségével (átlag, szórás, tartomány stb.), úgynevezett származtatott jellemzők készítése szükséges, így a származtatott jellemző egyetlen értékkel jellemzi a teljes beszédmintát. Erre azért van szükség, mert egyrészt az alacsony szintű jellemzők száma igen magas, másrésztől a legtöbb gépi tanulóljárás fix méretű jellemzővektor-bemenetet vár, ezzel szemben az alacsony szintű jellemzők száma függ a beszédminta hosszától. Egyes jellemzők, mint például a tranziensarány, eleve a teljes beszédmintán értelmezettek, így ott tipikusan nem valósul meg ez a kétlépcsős eljárás. A származtatott jellemzők készítésénél lehetőség van arra, hogy ne az összes kiszámított alacsony szintű jellemző értékéből hozzuk létre azokat, hanem az alacsony szintű jellemzőértékek egy adott részalmazából. Például lehetséges megvalósítani, hogy a beszédminta *e* hangjainak átlagos *jitter*értékét számíttassuk ki. Ennek az alkalmazása javíthatja a létrehozandó modellünk pontosságát (Kiss 2019), ugyanakkor megköveteli, hogy a beszédminta annotációja (címkézése) megfelelő időpontossággal rendelkezésünkre álljon.

A bemutatott eljárás esetében a BME LSA több nyelvre alkalmazható automatikus beszédhangszintű annotáló és szegmentáló eljárásának a segítségével elkészült a beszédminták beszédhangszintű annotációja (KISS – SZTAHÓ – VICSÍ 2013). A beszédminták elején és végén lévő felesleges részeket eltávolították. A beszédminta amplitúdó szerinti csúcsra normalizálását megvalósították a beszédminta teljes hossza alapján. Az alacsony szintű jellemzők kiszámolását 10 ms-os lépésközzel végezték, az alapfrekvencia, a *jitter* és a *shimmer* esetében 50 ms széles elemzési ablakot alkalmaztak, a többi alacsony szintű jellemző esetében pedig 30 ms széles volt az elemzési ablak. A 2. fejezetben bemutatott akusztikai-fonetikai jellemzők kiszámolását elvégezték. A származtatott jellemzőket (átlag, szórás és százalékos tartomány) statisztikai függvények segítségével a teljes beszédminta alapján, illetve az egyes beszédhangok esetében is kiszámolták.

#### 4.1.3. A tanítás

A modell tanítása esetén a bemeneten egy jellemzővektor-halmazt kell megadni, és a kiválasztott gépi tanulóljárás elvégzi a modell tanítását, létrehozását, ami a fejlesztendő eljárás motorjaként funkcionál. A gépi tanulóljárások, elsősorban a mély neurális hálók megjelenése,

forradalmasította a mesterségesintelligencia-alapú eljárásokat, és segítségükkel olyan megoldások születtek, amelyek 20 évvel ezelőtt elképzelhetetlenek tűntek. A matematikai leírásuk nem bonyolult, most mégis inkább az egyenletek mellőzésével, az alapötletüket mutatjuk be.

Legkönnyebben talán a kétosztályos osztályozáson keresztül érthető meg a gépi tanuló-eljárások alapja. Ebben az esetben a tanítás során rendelkezésre állnak a tanítóminták, a célváltozójukkal (0 vagy 1) és a leíró jellemzőikkel reprezentálva. A gépi tanulóeljárás feladata, hogy a tanítóminták alapján az  $n$  dimenziós teret felossza, és felcímkézze a két osztály szerint, ahol  $n$  a leíró jellemzők száma, tetszőleges számú térrész megadásával, vagyis az  $n$  dimenziós tér két színnel (0 és 1) való kiszínezése. Ezután egy új minta kiértékelése úgy történik, hogy az eljárás megvizsgálja, hogy az az  $n$  dimenziós tér melyik részéhez tartozik, és ez alapján sorolja a megfelelő osztályba. A modell tanítása során kettős célt kell megvalósítani: törekedni kell egyfelől arra, hogy a tanítóminták minél pontosabban illeszkedjenek a megvalósított felosztáshoz; másfelől pedig arra, hogy maga a felosztás minél „egyszerűbb” legyen, ezáltal biztosítva, hogy a modellnek megfelelő legyen az általánosítóképessége. A két cél maximalizálása általában egyszerre nem megvalósítható, mivel az egyik javítása a legtöbb esetben rontja a másik teljesítményét. Általánosságban kijelenthető, hogy a tanítómintákon tökéletesen teljesítő modell rossz általánosítóképességgel bír, mivel ebben az esetben – bár a tér felosztása tökéletes a tanítóminták tükrében, azonban – a tanítóminták általában nem tartalmaznak elégséges információt a célváltozóról, emellett lehet hibás mind a célváltozójuk értéke, mind a létrehozott reprezentációjuk. Így a tökéletes felosztás feltehetőleg csak az adott tanítóhalmazon teljesít jól, de a teljes populáción már nem. Ezt a jelenséget nevezzük túltanulásnak, és ennek az elkerülése fontos a modell létrehozása során. Ez az alapgondolat kiterjeszhető többosztályos osztályozás és regressziós eljárás megvalósítására is, ahol az előbbi esetén kettőnél több osztályra kell felosztani az  $n$  dimenziós teret, az utóbbi esetén pedig „végtelen” számú térrészre, vagyis a leíró jellemzők és a célváltozó közti függvény megtalálása a feladat.

A tényleges tanítás megkezdése előtt az egyes jellemzőket célszerű azonos skálára normalizálni, erre a két leggyakrabban alkalmazott eljárás az egyes jellemzők értékeinek  $-1$  és  $1$  közé skálázása vagy a jellemzők  $z$ -normalizálása, vagyis  $0$  átlagra és  $1$  szórásra skálázása. Erre elsősorban azért van szükség, mivel az eredetileg eltérő skálájú jellemzőket, a skálájuk természeténél fogva, eltérő súllyal venné figyelembe a gépi tanulóeljárás, ami általában rontja a létrehozott modell teljesítőképességét.

További fontos lépés még a jellemzőhalmazból a megfelelő jellemzővektor kiválasztása, vagyis a kevésbé fontos jellemzők elhagyása, ami a kis méretű tanítóadat-bázisok esetében kritikus lehet a modell teljesítőképességének maximalizálás érdekében. Az összes lehetséges jellemzővektor kipróbálására általában nincs lehetőség, mivel  $n$  elemű halmaznak  $2^n$  részhalmaza van, így már igen csekély számú jellemző halmaz esetében is ( $n > 20$ ) túl sok az összes lehetséges jellemzővektor száma.

A tanítás megkezdése előtt még azt érdemes megfontolni, hogy egy közös modellt hozunk létre az összes tanítóminta alapján, vagy azokat több részre bontva valamilyen jellemzőjük mentén több modellt alkossunk meg. Ennek fontossága elsősorban szintén a kis méretű

tanítóadat-bázisok esetében van, mert bár elvileg minden metaadat számszerűsíthető, így képesek vagyunk azokat megadni a gépi tanulóeljárásnak, ugyanakkor ha a tanítóadat-bázisban nagyon eltérő, jól körülhatárolható csoportok vannak, akkor kis méretű adatbázis esetében ez a jelenség ronthatja egy közös modell teljesítőképességét.

Az utolsó lépés a gépi tanulóeljárás kiválasztása, felparaméterezése és a tanítás megvalósítása. Egyes gépi tanulóeljárások csak adott típusú speciális probléma megoldására alkalmazhatók (például lineáris probléma megoldására), illetve általában eltérő lehet a tanítási idejük és a létrehozott modelljük teljesítőképessége is.

A bemutatott eljárás esetében a modell tanítása szupportvektor-regresszió (SVR) (DRUCKER et al. 1996) gépi tanulóeljárás segítségével történt. Az SVR használata nem szükségszerű a feladat megoldásához, neurális hálókkal vagy véletlen erdővel (*random forest*) hasonló teljesítményű modelleket lehet létrehozni, azonban kis méretű ( $n < 1000$ ) tanítóhalmazok esetében az SVM-es gépi tanulóeljárások hatékonyan taníthatók, ezáltal tipikusan valamivel jobb teljesítményű modelleket lehet létrehozni általuk, ugyanannyi erőforrás használatával.

Az SVR-eljárásnak, mint ahogy a többi eljárásnak is, számos kiválasztandó, optimalizálандó paramétere van. Az egyik legfontosabb a kernelfüggvény megválasztása, amit a probléma típusa alapján célszerű kiválasztani, vagy többféle kernelt kipróbálni. A bemutatott eljárás esetében az rbf-kernel bizonyult a leghatékonyabbnak. Azonban az adott kernelnek további beállítandó paraméterei, úgynevezett hiperparaméterei vannak (pl. *cost* és *gamma*). Az optimális hiperparaméterek kiválasztását legtöbbször előre definiált értékkombinációk kimerítő kereséssel végzik, azonban nagyon sok értékkombináció kipróbálása lényegesen megnöveli a tanítási időt.

A kiszámolt jellemzőhalmazból kikerültek azok a jellemzők, amelyek nem mutattak lineáris összefüggést a depresszió súlyosságával, majd a maradék jellemzőhalmazból, az úgynevezett *fast forward selection (FFS)* eljárás (KISS 2019) segítségével kapták meg a végleges jellemzőhalmazt. Külön modellt hoztunk létre a női és a férfi beszélők esetében, mivel így jobb eredményeket lehet elérni (KISS 2019).

## 4.2. A modell tesztelése

A modell tesztelése fontos lépés már a tanítás során, hiszen gyakran a modellszelekció is már ez alapján valósul meg. A kiválasztott végleges modell jellemzése szintén lényeges, hiszen ezáltal kaphatunk képet annak használhatóságáról. A modell tesztelése során az előfeldolgozó egység megalkotja a modell teszteléséhez szükséges jellemzővektort. Majd a létrehozott modell alapján a gépi tanulóeljárás általában egy értéket (osztályt vagy valós számot) rendel a beszédmintához. A modell kiértékelése során fontos, hogy olyan mintákkal végezzük a modell tesztelését, amelyek a modell tanítása során nem játszottak szerepet.

Osztályozás esetén a modell működésének jellemzését a **tévesztési mátrix** segítségével írhatjuk le tömören és veszteségmentesen (3. táblázat).

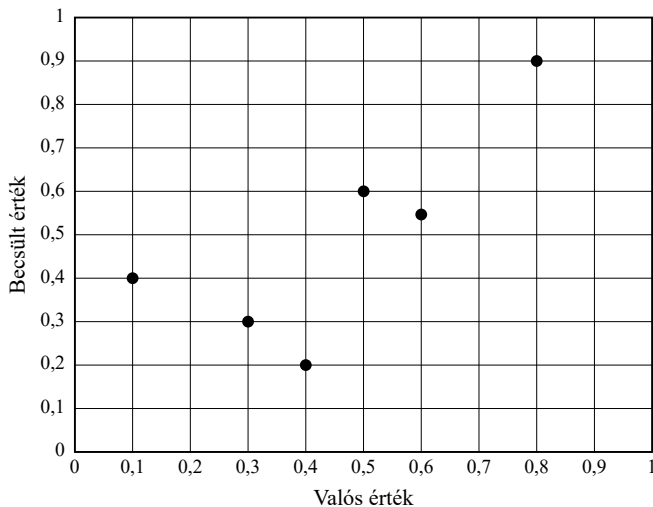
Az orvosi gyakorlatban, amikor egy betegség diagnózisa a feladat, azt kétosztályos osztályozással lehet megvalósítani. A két osztály a pozitív (beteg) és negatív (nem beteg). Ekkor a minták esetében négyféle döntés lehetséges: valós pozitív (VP), álpozitív (ÁP), valós negatív (VN) és álnegatív (ÁN) (3. táblázat).

3. táblázat

*Kétosztályos osztályozás tévesztési mátrixa az orvosi gyakorlatban*

	<b>Pozitív</b>	<b>Negatív</b>
<b>Pozitívnak döntött</b>	Valós pozitív (VP)	Álpozitív (ÁP)
<b>Negatívnak döntött</b>	Álnegatív (VN)	Valós negatív (VN)

Regresszió esetén a tévesztési mátrix nem alkalmazható, hiszen végtelen sok cellát tartalmazna, így ebben az esetben **szórásgörbén** (*scatterplot*) célszerű bemutatni a modell működését, ami a tesztminták valós és becslt értékeinek megjelenítését teszi lehetővé (3. ábra). Ebben az esetben is az átlóra ( $x = y$ ) eső minták jelzik a tökéletes becslést. Minden pont esetében leolvasható, hogy az adott minta esetében mekkora volt a valós és a becslt érték.



3. ábra

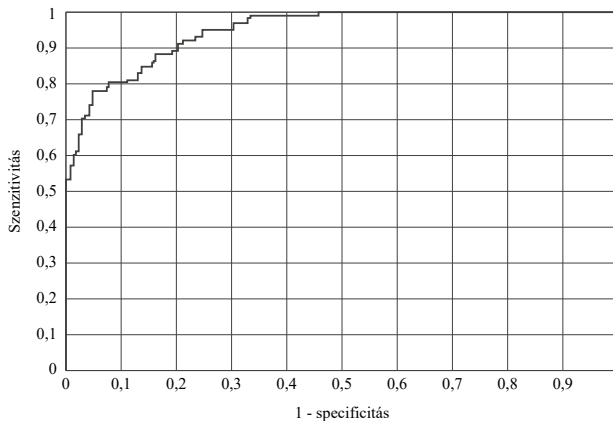
*Szórásgörbe szemléltetése*

#### 4.2.1. Kiértékelési metrikák

A **tévesztési mátrix**, illetve a **szórásgörbe** veszteségmentesen leírja a modell működését. Hátrányuk azonban, hogy két tévesztési mátrixról vagy szórásgörbéről általában nem dönthető el, hogy melyik a jobban teljesítő modell leírása. Emiatt célravezető a feladat szempontjából releváns kiértékelési metrikával (teljesítménymutató) számszerűsíteni a modell használhatóságát.

Általános osztályozási feladat tesztelése esetében a leggyakrabban a **pontosság**- (*accuracy*) metrikát alkalmazzák, ami a helyesen minősített minták számának aránya az összes tesztminta számához viszonyítva. Az orvosi gyakorlatban azonban a **szenzitivitás**metrika a szokásos elsődleges leíró jellemző, ami a helyesen pozitívnak osztályozott minták (VP) aránya az összes pozitív (VP + ÁN) mintához viszonyítva, vagyis a pozitív osztály pontossága. A szenzitivitás megadása mellett mindig szükséges a **specifitás**metrika meghatározása is, ami a helyesen negatívnak osztályozott minták (VN) aránya az összes negatív mintához (VN + ÁP) viszonyítva. Ha a tesztelés során az egyes osztályokban nem kiegyenlített a minták száma, akkor a pontosság helyett célravezetőbb az **átlagos pontosság**- (*balanced accuracy*) metrikát alkalmazni, ami kétosztályos osztályozás esetében a szenzitivitás és specifitás mértani közepe, míg az  $n$  osztályos osztályozásnál az egyes osztályok pontosságának mértani közepe.

Egy modell szenzitivitásának optimalizálása általában együtt jár a specifitásának a romlásával. A szenzitivitás- és specifitásmetrikák együttes jellemzésére ad lehetőséget a *Receiver Operating Characteristic (ROC)* görbe, ami a szenzitivitás- és az 1-specifitás-változók együttes mozgását ábrázolja, vagyis annak vizsgálata, hogy egy adott szenzitivitásérték mellett mekkora lesz a modell specifitásértéke. Az ROC-görbe szemléltetése a 4. ábrán látható. A véletlen osztályozás ROC-görbéje az  $x = y$  egyenes. Az ROC-görbe integrálásával megkapjuk a görbe alatti terület értékét (AUC), ami 0–1 tartományban változhat, és jól jellemzi egy modell működését. A véletlen osztályozás AUC-értéke 0,5; a tökéletes osztályozás értéke 1.



4. ábra

Az ROC-görbe szemléltetése

Emellett más metrikák is alkalmasak lehetnek a modell teljesítményének számszerű leírására. Fontos megjegyezni, hogy általában egy adott kiértékelési metrika maximalizálása magával vonja a többi kiértékelési metrika romlását. Emiatt az optimalizálandó kiértékelési metrika megválasztása nagyban meghatározza magának a létrehozott modellnek a használhatóságát is.

Regresszió esetében a leggyakrabban alkalmazott kiértékelési metrikák az átlagos abszolút hiba (*mean absolute error, MAE*), az átlagos hibanegyzet gyökének (*root mean square error,*



RMSE) megadása, illetve a Pearson-féle korrelációs együttható (mekkora a lineáris összefüggés a valós és a becsült értékek között). Itt is elmondható, hogy a feladat függvényében érdemes kiválasztani ezek közül az optimalizálandót, bár a feladattól függetlenül leggyakrabban az RMSE-értéket szokás kiválasztani. Érdemes megjegyezni, hogy az RMSE értéke nem lehet kisebb, mint a MAE értéke, és csak akkor egyezhetnek meg, ha a modell minden minta esetében ugyanakkora hibával becsül. Az RMSE-metrika egyszerre tartalmaz információt arról, hogy átlagosan mekkora a hiba mértéke, illetve arról, hogy a hibák nagyságának mekkora a szóródása.

#### 4.2.2. Kiértékelési módszerek

Egy adott modell használhatóságának jellemzéséhez szükséges a modell kiértékelése, valamilyen tesztmintahalmaz segítségével, a kiválasztott kiértékelési metrikák megadásával. Fontos azonban, hogy csak olyan mintával szabad tesztelni egy adott modellt, ami nem játszott szerepet a modell tanításában, ellenkező esetben a modell létrehozásakor biztosan túltanulás fog fellépni, így a modell általánosítóképessége és ezáltal a használhatósága megkérdőjelezhető, illetve a kiértékelési metrikák túlbecsülhetik a modell valódi teljesítményét.

A tanító- és a tesztalmozok megadását a legegyszerűbben úgy lehet megtenni, hogy a rendelkezésünkre álló adatbázist két részre bontjuk: tanító- és tesztelőhalmazokra. Azonban kis méretű adatbázisok esetében ez a megoldás nem alkalmazható, hiszen ezáltal tovább csökkenne a tanítóhalmazban szereplő minták száma, míg a gépi tanulójárások nagyszámú tanítómintahalmazok esetében képesek hasznos modelleket építeni. Így kis méretű adatbázisok esetében ( $n < 1000$ ) célravezetőbb  $n$ -fold keresztvalidációt (*n-fold crossvalidation*) alkalmazni. Az eljárás lényege, hogy a rendelkezésünkre álló adatbázist felosztjuk  $n$  részre, majd  $n$  darab külön modellt hozunk létre, ahol minden egyes modell létrehozásakor pontosan  $n-1$  rész kerül a tanítóhalmazba és 1 rész pedig a tesztalmozba, és minden részt pontosan egyszer használunk fel mint tesztalmozot. Ennek az előnye, hogy akár már alacsony  $n$  esetén is lényegesen több minta áll majd rendelkezésre a modell tanításához, szemben az egyszerű kettéválasztással, ráadásul az eljárásunkat a teljes halmazon tudjuk tesztelni. A hátránya, hogy nem egy modellt értékelünk ki valójában, hanem a teljes eljárásunk használhatóságáról kapunk képet.

A bemutatott eljárás esetében teljes keresztvalidáció alkalmazásával történt meg az eljárás kiértékelése, vagyis az  $n$  értéke megegyezett a rendelkezésre álló halmazok mintáinak számával.

## 5. Eredmények

A depresszió súlyosságát becslő modell létrehozása és az eljárás tesztelése az előző fejezetekben bemutatott módon valósult meg. Az eredmények vizsgálata és értelmezése előtt érdemes még egy fogalmat definiálni.

Vegyük észre, hogy osztályozás esetében egy adott pontosság értelmezhető, kézzelfogható képet ad a modell használhatóságáról. Ezzel szemben egy adott MAE- vagy egy RMSE-érték esetében ugyanez már nem mondható el. Az világos, hogy a tökéletes regresszió esetében mindkét metrika 0 értékű, azonban azt már nehezebb megmondani, hogy például egy 10 RMSE-értékű modell mennyiben használható. Mind a MAE-érték, mind az RMSE-érték elméletileg tetszőlegesen nagy lehet, és a konkrét érték jósága nagyban függ a célváltozó skálájától. Például ha az eljárásunk az emberek életkorát becsüli évben megadva, akkor a 10 RMSE-hibaérték egy elfogadható modell jellemzése lehet, ugyanez a hibaérték az emberek magasságát méterben becsülő modell esetében viszont teljesen elfogadhatatlan. Tehát a regressziós modell MAE és RMSE leíró jellemzőit mindig a használt célváltozó skálájának tartományában, szóródásában célszerű elemezni.

Emiatt érdemes definiálni a véletlen osztályozó fogalmát. Osztályozás esetében a véletlen osztályozó a vizsgált adatbázisban előforduló leggyakoribb osztályra dönt minden esetben, több leggyakoribb osztály esetén véletlenszerűen az egyikre. Így a véletlen osztályozó átlagos pontosság-értéke 50% lesz. Regressziós eljárás esetében a véletlen osztályozó a regressziós eljárás minden mintájához a célváltozó átlagát rendeli a tanítóadat-bázis alapján.

A bemutatott eljárás esetében a véletlen osztályozó RMSE-értéke 13,1; míg a MAE-értéke 11,4. A véletlen osztályozó RMSE-értéke megegyezik az adatbázisban található személyek BDI-értékének szórásával. A vizsgált modell(ek) kiértékelésénél érdemes megadni, hogy mekkora a relatív javulás a véletlen osztályozó teljesítményéhez képest, ezáltal segítve a modell használhatóságának megítélését.

A létrehozott modellek SVR típusú gépi tanulóeljárással készültek, ami regressziós feladat megoldására alkalmas, így a modellek kimenetele regresszió esetében közvetlenül vizsgálható.

A 4. táblázatban látható az optimalizált modell osztályozóképességének tesztelési eredménye, ha a becsült értékek konvertálását a BDI-skála alapján végezzük.

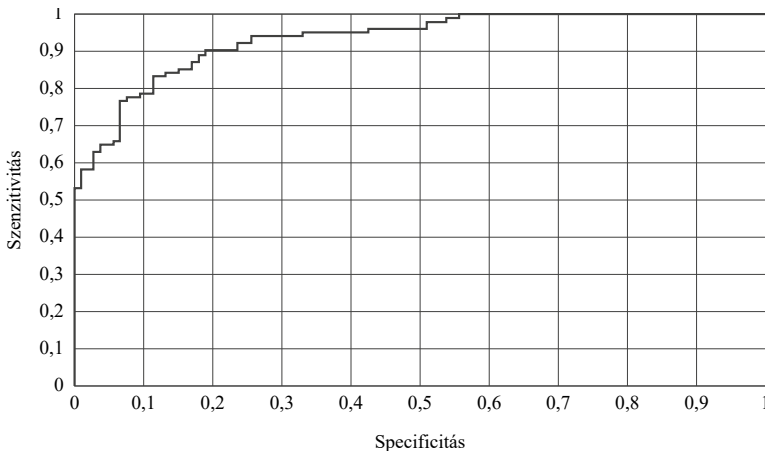
#### 4. táblázat

*A létrehozott modell osztályozóképességének tesztelése*

	Valós depressziós	Valós egészséges
<b>Depressziósnak becsült</b>	88	16
<b>Egészségesnek becsült</b>	15	90

A tévesztési mátrixból kiszámítható, hogy a modell átlagos pontosságértéke 85%, szenzitivitása 85%, és a specificitása is 85%. Vagyis az optimalizált modellt felhasználva az eljárás mind a depressziós, mind az egészséges személyek 85%-át osztályozná helyesen.

Az optimalizált modell ROC-görbéje az 5. ábrán látható, amiből az AUC-értékre 0,93 adódik.



5. ábra

*A létrehozott modell ROC-görbéje*

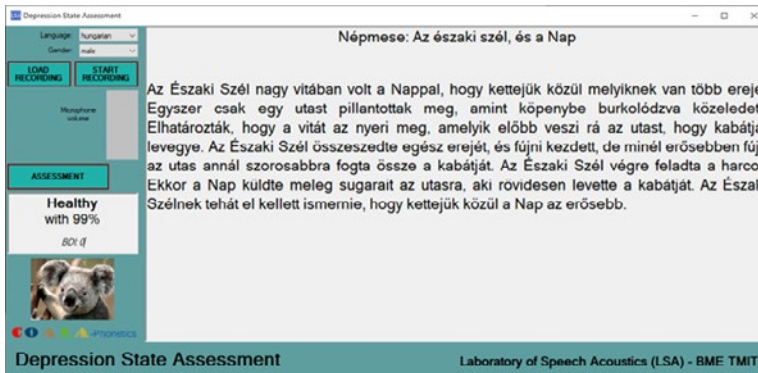
Az ROC-görbéről leolvasható, hogy ha legalább 90%-os szenzitivitást követelünk meg a rendszertől, akkor a specifitás értéke még mindig 81%-nak adódna, vagyis az egészségesnek diagnosztizált személyek közül csak minden ötödiket diagnosztizálna tévesen depressziósnak. Ha a szenzitivitásnak 99%-os szintet követelünk meg, a rendszer specifitása még akkor is 46% maradna, vagyis ha valaki valóban depressziós, akkor csak maximum 1%-ban nem figyelmeztetne a rendszer, bár ekkor már minden második egészséges embernél is figyelmeztetést adna ki.

A létrehozott modellt már megfelelő hatékonysággal lehetne előszűrő rendszerként használni, így meggyorsítva a betegség korai felismerését és a terápia megkezdését.

## 6. Alkalmazhatóság

A BME LSA megvalósította a depresszió automatikus felismerését támogató program prototípusát (*Depression State Assessment*), a European Space Agency (ESA) Coala projekt keretében. A projekt célja a Concordia kutatóállomáson áttelelő kutatók pszichológiai állapotának beszédhang-alapú megfigyelése volt, különösképpen a depressziós állapot kialakulásának jelzésére. Az állomáson dolgozó francia és olasz kutatók nagyjából kétheti rendszerességgel felolvasták *Az északi szél és a Nap* című mesét az anyanyelvükön. A program használatának az elején (6. ábra) be kell állítani a beszélő nemét és anyanyelvét, majd elvégezni a felolvasást, és az alkalmazás kimenetként megadja a beszélő becsült depressziós súlyosságát a BDI-skála szerint, illetve százalékban kifejezve annak a valószínűségét, hogy depressziótól szenved. Ennek a számítási ideje körülbelül

ugyanannyi, mint amennyi idő alatt elvégzi a felolvasást a vizsgált személy, vagyis körülbelül 1 percig tart a művelet.



6. ábra

*A depresszió automatikus felismerését megvalósító prototípus felhasználói felülete*

A modellek tanítása a 4. fejezetben bemutatott eljárással zajlott magyar, német és olasz nyelvű beszédminták felhasználásával. A prototípus nemcsak a Coala projektben használható, hanem jól alkalmazható előszűrésre a háziorvosi rendelőkben vagy otthoni környezetben önvizsgálatra a depresszió fennállásának figyelmeztetésére. Természetesen a program nem helyettesíti a pontos diagnózis felállítását, azonban segíthet a depressziós személyek felderítésében.

A program használatakor fontos több szempontot is figyelembe venni. Ilyen például, hogy a modellek egészséges és depressziós személyek beszédmintái alapján lettek létrehozva, így nem becsülhető pontosan, hogy milyen értéket adna vissza olyan személyek esetében, akik bár nem depressziósak, de olyan betegségtől szenvednek, amely hatással bír a beszédproduktumukra. Emellett a beszédproduktumra hatással van a vizsgált személy életkora, érzelmi állapota, fáradtsága stb., amelyek mind torzíthatják a kapott eredményt. Maga a program nem kíván nagy fokú kooperációt a vizsgált személytől, azonban ha nem megfelelően állítja be az anyanyelvét, nemét, illetve nem a kapott szöveget olvassa fel, akkor az eredmények sem tekintendők helytállóknak.

## 7. Kitekintés

A bemutatott eljárás alkalmas arra, hogy a vizsgált személy depressziós állapotát becsülje. Ugyanakkor a tanulmány terjedelme miatt több tervezési kérdést csak érinteni lehetett, vagy egyáltalán nem is esett szó róluk. A teljesség igénye nélkül álljon most itt három olyan kérdés,

amelyek aktuális kutatási problémák, és azok megválaszolása releváns egy valós előszűrő rendszer megtervezésében.

1. Vajon lehetséges-e olyan rendszert létrehozni, amely egyszerre több betegség felismerését valósítja meg? Mekkora pontosságú lehet egy ilyen rendszer (SZTAHÓ és mtsai 2018)?
2. Az olvasott vagy a spontán beszéd alkalmasabb-e a depresszió felismerésére (KISS–VICSÍ 2017a)?
3. Lehetséges-e több nyelvre érvényes eljárást kidolgozni? Mekkora pontossággal valósítható meg egy ilyen rendszer? Felhasználhatóak-e más nyelvű depressziós beszédatadatok mintái egy adott nyelvű eljárás javítására (KISS–VICSÍ 2017b)?

## FOGALMAK

*alapfrekvencia; formánsfrekvencia; formánsfrekvencia sávszélessége; intenzitásint; jütter; shimmer; pontosság (átlagos pontosság); specificitás; szenzitivitás; szórásgörbe; tévesztési mátrix*

## IRODALOM

- ALPERT, Murray – POUGET, Enrique R. – SILVA, Raul R. 2001. Reflections of depression in acoustic measures of the patient's speech. *Journal of Affective Disorders* 66/1. 59–69.
- BECK, Aaron T. – STEER, Robert A. – BALL, Roberta – RANIERI, William F. 1996. Comparison of Beck Depression Inventories-IA and -II in psychiatric outpatients. *Journal of Personality Assessment* 67/3. 588–597.
- CUMMINS, Nicholas – SCHERER, Stefan – KRAJEWSKI, Jarek – SCHNIEDER, Sebastian – EPPS, Julien – QUATIERI, Thomas F. 2015. A review of depression and suicide risk assessment using speech analysis. *Speech Communication* 71. 10–49.
- DRUCKER, Harris – BURGESS, Christopher J. – KAUFMAN, Linda – SMOLA, Alex – VAPNIK, Vladimir 1996. Support vector regression machines. *Advances in Neural Information Processing Systems* 9. 155–161.
- DWIVEDI, Yogesh – RIZAVI, Hooriyah S. – CONLEY, Robert R. – ROBERTS, Rosalinda C. – TAMMINGA, Carol A. – PANDEY, Ghanshyam N. 2003. Altered gene expression of brain-derived neurotrophic factor and receptor tyrosine kinase B in postmortem brain of suicide subjects. *Archives of General Psychiatry* 60/8. 804–815.
- FRIEDRICH, Mary Jane 2017. Depression is the leading cause of disability around the world. *Jama* 317/15. 1517.
- GOSZTOLYA, Gábor – BUSA-FEKETE, Róbert – GRÓSZ, Tamás – TÓTH, László 2017. DNN-based feature extraction and classifier combination for child-directed speech, cold and snoring identification. In *Interspeech 2017*. 3522–3526. [https://www.isca-speech.org/archive/Interspeech\\_2017/pdfs/0905.PDF](https://www.isca-speech.org/archive/Interspeech_2017/pdfs/0905.PDF) (A letöltés ideje: 2021. január 6.)
- HAMILTON, Max 1986. The Hamilton rating scale for depression. In SARTORIUS, Norman – BAN, Thomas A. (eds): *Assessment of Depression*. Springer Verlag, Berlin–Heidelberg. 143–152.

- HAWTON, Keith – CASAÑAS I COMABELLA, Carolina – HAW, Camilla – SAUNDERS, Kate 2013. Risk factors for suicide in individuals with depression: A systematic review. *Journal of Affective Disorders* 147/1–3. 17–28.
- HORII, Yoshiyuki 1982. Jitter and shimmer differences among sustained vowel phonations. *Journal of Speech, Language, and Hearing Research* 25/1. 12–14.
- JACKSON, Stanley W. 1986. *Melancholia and Depression: From Hippocratic Times to Modern Times*. Yale University Press, New Haven.
- KISS, Gábor – SZTAHÓ, Dávid – VICSÍ, Klára 2013. Language independent automatic speech segmentation into phoneme-like units on the base of acoustic distinctive features. In *4th International Conference on Cognitive Infocommunications (CogInfoCom)*, December 2013, Budapest. 579–582.
- KISS, Gábor – VICSÍ, Klára 2017a. Comparison of read and spontaneous speech in case of automatic detection of depression. In *8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, September 2017, Debrecen. 213–218.
- KISS, Gábor – VICSÍ, Klára 2017b. Mono-and multi-lingual depression prediction based on speech processing. *International Journal of Speech Technology* 20/4. 919–935.
- KISS, Gábor 2019. *A depressziós beszéd akusztikai-fonetikai jellemzőinek vizsgálata*. PhD-értekezés. BME VIK, Budapest.
- KLEIN, Daniel N. 2008. Classification of depressive disorders in the DSM-V: Proposal for a two-dimension system. *Journal of Abnormal Psychology* 117/3. 552.
- KOVÁCS, Annamária – KISS, Gábor – VICSÍ, Klára – WINKLER, István – COATH, Martin 2015. Comparison of skewness-based salient event detector algorithms in speech. In *6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, October 2015, Győr. 285–290.
- KRAEPELIN, Emil 1921. Manic depressive insanity and paranoia. *The Journal of Nervous and Mental Disease* 53/4. 350.
- LOW, Daniel M. – BENTLEY, Kate H. – GHOSH, Satrajit S. 2020. Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope Investigative Otolaryngology* 5/1. 96–116.
- LI, Xiang – TAN, Nanlin – WANG, Tianlei – SU, Shuqiang 2014. Detecting driver fatigue based on nonlinear speech processing and fuzzy SVM. In *12th International Conference on Signal Processing (ICSP)*, October 2014, Hangzhou. 510–515.
- MATHERS, Colin D. – LONCAR, Dejan 2006. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Medicine* 3/11. e442. <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0030442>
- OLASZY Gábor 2010a. A beszédképzés folyamata. In NÉMETH Géza – OLASZY Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémia Kiadó, Budapest. 19–26.
- OLASZY Gábor 2010b. A beszéd szupraszegmentális szerkezete. In NÉMETH Géza – OLASZY Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémia Kiadó, Budapest. 171–205.

- OLESEN, Jes – GUSTAVSSON, Anders – SVENSSON Mikael – WITTCHEM, H. U. – JÖNSSON, Bengt – CDBE 2010 Study Group, and European Brain Council 2012. The economic cost of brain disorders in Europe. *European Journal of Neurology* 19/1. 155–162.
- PREISS, Kymerlie – BRENNAN, Leah – CLARKE, David 2013. A systematic review of variables associated with the relationship between obesity and depression. *Obesity Reviews* 14/11. 906–918.
- SATORI, Hassan – ZEALOUK, Ouissam – SATORI, Khalid – ELHAOUSSI, Fatima 2017. Voice comparison between smokers and non-smokers using HMM speech recognition system. *International Journal of Speech Technology* 20/4. 771–777.
- SZTAHÓ, Dávid – TULICS, Miklós G. – VICSÍ, Klára – VALÁLIK, István 2017. Automatic estimation of severity of Parkinson’s disease based on speech rhythm related features. In *8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, September 2017, Debrecen. 11–16.
- SZTAHÓ, Dávid – KISS, Gábor – TULICS, Miklós G. – VICSÍ, Klára 2018. Automatic separation of various disease types by correlation structure of time shifted speech features. In *41st International Conference on Telecommunications and Signal Processing (TSP)*, July 2018, Athen. 1–4.
- VASHKEVICH, Maxim – AZAROV, Elias – PETROVSKY, Alexander – RUSHKEVICH, Yuliya 2018. Features extraction for the automatic detection of ALS disease from acoustic speech signals. *Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, September 2018, Poznan. 321–326.
- VETRÁB Mercedes – GOSZTOLYA Gábor 2019. Érzelmek felismerése magyar nyelvű hangfelvételekből akusztikus szózsák jellemzőreprezentáció alkalmazásával. In BEREND Gábor – GOSZTOLYA Gábor – VINCZE Veronika (szerk.): *XV. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem Informatikai Intézet, Szeged. 265–274.
- VICSÍ Klára 2010a. A beszéd fizikai jellemzése. In NÉMETH Géza – OLASZY Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémia Kiadó, Budapest. 39–56.
- VICSÍ Klára 2010b. Pszichofizikai tényezők. In NÉMETH Géza – OLASZY Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémia Kiadó, Budapest. 56–69.
- ZHANG, Yue – WENINGER, Felix – SCHULLER, Björn W. 2017. Cross-domain classification of drowsiness in speech: The case of alcohol intoxication and sleep deprivation. In *Interspeech 2017*. 3152–3156. [https://www.isca-speech.org/archive/Interspeech\\_2017/pdfs/1015.PDF](https://www.isca-speech.org/archive/Interspeech_2017/pdfs/1015.PDF) (A letöltés ideje: 2021. január 10.)

## KÖSZÖNETNYILVÁNÍTÁS

A K1285968 számú projekt a Nemzeti Kutatási Fejlesztési és Innovációs Alapból biztosított támogatással, a K 18 pályázati program finanszírozásában valósult meg.

# A nyelv és a beszéd sajátosságai neurodegeneratív kórképekben

Hoffmann Ildikó

NYELVTUDOMÁNYI KUTATÓKÖZPONT  
SZEGEDI TUDOMÁNYEGYETEM

Svindt Veronika

NYELVTUDOMÁNYI KUTATÓKÖZPONT

Bóna Judit

ELTE EÖTVÖS LORÁND TUDOMÁNYEGYETEM

## 1. Neurodegeneratív betegségek

A **neurodegeneratív betegségek** a központi idegrendszert (vagyis az agyat és a gerincvelőt) érintő krónikus betegségek összefoglaló elnevezése. Jelenlegi tudásunk szerint ezek a betegségek nem gyógyíthatók, mivel az idegsejtek progresszív (folyamatosan zajló) leépülésével és/vagy elhalásával járnak együtt. A leggyakrabban előforduló neurodegeneratív betegségek: az Alzheimer-kór (AK); a Parkinson-kór (PK); a szklerózis multiplex (lat. *sclerosis multiplex*, SM); a Huntington-kór (HK); valamint az amiotrófiás laterálszklerózis (ALS).

A neurodegeneratív betegségek hátterében egyaránt állhatnak genetikai és környezeti tényezők. Minden neurodegeneratív betegség közös mechanizmusa, hogy kóros, sejten kívüli vagy sejten belüli fehérjecsoportosulások (ún. fehérjeaggregátumok) jönnek létre. Miközben a fehérjeaggregátumok típusa több neurodegeneratív betegség hátterében hasonló, a szakembereket a diagnosztizálás folyamán az idegsejtpusztulás tipikus helye orientálja, vagyis ezek a betegségek jellemzően adott agyterületeken végbemenő sejtpusztulással jellemezhetők, azaz régióspecifikusak (PENKE–DATKI–ZARÁNDI 2003).

A kognitív és a nyelvi rendszer változásai szoros összefüggésben állnak egymással. A neurodegeneratív betegségeket számos tünet jellemezheti (KEMPLER 2004). A kognitív folyamatok lelassulnak, az összetett, bonyolult kognitív működésért felelős ún. **végrehajtó funkciók** kapacitása korlátozódik. Végrehajtó funkcióknak nevezzük összefoglalóan a magasabb szintű kognitív működéseket, amelyeknek körébe tartozik például a tervezés, a döntéshozás, a következtetés, a váltás és válaszgátlás, a kognitív flexibilitás, valamint a **munkamemória**. A munkamemória a nyelvi működésért is felelős, korlátozott kapacitású emlékezeti rendszer, amely hozzáférést biztosít az eltérő reprezentációkhoz, fenntartja és manipulálja az információt, és független információkat hoz interakcióba.



Neurodegeneratív kórképekben megváltozhat a beszéd, **afáziára** (agysérülés, pl. stroke, trauma, tumor következtében előálló nyelvi és beszédzavarok összefoglaló neve; BÁNRÉTI 2014) jellemző tünetek jelentkezhetnek a nyelvi tevékenységben: **fonemikus parafázia** (a célszóhoz hangzásában hasonló, de jelentéssel nem rendelkező szó, pl. *piros* helyett *kiros*; SZÉPE 2009), **szemantikus parafázia** (a célszó helyett a kategória egy másik elemének kiválasztása, pl. *kalapács* helyett *csavarhúzó*), **diszfluencia** (folyamatában akadályoztatott beszéd), **agrammatizmus** (grammatikai vagy morfológiai hibázás a mondatban, pl. *a macska kergeti az egér*). Megértési zavarok is előfordulhatnak. Az egyik legjellemzőbb nyelvi tünet a szótalálási probléma, melynek hátterében a szemantikus emlékezet zavara áll. Változás következhet be az írott nyelv felismerésében és produkciójában is **alexia** (olvasási zavar) és **agráfia** (írászavar, vagyis a fonéma–graféma konvertálás zavara) formájában. A jobb félteke bizonyos területein bekövetkező változások pragmatikai és kommunikációs deficiteket eredményezhetnek, melyek érintik a prozódiaát, valamint a nem szó szerinti nyelvet (pl. idiómák, ironikus és szarkasztikus kifejezések megértése), illetve magukba foglalhatják a társalgásban való részvétel szabályainak figyelmen kívül hagyását és/vagy meg nem értését. Jelentkezhet továbbá akalkulia (számolási zavar), téri-vizuális zavar, agnózia (a tárgyak és emberek felismerésének zavara), auditoros és taktilis deficit, diszfunkció a végrehajtó rendszer működésében (fluencia, megosztott figyelem, előtervezés, döntéshozás). Az emlékezet sérülései különböző memória-alrendszereket érinthetnek, melyek részt vesznek a nyelvi folyamatokban, így a rövid távú, a hosszú távú, azon belül például a **procedurális** (mentális nyelvtan működtetéséért felelős) vagy a **deklaratív** (a szavak tárolásáért felelős) emlékezetet. Megjelenhet **demencia**, azaz időskori kóros kognitív leépülés, s ezzel párhuzamosan hangulat-zavarok (pl. depresszió) léphetnek fel. Olyan nem kognitív pszichés tünetek is előfordulhatnak, mint a hallucináció, a személyiségváltozás, az agresszivitás vagy a téveszmék (TARISKA 2000). Mindezeket bizonyos esetekben a mozgás zavarai kísérhetik.

Demenciában folyamatos zavar van jelen a mentális funkciók legalább három területén, melyek a következők lehetnek: emlékezet, nyelv, téri-vizuális készség, absztrakció, számolás, ítélőképesség, érzelmi állapot és személyiség (CUMMINGS–MEGA 2003). A különböző funkciók egy korábbi magasabb szintjéhez képest változás, hanyatlás következik be (DSM-5, APA 2013). A demenciát szórt (ún. diffúz) vagy többgócós agyi diszfunkciók kísérik. A változások során az emlékezet és a nyelv szoros kapcsolatát jól nyomon követhetjük. A demenciának sokféle típusa és besorolási lehetősége van (TÚRI–NÉMETH–HOFFMANN 2014). Míg a degeneratív demenciák folyamata nem megállítható (pl. Alzheimer-kór, Parkinson-kór, Huntington-kór), a nem degeneratív demenciák folyamatos romlására hatással lehetünk, lassíthatók, sőt vissza is fordíthatóak. A legáltalánosabb felosztás szerint két nagy csoportot alkotnak, mely alapján az agykéreg érintő, ún. kortikális (Alzheimer-kór, frontotemporális demencia) és az agykéreg alatti területeket érintő, ún. szubkortikális demenciáról (Parkinson-kór, Huntington-kór) beszélhetünk.

A neurodegeneratív betegségek többségében olyan agyterületek érintettek, amelyek nem tartoznak az általánosan ismert „klasszikus” nyelvi folyamatokért felelős régiók közé. Ennek ellenére az alább bemutatott betegségekben gyakran számos nyelvi és beszédfolyamat is érintett lehet. Ez az agy hálózatos felépítéséből fakad, melynek következménye a nyelvi folyamatoknak

más kognitív folyamatokkal való szoros összefüggése és együttműködése. A nyelvi funkciók vizsgálatának fontos szerepe van a különböző neurodegeneratív zavarokban megjelenő korlátozottságok feltárásában (GARRARD et al. 2005), az állapotváltozások követésében, és segítségre lehet a diagnózis során is.

Az alábbiakban elsőként röviden bemutatunk néhány olyan neurodegeneratív zavart, melyben a nyelvi változások jól követhetőek, a későbbiekben pedig az Alzheimer-kórt, a Parkinson-kórt és a szklerózis multiplexet ismertetjük részletesen.

**Primer progresszív afázia (PPA)** során első tünetként (erre utal a primer jelző) progresszív, vagyis folyamatosan súlyosbodó beszédzavar jelenik meg, megelőzve minden más kognitív tünetet (MESULAM 1982; GORNO–TEMPINI–MILLER 2013), majd körülbelül két évet követően demencia alakul ki. Az érintett agyterület és a tünetek alapján a PPA és az Alzheimer-kór (AK) nagyon hasonló (vö. szóalálási nehézség), ugyanakkor a különbséget az adja, hogy míg AK-ban a memória és a nemverbális kogníció sérül, PPA-ban ezek épek maradnak, s a zavar az első két évben csak a beszédet érinti (KERTÉSZ 2004). Két típusát különböztetjük meg: a fluens (folyamatos) és a nonfluens (nem folyamatos) PPA-t. A fluens primer progresszív afáziát szemantikus demenciának is nevezzük. Az artikuláció, a fonológia és a szintaxis ép marad, az akusztikus-fonetikus feldolgozás kismértékben érintett. Sérül a lexikai-szemantikai háló, a lexikon elemeinek elérése, problémák jelentkeznek az auditoros lexikai és a szemantikai hozzáférésben (HODGES–MILLER 2001). A betegek beszédében szemantikus parafáziák és szemantikus anómia (szóalálási nehézség, mely főleg a főneveket érinti, pl. a páciens *asztal* helyett *széket* mond) jelentkeznek. Redukálódik a beszéd megértése. Sok tünetében hasonlóságot mutat az agykéreg hátulsó (ún. posterior) területeit érintő afáziával, a különbség abból adódik, hogy míg fluens PPA-ban a kifejezések (frázisok) szintjének a megértése ép marad, addig ez posterior típusú afáziában sérül. Ezzel szemben a nonfluens primer progresszív afázia vezető tünete a fonológiai parafázia, kísérő tünete a **dizartéria**, vagyis olyan hangképzési és beszédzavar, ami a beszédképző szervek mozgató beidegzésének agyi eredetű sérülése következtében alakul ki. A betegek beszédmegértési nehézséggel küzdenek, melynek oka az akusztikus-fonetikus feldolgozás, a rövid távú auditoros emlékezet és a fonológiai bemeneti lexikon sérülése (CLARK et al. 2005). Sok tünetben egyezik az anterior (elülső agyterületi) afáziával, de nonfluens PPA-ban az agrammatizmus nem jellemző.

A **frontotemporális demencia (FTD)** vagy Pick-betegség (KERTÉSZ 2004) klinikai szempontból nagyon hasonlít az Alzheimer-kórhoz, ugyanakkor míg az AK kezdeti szakaszában az emlékezet hanyatlása az egyik vezető tünet, addig FTD-ben a személyiség megváltozása (TARISKA 2000). **Perszeverációs** (pl. szavak, cselekedetek céltalan ismételtetése) és sztereotip magatartás (pl. járkálás, dobolás a szorongás csökkentésére) jelentkezik. A beszéd megértése redukálódik. A mondatmegértési zavar elsődleges oka a mondat központi elemének, az igének az elérhetetlensége. A beszédprodukción lelassul, a prozódia monotonná válik, valamint megnevezési zavar lép fel. A beszédet gyakran kíséri dizartria, mely súlyos esetben mutizmusig (némaság) vezethet (KERTÉSZ 2004).

A **Huntington-kór (HK)** tisztán genetikai eredetű, ritka, dominánsan öröklődő progresszív neurodegeneratív betegség (KREMER et al. 1994; RAINER–DRAGATSI–DIETRICH 2011;

RUBINSZTEIN et al. 1996). Neurológiai háttérben a nagyagy fehérállományának leépülése, valamint az információ szervezésében, irányításában, összekötésében és továbbításában szerepet játszó bizonyos agykéreg alatti területek leépülése áll. Előfordulásának gyakorisága földrajzilag változó. Tipikus indulása 40–50 éves kor közöttre tehető, de a betegek mintegy 5%-ánál gyermekkorban is kezdődhet (SZIRMAI 2006). Mozgásos (ún. motoros) tünetek, lelassult információfeldolgozás, pszichopatológiai zavarok és kognitív hanyatlás jellemzi. A betegség együtt jár depresszióval és a személyiség változásával, előrehaladtával pedig minden esetben demencia kíséri (SZIRMAI 2006). A betegség kezdeti és egyik vezető tünete az arcizmok szabálytalan és nem akaratlagos mozgásai, görcsei és grimaszba rándulásai, a nyelés nehezítettsége, mely szükségszerűen mind a hangképzési, mind a beszédképzési folyamatokra jelentős hatással van. Lelassul a beszéd, megnő a szünettartás ideje és gyakorisága, sérül az intonáció, szótalálási és szóproduktions zavarok lépnek fel, parafáziák, kényszeres túltoldalékolás, felesleges szabályalkalmazás és agrammatizmus jelennek meg a beszédprodukció során.

A továbbiakban a neurodegeneratív zavarok közül hármat, az Alzheimer-kórt, a Parkinson-kórt és a szklerózis multiplexet mutatjuk be részletesen, ismertetjük a beszédtevékenység során bekövetkező változásokat, majd kitérünk a nyelvészeti vizsgálatok szerepére a diagnosztikában és a terápiában.

## 2. Alzheimer-kór (AK)

Az **Alzheimer-kór (AK)** egy lassú, folyamatos kognitív károsodás. Kialakulásának pontos oka nem ismert. A neurodegeneratív folyamat során az agykéregben található ún. piramis sejtek és az összekötő idegsejtek (ún. interneuronok) fokozatosan elveszítik az információ továbbítását végző (ún. axonok), valamint az információt fogadó nyúlványaikat (ún. dendritek). Mivel az idegsejtek (neuronok) ezeken a nyúlványokon keresztül kapcsolódnak össze, és hoznak létre ún. szinaptikus kapcsolatokat, a károsodás éppen ezeket érinti: a korábban kialakult szinapszisok megszűnnek, az információ továbbítása ellehetetlenül (BECKER–GIACOBINI 1996). A szinaptikus kapcsolatok progresszív pusztulása elsősorban az agykérgi és a kéreg alatti (ún. bazális előagy) neuronhálózatot és összeköttetéseit érinti, mely a tanulási, az emlékezési és a végrehajtó működési folyamatokért felelős, de a teljes neuronhálózat működésére is kihat (MARKESBERY 1998; TÚRI–NÉMETH–HOFFMANN 2014). Indulását tekintve beszélhetünk korai (40–45 év) és késői (65–70 év) kezdetről. A betegségben enyhe, közepes és súlyos szakaszt különböztetünk meg.

Az Alzheimer-kór kialakulását enyhe kognitív hanyatlás előzi meg, melyet enyhe kognitív zavarnak vagy Alzheimer-prodrómának hívunk. **Enyhe kognitív zavarban (EKZ)** a különböző mentális funkciók egy megelőző magasabb szintjéhez képest tapasztalható enyhe változás, mely érintheti a rövid távú emlékezetet, a szókeresést vagy a figyelmi folyamatokat is. Az EKZ az egészséges és a demens állapot közötti skálán jelenik meg, előjele lehet egy később kialakuló

demenciának (TÚRI–NÉMETH–HOFFMANN 2014). Az emlékezeti rendszerek közül az **implicit** és a **procedurális memória** ép marad. Az implicit memória olyan emlékezeti rendszer, amely az előhívás során nem igényel tudatos erőfeszítést. Itt tárolódnak a készségszintű, automatikussá vált ismereteink, valamint a mentális nyelvtan, vagyis a nyelvtani szabályok (és kivételek) összessége. A tanulmányok egy része érintettnek találja a munkamemóriát és a szemantikus memóriát, más részük szerint ezek intaktak (nem érintettek). Az eredmények attól is függenek, hogy milyen eljárásokkal történik a vizsgálati személyek tesztelése. Az EKZ-ban megjelenő nyelvi tüneteket az Alzheimer-kór nyelvi változásaival együtt ismertetjük.

Az Alzheimer-kór első nyelvi deficitjei nagyon enyhék, ezért nehéz őket megkülönböztetni a normál időskorban is bekövetkező emlékezeti problémáktól és szóalálási nehézségektől (MORRIS 1996). Később a kezdeti szimptómák felerősödnek, progresszív zavarok jelentkeznek, melyek érintik a különböző nyelvi szinteket, az emlékezeti folyamatokat, a mozgást, a téri-vizuális készségeket és a végrehajtó funkciókat. A beszédtevékenységet a beszélt és az írott nyelv megértési és produkciós zavarai, szókeresési és szóalálási problémák, fonemikus és szemantikus parafáziák jellemzik, kiüresedik a beszéd, vagyis a beteg beszédében túlsúlyba kerülnek a funkciószavak (pl. kötőszók), és „elfogynak” a tartalmas szavak, miközben az ismétlési (utánmondási) képesség relatíve sokáig érintetlen marad (BAYLES–TOMODEA–TROSSET 1992; APPELL–KERTESZ–FISHMAN 1982; CROOT et al. 2000). A jól működő nyelvi funkciók háttérben jól működő emlékezeti folyamatokat feltételezünk. A kór lefolyása során a nyelv és az emlékezet funkcióinak szoros összefüggéseit jól nyomon lehet követni: a beszédmegértés, a beszédprodukció és az emlékezeti funkciók nehézségei átfedik egymást (HOFFMANN–SZATLÓCZKI–KÁLMÁN 2016).

Az AK során végbemenő nyelvi változások vizsgálhatók a betegség szakaszait tekintve és nyelvi szintenként is. Az AK enyhe szakaszában megnő a hezitációk (kitöltött szünetek, pl. *ő*) száma a spontán beszédben, szófelidézési zavarok, szóalálási problémák, fonemikus parafáziák jelentkeznek, ugyanakkor a szintaxis és a társalgási struktúrák érintetlenek maradnak. A középsúlyos szakaszban az anómia súlyosbodik, fonemikus és szemantikus parafáziák jelennek meg, a társalgási struktúrák megértése és produkálása időnként már zavart szenvedhet, miközben a szintaxis továbbra sem érintett. Súlyos szakaszban a procedurális emlékezet erős deficitje miatt agrammatizmus lesz jellemző, a mondatok nem interpretálhatók, vagy teljesen üresek, a parafáziák felerősödnek, a társalgási struktúrák megértése és produkálása erősen zavart, a beteg a társalgási partnert vagy önmagát ismétleti, de gyakori a mutizmus is.

A nyelvi szintek érintettsége különböző módon nyilvánul meg a betegség lefolyása során (BAYLES–BOONE 1982; SZATLÓCZKI et al. 2015). A fonetikai-fonológiai szinten már EKZ-tól kezdve változások következnek be a beszéd temporális jellemzőiben. Enyhe lassúbbodás figyelhető meg a beszédtempóban, a hezitációk (kitöltött szünetek) száma és időtartama megnövekszik (ROARK et al. 2011; SATT et al. 2014; JARROLD et al. 2014; TÓTH et al. 2018; GOSZTOLYA et al. 2019; VINCZE et al. 2020). A beszéd temporális változása az EKZ egyik legkorábbi tünete.

A spontán beszéd időviszonyainak változását AK-ban több kutatás is vizsgálta (HOFFMANN et al. 2010; LÓPEZ-DE-IPINA et al. 2013; MEILÁN et al. 2012; SATT et al. 2014; JARROLD et al. 2014). A betegség előrehaladtával az artikulációs és a beszédtempó fokozatosan lassul,

a szünetek száma és időtartama megnövekszik. Míg az enyhe szakaszban a hezitálások, addig a középsúlyos és súlyos szakaszban a néma szünetek jellemzőek. A betegek már a kezdeti szakaszban is lassabban beszélnek, az egészségesekhez képest gyakrabban és hosszabb szüneteket tartanak, a kimondandó szó megtalálásához több időre van szükségük, mely diszfluenciát okoz (SZATLÓCZKI et al. 2015). A spontán beszédben fonemikus parafáziák is jelentkezhetnek (CROOT et al. 2000; FORBES–VENNERI–SHANKS 2002; ROARK et al. 2011; WUTZLER et al. 2013; SATT et al. 2014; JARROLD et al. 2014).

A változásban az olvasási folyamatok is érintettek. Hangos olvasás során lassul az artikulációs és a beszédtempó, több időre van szükség a hangképzéshez, megnő a szünettartás száma és ideje is (MARTINEZ-SÁNCHEZ et al. 2013). A fonetikai-fonológiai szinten bekövetkező változásokat az *1. táblázat* foglalja össze (HOFFMANN–SZATLÓCZKI–KÁLMÁN 2016).

### 1. táblázat

*A nyelvi funkciók változása fonetikai-fonológiai szinten az enyhe kognitív zavarban (EKZ) és az Alzheimer-kór (AK) különböző stádiumaiban (HOFFMANN–SZATLÓCZKI–KÁLMÁN 2016 nyomán):*

*+ = a sérülés mértéke*

<b>A nyelvi jellemzők változása</b>	<b>EKZ</b>	<b>enyhe AK</b>	<b>középsúlyos AK</b>	<b>súlyos AK</b>
temporális változások a spontán beszédben (hezitációk számának és idejének megnövekedése)	+	+	++	+++
fonemikus parafázia	+	+	++	+++

A lexikai-szemantikai szinten lévő változások is korán jelentkeznek (*2. táblázat*). A szemantikus emlékezet felelős például a megnevezés, a tárgyfelismerés és a verbális fluencia (egy adott szabály szerinti szóaktiválás megadott időn belül, a feladat ismertetését lásd az *5. fejezetben*) épségéért. A szemantikus emlékezet zavara miatt korlátozódik a mentális lexikonhoz való hozzáférés, így a szótalálásra fordított idő megnövekedik, ami először beszédproduktív zavarhoz vezet, később pedig megértési nehézségeket is okoz (KEMPLER 2004). Már EKZ-val élő személyeknek is problémát okozhat spontán beszéd során a megfelelő szavak megtalálása (FRASER et al. 2014; GARRARD et al. 2014), a verbálisfluencia-tesztekben gyengébben teljesítenek, mint az egészséges kontrollszemélyek (BALTHAZAR et al. 2007).

Az AK egyik legkorábbi és legszembetűnőbb nyelvi problémája a szótalálási zavar és a szemantikai hibázások (CROOT et al. 2000; DOS SANTOS et al. 2011; FRASER et al. 2014; LASKE et al. 2014; GARRARD et al. 2014). Szemantikus parafázia során jellemző hiba a célszó helyett egy fölérendelt kategória kiválasztása (pl. *kutya* helyett *állat*) (EMERY 2000). A betegeknek nehézsége van a képmegnevezésben és a konfrontációs képkiválasztásban (pl. négy képből a célszónak/célmondatnak megfelelő kiválasztása) (PETERSEN et al. 1999, 2001; RITCHIE et al. 2001; DUONG et al. 2006). Lassul a lexikai döntés (GIFFARD et al. 2001, 2002), a szemantikai priming (előfeszítési) feladatban a reakcióidők megnövekednek (GIFFARD et al. 2001, 2002;

DUONG et al 2006). A fonemikus- és a szemantikusfluencia-feladatok során a betegek kevesebb elemszámot tudnak produkálni, mint a kontrollbeszélők (APPELL–KERTESZ–FISHMAN 1982; BAYLES–KASZNIK–TOMODEA 1987; BARTH et al. 2005; JUNCOS-RABADÁN et al. 2010; ROARK et al. 2011; SATT et al. 2014; JARROLD et al. 2014).

2. táblázat

*A nyelvi funkciók változása lexikai, szemantikai szinten az enyhe kognitív zavarban (EKZ) és az Alzheimer-kór (AK) különböző stádiumaiban (HOFFMANN–SZATLÓCZKI–KÁLMÁN 2016 nyomán):*  
+ = a sérülés mértéke, - = intakt

A nyelvi jellemzők változása		EKZ	enyhe AK	középsúlyos AK	súlyos AK
szótalálási és szóvisszahívási zavarok		+	+	++	+++
a verbális fluencia korlátozódása	fonemikus	+	+	++	+++
	szemantikus	+	+	++	+++
szemantikus parafázia		-	+	++	+++

A szintaktikai korlátozódás jellemzően a középsúlyos és a súlyos AK-t érinti. Szintaktikai szinten (3. táblázat) a mondatmegértés és a mondatprodukciónál az AK-val élő személyek hibamintázatai részlegesen hasonlítanak az afáziával élő személyek szintaktikai korlátozottságához (ULLMAN 2001; BÁNRÉTI–HOFFMANN–VINCZE 2016). Sérül a mondatmegértés, a rendhagyó tematikus szerepek kiosztásának megértése, a szintaktikailag komplex szerkezetek létrehozása és a narratív beszéd szintaxisa (JUNCOS-RABADÁN et al. 2010; MANOUILIDOU–DE ALMEIDA 2009; MANOUILIDOU et al. 2009). A súlyos szakaszban felerősödik az agrammatizmus (PINANGO 2006; FYNDANIS et al. 2013; BÁNRÉTI–HOFFMANN–VINCZE 2016), kiüresednek a mondatok, gyakori a mutizmus.

3. táblázat

*A nyelvi funkciók változása szintaktikai szinten az enyhe kognitív zavarban (EKZ) és az Alzheimer-kór (AK) különböző stádiumaiban (HOFFMANN–SZATLÓCZKI–KÁLMÁN 2016 alapján):*  
+ = a sérülés mértéke, - = intakt

Nyelvi jellemzők változása	EKZ	enyhe AK	középsúlyos AK	súlyos AK
korlátozott szintaktikai komplexitás	-	-	+	+++
agrammatizmus	-	-	-	+++

A pragmatikai szint vizsgálatára eddig kevés kutatás született (4. táblázat). A társalgási struktúrák középsúlyos állapottól kezdve sérülnek. A kór előrehaladtával egyre rövidebb szövegeket produkálnak a betegek, melyek kevesebb releváns információval, ugyanakkor több hibázási

típussal rendelkeznek, oda nem illő frázisokat és szemantikai parafáziákat tartalmaznak (JUNCOS-RABADÁN et al. 2010; RAPP-WILD 2011; TSANTALI-ECONOMIDIS-TSOLAKI 2013). Problémát okoz a hallott vagy olvasott szöveg tartalmának visszaidézése (TALER-PHILLIPS 2008). A tudatelmélet (*ToM, Theory of Mind*), azaz az a képesség, hogy szándékot, vélekedést, vágyakat tudunk tulajdonítani másoknak, s ezáltal meg tudjuk jósolni, magyarázni és értelmezni viselkedésüket, az AK középsúlyos állapotától kezdve sérül, és ez mind az elsődleges (más emberek mentális állapotának kikövetkeztetése), mind a másodlagos (beágyazott mentális állapottulajdonítás, azaz egy személynek egy másik személlyel kapcsolatos vélekedéseinek kikövetkeztetése, bejósolása) hozzáférést érinti (vö. BÁNRÉTI-HOFFMANN-VINCZE 2016).

#### 4. táblázat

*A nyelvi funkciók változása pragmatikai szinten az enyhe kognitív zavarban (EKZ) és az Alzheimer-kór (AK) különböző stádiumaiban (HOFFMANN-SZATLÓCZKI-KÁLMÁN 2016 alapján):*

+ = a sérülés mértéke, - = intakt

A nyelvi jellemzők változása	EKZ	enyhe AK	középsúlyos AK	súlyos AK
produktív és receptív diskurzusszintű feldolgozás	-	+	++	+++
elsődleges tudatelméleti hozzáférés	-	-	+	+++
másodlagos tudatelméleti hozzáférés	-	-	++	+++

A nyelvi deficit tehát már EKZ-tól kezdve jelen van. A betegséget vizsgálva a kutatási eredmények azt az elképzelést erősítik, hogy a nyelvi és az emlékezeti sérülések szorosan együtt járnak (SZATLÓCZKI et al. 2015).

### 3. Parkinson-kór

A **Parkinson-kór (PK)** egy progresszív lefolyású degeneratív központi idegrendszeri betegség, amely az agy egyes részein kialakuló dopaminhiánnyal, illetve kolinergetúlsúllyal áll összefüggésben (HIDASÍ 2010). A dopamin az agyban termelődő ingerületátvivő vegyület, míg a kolinerger idegsejtek egy másik ingerületátvivő anyagot, acetilkolinot termelnek. A Parkinson-kór neurológiai hátterében az agy bizonyos kéreg alatti területeinek, az ún. bazális ganglionoknak a degenerációja és kóros működése áll (HIDASÍ 2010). A bazális ganglionok vagy törzsdúcok az agy kéreg alatti területein elhelyezkedő szürkeállományi réteg, amely erős összeköttetéssel rendelkezik többek között a kérgi területekkel, szerepe van az akaratlagos

mozgások koordinálásában és a magasabb szintű kognitív folyamatok (pl. emlékezet, tanulás) működésében. A nem akaratlagos mozgásokért felelős (ún. extrapiramidális) mozgatórendszer károsodása miatt a Parkinson-kór fő tünetei a nyugalmi remegés, az izommerevség és a hypo-, illetve bradykinézis (meglassultság) (KEMPLER–LANCKER 2002). Mind a mozgásban, mind a pszichés működésben változást okoz: kezdetben féloldali, később mindkét oldali (de aszimmetrikus) remegés, lelassult mozgás, apró léptű, csoszogó járás, mozgás közbeni „lefagyás”; alvászavar, depresszió, emésztési panaszok jellemezhetik (HIDASÍ 2010).

A Parkinson-kór emellett gyakran a beszédben is elváltozásokat okoz. A kór miatt kialakult beszédzavarokat hypokinetikus dizartriának nevezik (a betegek 90%-ánál jelen van), amely a Parkinson-betegek esetében dadogás- és hadarásszerű megakadásokkal jár(hat) együtt. A **hypokinetikus dizartria** agyi eredetű, lelassult mozgással járó kiejtési zavart jelent, mely érintheti a zöngképzést, az artikulációt és a szupraszegmentumok képzését is. Ehhez társulhatnak a Parkinson-kór súlyosságától függően szótalálási nehézségek, mondatmegértési zavarok, légzési és nyelési nehézségek, lassú válaszadás. A beszédbeli tünetek a következők lehetnek: pontatlan artikuláció, monoton beszéd, szabálytalan zöngképzés, nem változó hangerő, csökkent hangsúlyozás, extrém hosszúságú szünetek, az artikulációs tempó megváltozása (a tempó vagy túl gyorsá vagy túl lassúvá válik), gyakori ismétlések (HO et al. 1998; ACKERMANN–ZIEGLER 1991; KEMPLER–LANCKER 2002; GOBERMAN–BLOMGREN–METZGER 2010).

Az állkapocs, a nyelv és az ajkak meglassult működése pontatlanabb artikulációval jár együtt. Ez a beszédhangok akusztikai sajátosságaiban is felfedezhető. Parkinson-betegek és kontrollszemélyek magánhangzóinak formánsait és a magánhangzóteret vizsgálva azt találták, hogy az úgynevezett magánhangzó-artikulációs index (*Vowel Articulation Index*), amelyet az /a, i, u/ hangok első két formánsai alapján számítanak ki, szignifikáns különbséget mutat a két csoport között: a betegeknél kisebb értékű. Ez azt jelenti, hogy a Parkinson-kóros betegek kisebb térben valósítják meg a magánhangzókat, azaz a magánhangzóik kevésbé különülnek el egymástól, mint az egészséges személyeké. A vizsgálatok szerint a magánhangzó artikulációja nem mutatott összefüggést sem a Parkinson-kór súlyosságát megítélő skálán elért pontszámmal, sem a betegség stádiumával (SKODDA–VISSER–SCHLEGEL 2011). RUSZ és munkatársai (2013) azt találták, hogy a beszéd típus is meghatározó abban, hogy kimutathatók-e, illetve milyen mértékű eltérések mutathatók ki a beszéd akusztikai szerkezetében a kezdeti stádiumban lévő Parkinson-kór esetén a kontrollbeszélőkhöz képest. Eredményeik szerint a kitarított hangokon kevésbé manifesztálódnak a Parkinson-kóros és az egészséges beszéd közötti artikulációs különbségek, míg leginkább a spontán narratívákban jelennek meg a betegségre jellemző eltérések.

A beszédtempót és az artikulációs tempót illetően ellentmondó eredmények olvashatók a szakirodalomban. Egyes kutatók nem találtak szignifikáns különbséget a Parkinson-betegek és a kontrollbeszélők artikulációs tempójában (DE LETTER et al. 2006; SKODDA–SCHLEGEL 2008), míg mások szignifikánsan gyorsabb (ADAMS 1994), megint mások lassabb tempót mértek a betegeknél (MARTÍNEZ-SÁNCHEZ et al. 2014; HSU et al. 2017). A tempógyorsulást elsősorban az angolra igazolták (ADAMS 1994). Ugyanakkor MÁRTÍNEZ-SÁNCHEZ és munkatársai (2016) spanyol anyanyelvű betegek beszédében lassabb tempót mértek a betegeknél a kontrollokhoz képest, és



ez a tempólassulás nem függött össze a gyógyszeres kezeléssel, a beteg életkorával és a betegség stádiumával sem. Hasonlóan lelassult tempót igazoltak Hsu és munkatársai (2017) mandarin anyanyelvű betegek esetében. Magyar anyanyelvű betegeknél diadochokinetikus beszédmozgás-vizsgálat során rögzített hangfelvételeken elemezték az artikulációs tempót (Kis és mtsai 2020). A vizsgálat lényege az, hogy az adatközlőnek egy CV-kapcsolatokból álló hangsort (például *pataka, badaga*) kell folyamatosan ismételnie tíz másodpercig. Az eredmények azt mutatták, hogy a különböző neurológiai kórképeket mutató betegek artikulációs tempója szignifikánsan különbözött egymástól: a Parkinson-betegek tempója volt a leggyorsabb, ezután következett a szklerózis multiplexszel élő betegeké, míg a leglassabban a stroke-on átesett betegek artikuláltak.

A megváltozott beszédtempó dadogásszerű és hadarásszerű tünetekkel is együtt jár (Ho et al. 1998). Ezeket a szakirodalom neurogén megakadásoknak hívja (GOBERMAN–BLOMGREN–METZGER 2010), mivel nem fejlődési eredetűek, hanem szerzett agyi zavar okozza őket. A megakadások jellemzőit a PK-s betegek beszédében több tanulmányban vizsgálták például a gyógyszeres (levodopa-)kezeléssel összefüggésben is. Egyes szerzők a levodopa hatására kevesebb dadogásszerű megakadást találtak (LEDER 1996), mások többet (LOUIS et al. 2001), míg KOLLER (1983) mindkét esetre talált példát. A gyógyszer hatására bekövetkező változások vizsgálatakor úgynevezett *off* és *on* állapotban is szokták a beszédet elemezni, azaz a gyógyszer bevétele előtt (*off*) és után (*on*). Ezek az elemzések csoportszinten nem találtak különbséget a két állapot között a beszéd folyamatosságát tekintve (bár egyéni eltérések lehettek) (BENKE et al. 2000; GOBERMAN–BLOMGREN 2003). GOBERMAN és munkatársai (2010) szerint a betegek szignifikánsan több megakadást produkáltak felolvasáskor, mint a kontrollbeszélők. A megakadások típusait tekintve gyakoribbak voltak a betegeknél a szó belseji (dadogásszerű) megakadások (újraindítás, szünet a szóban, nyújtás) a kontrollokhoz képest, míg a szavak közötti megakadások (egyéb típusú megakadások: kitöltött szünet / hezitálás, ismétlés, töltelékszó) gyakoriságában nem volt különbség a két csoport között.

A Parkinson-kór kezelése főként gyógyszeres úton történik, de súlyos stádiumban lehetőség van mély agyi stimulációra is. Ez tüneti kezelés, a betegség előrehaladását nem befolyásolja, viszont javít a beteg életminőségén. A kutatások szerint egyes betegeknél a beszéd javulását okozza (WALKER et al. 2009), míg mások szerint a beszéd romlását (BURGHAUS et al. 2006); és mindkét esetre talált példát MURDOCH (2010). SNYDER és munkatársai (2018) azt találták, hogy a mély agyi stimuláció még több dadogás- és hadarásszerű megakadást okozott a beszédben.

A dizartria beszédbeli megjelenése mellett a magasabb szintű nyelvi funkciók is károsodást mutathatnak Parkinson-kórban. A betegség együtt járhat demenciával, de a különböző kognitív területek, például végrehajtó funkciók, nyelv, memória és vizuális térbeli képességek károsodását leírták olyan betegeknél is, akiknél nem volt nyilvánvaló demencia. Ezek a károsodások már a betegség korai szakaszában is jelentkezhetnek. Habár nem minden neuropszichológiai kutatás szerint van kimutatható nyelvi károsodás a Parkinson-kórban, a szóelőhívási képességet vizsgáló szemantikusfluencia- és az igefluencia-feladatban a betegek teljesítménye elmarad a kontrollokhoz képest, éppúgy, mint a különböző ismétléses/visszamondásos feladatokban (CABALLOL–MARTÍ–TOLOSA 2007). A betegség előrehaladtával csökken a spontán közlések

szintaktikai komplexitása (ILLES et al. 1988), és sérül a mondatmegértés és a pragmatikai jelenségek megértése is (PELL–MONETTA 2008).

A Parkinson-kóros beszéd érthetőségét nagymértékben befolyásolja a beszéd típusa. KEMPLER és LANCKER (2002) kísérletének tanúsága szerint egy 74 éves Parkinson-beteg férfi beszéde legkevésbé a spontán megnyilatkozás során volt érthető (mindössze 29%-ban), az olvasás során az érthetőség 78%-os volt, míg spontán éneklés során 88%-os.

A Parkinson-kórt általában időskori betegségnek tartják, de előfordulhat fiatal korban is. A 21 és 39 év között kezdődő betegséget fiatalkori Parkinson-kórnak, a 20 éves kor alatti betegséget juvenilis Parkinson-kórnak nevezik (CALNE–KUMAR 2008). Ezek a típusok csak igen ritkán fordulnak elő. A nemzetközi és a hazai szakirodalomban egyaránt alig olvasható információ a fiatalkori betegségről, általában annyi tudható, hogy a betegség lefolyása enyhébb, de hosszabb tartamú, mint idősebb korban (GOLBE 1991; SCHRAG et al. 2003; CALNE–KUMAR 2008). A fiatal Parkinson-betegek esetében általában nem társulnak egyéb betegségek a kórképhez; ugyanakkor ezeknek a betegeknek számos olyan életmódbeli kihívással kell szembenézniük, amelyek pszichésen különös nehézséget jelentenek a számukra. A fiatalkori és az időskori Parkinson-kór beszédbeli tüneteit összevetve megállapították (MESSINIS–ANTONIADIS–BIRIS 2001), hogy a fiatalkori betegség esetén ritkább tünet a levegős zöngképzés, a hipernazalizálás és a beszéd érthetőségének csökkenése, mint időskorban. Ugyanakkor az időskori betegségnél gyakoribb, szembetűnőbb jellegzetességei a fiatalkorú Parkinson-betegek beszédének a rekedtes zöngképzés, az artikulációs zavarok (pontatlan artikuláció), a monotonná váló beszéd, illetve a beszédritmus zavarai (MESSINIS–ANTONIADIS–BIRIS 2001).

## 4. Szklerózis multiplex

A **szklerózis multiplex (SM)** a központi idegrendszer krónikus, fehérállományt érintő megbetegedése, melynek kiváltó oka autoimmun gyulladáshoz vezető reakció. A betegség ún. *shubok*-ban (tünetes állapotok) zajlik, a mielinpusztulás következtében az agyban gyulladáshoz vezető góccok (ún. plakkok) alakulnak ki. A sokéves kórlefolyás során az idegrendszer tartós károsodása alakul ki. A betegség 3:2 arányban érinti a nőket és a férfiakat, jellemzően 20–40 éves kor között kezdődik. Földrajzi elterjedése szempontjából nem egységes a kép, mivel a mérsékelt égövön gyakori, míg a trópusi éghajlaton előfordulása rendkívül ritka (SZIRMAI 2006).

Az SM-nek három fő típusa különböztethető meg. A betegség az esetek nagyjából 80-85%-ában ún. relapszáló-remittáló módon indul, vagyis az előzmények nélkül föllépő *shubok*at, más néven relapszusokat (tünetes állapotok) teljesen vagy nagyrészt tünetmentes, remissziós állapotok követik. Ezek az időszakok hosszabb-rövidebb időközönként váltakoznak. Ezt a típust nevezzük **relapszáló-remittáló SM-nek (RRSM)**. Ez a típus az esetek mintegy 60%-ában egy idő után progrediál, vagyis kialakul egy tartósan romló állapot, amelyben a relapszusokat nem

követik tünetmentes időszakok. Ilyenkor a tünetek állandósulnak, majd elindul a folyamatos romlás. Ez a típus a **másodlagos progresszív SM**. A betegség az érintettek körülbelül 20%-ánál pedig eleve progresszív módon indul, vagyis az első tünetek jelentkezésétől kezdve nincsenek tünetmentes időszakok, hanem folyamatos állapotromlás tapasztalható, ez az ún. **elsődleges progresszív SM**.

Az SM-ben megjelenő plakkok leggyakrabban az agytörzsben, a kisagyban (*cerebellum*), a két agyféltekét összekötő fehérállományi idegrostkötegben, vagyis a kérgestestben (*corpus callosum*), a látóidegben (*nervus opticus*), a gerincvelőben, a bazális ganglionokban, valamint a thalamusban jönnek létre (WYLEZINSKA 2003; HABEK 2013; MINAGAR et al. 2013). Az SM tünetei a betegség változatos idegrendszeri háttere miatt mind jellegükben, mind súlyosságukban meglehetősen sokfélék. A tünetek lehetnek 1. kognitív (pl. krónikus fáradtság, depresszió, demencia, figyelemzavar); 2. szenzoros (pl. fokozott fájdalom és hidegérzékelés); 3. motoros (pl. végtaggyengeség, zsibbadás, izommerevség); 4. agytörzsi (pl. szemmozgás-zavar, kettős látás, nyelési és rágási nehézség, szédülés); 5. kisagyi tünetek (pl. egyensúly- és járászavar, dizartria, diszprozódias beszéd); valamint 6. gerincvelői eredetű zavarok (pl. vizeletürítési és emésztési problémák) (SZIRMAI 2006). A betegség leggyakrabban látóideg-gyulladással, látásromlással kezdődik.

A nyelvi és beszédfolyamatok megfelelő működése számos kognitív, szenzoros és motoros működés dinamikus összehangolását igényli. SM-ben az információfeldolgozási sebesség a gyulladással kapcsolatos plakkokkal kapcsolatban álló agyterületeken szükségszerűen lelassul, mivel a neuronok axonjait körülvevő mielinhüvelyek károsodása vagy pusztulása éppen az információfeldolgozás sebességét és hatékonyságát csökkenti. A csökkent információáramlási képesség számos ponton lehet hatással a nyelvi és beszédfolyamatokra. Átmeneti vagy állandó nyelvi és beszédzavarok az SM-mel élő betegek mintegy 60 százalékánál fordulnak elő, ezek súlyosságukat és jellegüket tekintve az imént ismertetett okok következtében igen különbözőek lehetnek. Megjelenhetnek a nyelvi és beszédtevékenység bármely folyamatában (percepció, megértés, produkció) és szintjén (az artikulációtól a pragmatikáig) (lásd RENAULD et al. 2016 részletes áttekintő tanulmányát). Az SM-ben előforduló leggyakoribb nyelvi és beszédzavarok a következők: 1. a beszéd- és hangképzés zavarai: a) dizartria, b) diszfónia, c) az artikulációs és beszédtempó lassulása; 2. nyelvi zavarok mind a produkció, mind a percepció oldalán: d) szókeresési és szóelőhívási nehézségek, e) megnevezési zavar, f) a verbális fluencia csökkenése, g) szemantikai parafáziák, h) megértési zavarok. Ritkán afáziás tünetek is előfordulhatnak, súlyos (de ritka) esetben akár globális afázia is.

Az SM-betegekkel végzett nyelvi és beszédvizsgálatokban általában kétféle módszertan alapján csoportosítják a résztvevőket a kutatások. Így a vizsgált páciensek megkülönböztetését vagy a betegség típusa szerint (relapszáló-remittáló vs. progresszív), vagy pedig a kimutatható kognitív hanyatlás mértéke szerint (nincs kognitív hanyatlás – enyhe/kezdődő hanyatlás – közép-súlyos vagy súlyos hanyatlás) teszik meg. Noha a progresszív típusokkal ellentétben a relapszáló-remittáló típusban nem gyakori a kognitív hanyatlás, az RRSMBől progresszívvá váló állapotromlás általában egy hosszabb (akár években mérhető) folyamat eredménye, ezért nem mindig egyértelmű az SM aktuális típusának meghatározása. Így pedig többnyire a kognitív romlás

megindulásának sincsenek feltűnő, hirtelen felbukkanó jelei, egy ideig célzott vizsgálatok nélkül még nem kimutatható. Ezért szükséges, hogy egy-egy nyelvi vizsgálat kiegészüljön a beteg kognitív állapotának felmérésével, mivel ez teszi lehetővé az eredmények megfelelő értékelését. Az elmúlt két évtized kutatásai azt mutatják, hogy már a kezdődő kognitív hanyatlás is kimutatható a nyelvi és beszéd folyamatokban megfelelően célzott vizsgálatokkal (KUJALA et al. 1996).

A beszéd- és a hangminőség változásairól a betegek mintegy 50-60%-a számol be (BAUER et al. 2013; HARTELIUS et al. 2000). Az SM-ben megjelenő leggyakoribb beszédzavar tehát a dizartria, valamint az ezzel gyakran együtt járó hangképzési zavar, a diszfónia. A dizartria a motoros, vagyis a mozgások kivitelezésével összefüggő beszédzavarok közé tartozik. Hátterében a központi és perifériás idegrendszeri képletek gyulladása áll, amelyek izombénulást vagy izomgyengeséget okoznak, emellett gyakran spasztikuság (görcsösség), izomremegés (*tremor*) és az izommozgások összerendezésének képtelensége (*ataxia*) is kíséri. A diszfónia általában a dizartria egy összetevőjeként van számontartva, a hangszalagok működési nehézsége következtében alakul ki. E két zavar, a diszfónia és a dizartria számos tünetet produkál SM-ben, úgymint a hangerőszabályozás zavara, rekedtes zöngemínőség, pontatlan artikuláció, monoton hangsúlyozás, monotonabb beszéddallam, csökkent vitálkapacitás és légzési zavarok, hipernazalitás (MILLER 2011; BAUER et al. 2013). A dizartriás tünetek súlyosságát befolyásolhatja a fáradtság aktuális mértéke. Fontos megjegyezni, hogy az SM-ben megjelenő dizartriára gyakran másféle, irreguláris artikulációs mintázatok jellemzőek, ezért az általánosan használt artikulációs tesztek helyett célszerűbb spontánbeszéd-felvételeken keresztül vizsgálni (MILLER 2011). Az ataxiás dizartriára jellemző, hogy a beszéd egyes hangzói lassan ejtettek, és a hangzók között szünet hallható.

A beszédparaméterekre nem csupán a motoros nehézségek lehetnek hatással SM-ben, hanem a kognitív zavarok is okozhatnak változást a beszédprodukciónban (FEENAUGHTY et al. 2013, 2018; DE LOOZE et al. 2019; RODGERS et al. 2013). Bizonyos kognitív képességek, ezen belül is elsősorban az információfeldolgozás működésében történt változások hatással lehetnek a beszéd bizonyos paramétereire. DE LOOZE és munkatársai (2019) azt találták, hogy azoknak az SM-betegeknek a beszédét, akiknél kognitív zavar is kimutatható volt, lassabb beszédtempó, több megakadásjelenség és több néma szünet jellemezte, mint azokét az SM-betegeket, akiknél nem volt kimutatható kognitív zavar. FEENAUGHTY és munkatársainak (2013) kutatása ezzel szemben mind a kognitív szempontból ép, mind a sérült SM-es csoportban különbséget talált bizonyos beszédparaméterekben az egészséges kontrollszemélyekhez képest. Kutatásukban lassabb beszédtempót és gyorsabb artikulációs tempót mértek, hosszabb szünetekkel és megnövekedett szünetarányal az SM-es csoportban. Az SM-mel élők beszédprodukcójára gyakran nagyobb hatással van az adott nyelvi tevékenység létrehozásához szükséges kognitív erőfeszítés mértéke, mint amekkora változást a kognitív terhelés növelésével az egészséges személyek beszédprodukcójában láthatunk. Minél nagyobb ugyanis a végrehajtáshoz szükséges erőfeszítés, annál lassabb beszédtempóval, valamint több és hosszabb szünettel valósul meg a személy beszéde (ARNETT et al. 2008; DE LOOZE et al. 2019). TJADEN és WATLING (2011) megfigyelései szerint az SM-mel élők az egészséges kontrollszemélyekhez képest hajlamosabbak voltak szintaktikailag nem megfelelő helyen (vagyis nem valamely szintaktikai vagy szemantikai egység határán) tartani szünetet.

A lexikai-szemantikai előhívás nehézségeire utal a megnevezési tesztekben megjelenő hibák száma és típusa. BEATTY és MONSON (1990) kutatása szerint az RRSM-betegek 20%-ánál, a progresszív betegeknek pedig a 40%-ánál mutatható ki számottevő megnevezési zavar. Kutatásukban a megnevezési zavart együtt vizsgálták a beteg kognitív állapotával, és azt találták, hogy a kognitív zavart mutató betegek a megnevezési teszten is rosszabb eredményt értek el, mint azok, akik nem mutattak kognitív zavart. FRIEND és munkatársai (1999) azt találták, hogy mind RRSM-ben, mind a progresszív SM-típusokban szignifikánsan rosszabb eredményt értek el a résztvevők a vizuális megnevezési feladatokban a kontrollszemélyeknél. Hasonló eredményre jutottak a beszédmegértési teszt eredményeiben is. LETHLEAN és MURDOCH (1993) megállapították azonban azt is, hogy míg a megnevezési hibák számában a vizsgált SM-betegek rosszabb teljesítmény nyújtottak, mint az egészséges beszélők, addig a hibák jellege és típusai mindkét vizsgálati csoportban azonosak voltak. KUJALA és munkatársai (1996) ettől részben eltérően azt találták, hogy a kognitív zavart nem mutató SM-csoport az egészséges résztvevőkhöz hasonló eredményt mutatott, míg a kognitív képességeiben érintett csoport jellegzetes szemantikai hibákat produkált. Ezek a hibák abból származtak, hogy az érintett betegek egy-egy kép megnevezése során a megfelelő szemantikai kategóriát megtalálták ugyan, de a képen láthatónál egy általánosabb kategóriát neveztek meg (pl. a képen egy kutya látható, a páciens pedig azt mondja rá, hogy *háziállat*). Ez az eredmény arra utal, hogy bár a tágabb szemantikai kategóriák elérhetőek, de a konkrét egyedekhez vezető feldolgozási út hozzáférése nem működik.

A szókeresési nehézség kimutatására neurodegeneratív betegségekből különösen alkalmas módszer az ún. **verbális fluencia** mérése (a feladat ismertetését lásd az 5. fejezetben). HENRY és BEATTY (2006) metaanalízisükben megállapították, hogy a verbálisfluencia-feladatok a legérzékenyebb mérőeszközei lehetnek az SM-ben megjelenő kognitív hanyatlásnak. Eredményeik azt mutatták, hogy mind a relapszáló-remittáló, mind a progresszív csoport szignifikánsan rosszabb eredményt ért el az egészséges résztvevőknél, valamint szignifikáns különbséget találtak a relapszáló-remittáló és a progresszív csoport között is (FRIEND et al. 1999). A verbálisfluencia-feladatok előnye, hogy nem pusztán a szókeresési nehézségekről ad képet, hanem az SM neurológiai háttéréből fakadó lassabb információfeldolgozási mechanizmus feltárására is alkalmas lehet.

Az SM-ben megjelenő kommunikációs nehézségek számbavételekor a magasabb szintű nyelvi folyamatokban bekövetkező zavarokat is vizsgálni kell. Ezek a folyamatok azért mutathatnak korlátozódást SM-ben annak ellenére, hogy a komplex nyelvi feldolgozásért és produkcióért felelős agyterületeket az SM többnyire nem érinti, mert e folyamatok működéséhez szükség van az emlékezeti és figyelmi folyamatokra, hatékony információfeldolgozásra, téri-vizuális feldolgozásra, valamint a következtetési folyamatok működésére (ARNOTT et al. 1997). Az SM-ben megjelenő szemantikai és pragmatikai nehézségekről meglehetősen kevés kutatás készült (LAAKSO et al. 2000; ARRONDO et al. 2010; CAROTENUTO et al. 2018a, b; LETHLEAN–MURDOCH 1997). A pragmatikai képességekben a vizsgált betegek 55%-ánál találtak nehézségeket, mégpedig a beteg állapotától és a betegség típusától, valamint a kognitív állapotuktól függetlenül (CAROTENUTO 2018a, b). Eszerint a pragmatikai képességekben érintett betegek az egészséges

beszélőkhöz hasonlóan voltak képesek részt venni az általános hétköznapi kommunikációs helyzetekben, de kevesebb információt vagy a diskurzus szempontjából kevesebb releváns információt osztottak meg, mint a kontrollszemélyek, kevésbé vagy félreértették a nem szó szerinti jelentéstartalmakat (mint a humor, metafora vagy irónia).

A narratív képességekben potenciálisan fennálló zavarokat számos módszerrel vizsgálhatjuk: képleírási tesztekkel, megnevezési feladatokkal, irányított spontán beszédet előhívó feladatokkal, szövegfelidézési gyakorlatokkal. A narratív képességekben eltérés mutatható ki az SM-betegek egy részénél az egészségesekhez képest: a produkált szavak száma alacsonyabb, a frázisok rövidebbek, kevesebb a mellékmondat, a grammatikai szerkezetek egyszerűbbek, és esetleg agrammatikus fordulatok is megjelennek, az információtartalom és a szókincs kisebb, a szemantikai tartalom kevésbé specifikus, gyengébb a szófluencia, nagyobb a mondatban funkció nélkül megjelenő szavak és a pragmatikai jelölők aránya, valamint a szóelőhívási folyamatok neheztettebbek (WALLACE et al. 1993; CUETOS et al. 2007; ARRONDO et al. 2010). ARNOTT és munkatársai (1997) azt találták, hogy bár a megnyilatkozások hosszában nem mérhető különbség az SM- és a kontrollcsoport között, a tartalomban igen: kevesebb konkrét információt és több kétértelmű referenciális elemet tartalmazott az SM-betegek beszéde, mint a kontrolloké.

## 5. A nyelvészeti vizsgálatok szerepe a klinikai kutatásokban és a diagnosztikában

Az ép nyelvi tevékenység mögött ép kognitív folyamatokat feltételezünk. Az agy hálózatos felépítése miatt bármely kognitív folyamat korlátozódása vagy sérülése hatással lehet a nyelvi folyamatok működésére is. Az általános kognitív állapotot felmérő neuropszichológiai tesztek mindegyike éppen ezért nyelvi funkciókat mérő feladatokat is tartalmaz (pl. tárgymegevezés, mondatisméltés, verbális fluencia). A hazai klinikumban használt tesztek között találjuk többek között a Montreal Kognitív Felmérést (MoCA, NASREDDINE et al. 2005), az Addenbrooke Kognitív Vizsgálatot (AKV, DUDAS et al. 2002), melynek rövidített változata az ún. Mini Mental Teszt (MMT, FOLSTEIN–FOLSTEIN–MCHUGH 1975), a Korai Mental Tesztet (KMT, KÁLMÁN et al. 2013) és az Alzheimer's Disease Assessment Scale – Cognitive Subscale-t (ADAS-Cog, ROSEN–MOHS–DAVIS 1984; PÁKÁSKI et al. 2012).

Az átfogó kognitív állapotot felmérő tesztek és mérőeszközök mellett hasznos információval szolgálhatnak a beteg egy-egy rész-képességét alaposabban felmérő eszközök. Különösen hasznosnak bizonyulnak azok a tesztek, amelyek egyszerűen, rövid idő alatt felvehetőek, mégis komplexek, vagyis egyszerre több rész-képesség működésére következtethetünk általuk. Neurodegeneratív betegségekben az egyik ilyen, leggyakrabban alkalmazott nyelvi, emlékezeti és végrehajtó funkciókat is mérő teszt az ún. **verbálisfluencia-teszt**. Segítségével képet kaphatunk a beteg információfeldolgozási sebességéről (illetve annak lassulásáról), munkamemóriájáról, szemantikus emlékezetének

működéséről, szókeresési és szótalálási stratégiáiról, kognitív flexibilitásáról. A fluenciateszteknek három típusát szokás megkülönböztetni: a fonemikus- vagy betűfluencia-feladatban a betegnek megadott időn keresztül (1 perc) megadott hanggal kezdődő szavakat kell sorolni; a szemantikus-fluencia-feladat bizonyos kategóriába tartozó szavak felsorolásából áll (pl. állatok, testrészek, színek, élelmiszerek); az akció- vagy igefluencia-feladatban a betegnek cselekvéseket/igéket kell sorolniuk. A fluenciatesztek előnye, hogy gyorsan felvehetőek, a beteg számára minimális stresszel járnak, valamint végrehajtásuk során nem érintettek a neurodegeneratív zavarokban, így például az SM-ben gyakran általánosan sérült motoros és vizuális folyamatok. Számos kutatás igazolja ezeknek az egyszerű teszteknek az állapotmegismerésben való hatékonyságát.

Bizonyos esetekben a klinikai gyakorlatban hazánkban is alkalmazott, a szerzett agysérülések (pl. stroke vagy traumatikus agysérülés) következményeit felmérő afázia-tesztek (pl. Western Aphasia Battery, WAB, KERTESZ 1982) használata is hasznos és indokolt lehet. Ezek a tesztek komplex módon mérik fel az egyén beszédprodukciós, beszédmegértési, megnevezési, utánmondási, olvasási és írási képességeit.

Mivel a fent ismertetett neurodegeneratív betegségekben gyakran megjelenő tünetként kell számontartanunk a megnevezési zavarokat, ezért ennek mérésére megfelelő eszköz lehet például a magyar nyelvre is validált Boston Megnevezési Teszt (KAPLAN–GOODGLASS–WEINTRAUB 1983). Ugyanígy alkalmazhatók ennek más résztesztjei, a képleíráshoz használt tesztek is. Azokban az esetekben, ahol feltételezhető, hogy a megértési folyamatok is sérültek, alkalmazható a – szintén az afázia diagnosztikus gyakorlatában alkalmazott – DE RENZI-féle Token beszédmegértési teszt is (DE RENZI–VIGNOLO 1962).

A mesterséges intelligencia és a számítógépes nyelvészet segítségével olyan automatikus beszédfelismerő rendszerek készültek, melyek eredményesen használhatóak a klinikai differenciáldiagnosztikában. Az elmúlt néhány évben több kutatás is született a spontán beszéd temporális paramétereinek mérésére. A szakirodalom nem egységes abban, hogy mely paraméterek tudják elválasztani az enyhe AK-val élők beszédét az egészséges kontrollok beszédétől, megegyezik ugyanakkor abban, hogy a spontán beszéd temporális analízise hatásos módszernek bizonyul az enyhe AK felismerésére (HOFFMANN–SZATLÓCZKI–KÁLMÁN 2016). A spontán beszéd temporális paramétereinek automatikus elemzésére rendelkezésre áll egy hazai fejlesztésű teszt, a Speech-Gap teszt (S-GAP), mely egy spontánbeszéd-feladat során 86%-os pontossággal tudja megkülönböztetni az egészséges, az enyhe kognitív zavarral élők és az enyhe AK-val élők beszédét (GOSZTOLYA et al. 2019).

A nemzetközi gyakorlatban a beszéd- és hangképzési nehézségeknek a hétköznapi életminőségre gyakorolt hatása sztenderd módszerrel mérhető. A Voice Handicap Index (JACOBSON et al. 1997) a betegek hangképzési zavarainak a pszichoszociális életükre gyakorolt hatását mérő eszköz. Emellett több olyan automatikus mérési módszer is létezik, amely a beszéd alapján igen nagy valószínűséggel tudja diagnosztizálni a Parkinson-kórt. Egyes eljárások a zöngé alaphangjának tulajdonságai és a magánhangzók akusztikai szerkezetének megváltozása alapján akár 96,29%-os pontossággal különítik el a betegeket az egészséges csoporttól (KARAN–MAHTO–SAHU 2018), más eljárásokban több ezer (előre nem meghatározott)

beszédparamétert vizsgálva, gépi tanulással sikerült jól elkülöníteni a két csoportot egymástól (GRÓSZ et al. 2015). A nyelvészeti és beszédkutatók eredményei tehát jól hasznosíthatók lehetnek a klinikai kutatásokban és a diagnosztika folyamatában is.

## FOGALMAK

*neurodegeneratív betegségek; végrehajtó funkciók; afázia; diszfluencia; fonemikus parafázia; szemantikus parafázia; agrammatizmus; alexia; agráfia; procedurális memória; deklaratív memória; demencia; primer progresszív afázia; dizartria; frontotemporális demencia; perszeveráció; Huntington-kór; Alzheimer-kór; enyhe kognitív zavar; Parkinson-kór; szklerózis multiplex; relapszáló-remittáló SM; másodlagos progresszív SM; elsődleges progresszív SM; verbális fluencia, verbálisfluencia-teszt*

## IRODALOM

- ACKERMANN, Hermann – ZIEGLER, Wolfram 1991. Articulatory deficits in parkinsonian dysarthria: An acoustic analysis. *Journal of Neurology, Neurosurgery & Psychiatry* 54/12. 1093–1098.
- ADAMS, Scott G. 1994. Accelerating speech in a case of hypokinetic dysarthria: Descriptions and treatment. In TILL, James A. – YORKSTON, Kathryn M. – BEUKELMAN, David R. (eds): *Motor Speech Disorders: Advances in assessment and treatment*. Brookes, Baltimore. 213–228.
- AMERICAN PSYCHIATRIC ASSOCIATION 2013. *Diagnostic and Statistical Manual of Mental Disorders (5th edn, text revision) (DSM–5)*. American Psychiatric Publishing, Washington.
- APPELL, Julian – KERTESZ, Andrew – FISHMAN, Michael 1982. A study of language functioning in Alzheimer patients. *Brain and Language* 17/1. 73–91.
- ARNETT, Peter A. – SMITH, Megan M. – BARWICK, Fiona – BENEDICT, Ralph H. B. – AHLSTROM, Brian P. 2008. Oralmotor slowing in multiple sclerosis: Relationship to neuropsychological tasks requiring an oral response. *Journal of the International Neuropsychological Society* 14/3. 454–462.
- ARNOTT, Wendy L. – JORDAN, Faye M. – MURDOCH, Bruce E. – LETHLEAN, Jennifer B. 1997. Narrative discourse in multiple sclerosis: An investigation of conceptual structure. *Aphasiology* 11/10. 969–991.
- ARRONDO, Gonzalo – SEPULCRE, Jorge – DUQUE, Beatriz – TOLEDO, Jon – VILLOSLADA, Pablo 2010. Narrative speech is impaired in multiple sclerosis. *European Neurological Journal* 2/1. 1–8.
- BALTHAZAR, Marcio L. F. – MARTINELLI, José E. – CENDES, Fernando – DAMASCENO, Benito P. 2007. Lexical semantic memory in amnesic mild cognitive impairment and mild Alzheimer's disease. *Arquivos de Neuropsiquiatria* 65/3A. 619–622.
- BÁNRETI Zoltán 2014. *Az afázia*. In PLÉH Csaba – LUKÁCS Ágnes (szerk.): *Pszicholingvisztika 1–2*. Akadémiai Kiadó, Budapest. 1187–1261.
- BÁNRETI, Zoltán – HOFFMANN, Ildikó – VINCZE, Veronika 2016. Recursive subsystems in Aphasia and Alzheimer's Disease: Case studies in Syntax and Theory of Mind. *Frontiers in Psychology* 31. March. <http://dx.doi.org/10.3389/fpsyg.2016.00405>.



- BARTH, Sonja – SCHÖNKNECHT, Peter – PANTEL, Johannes – SCHRÖDER, Johannes 2005. Mild cognitive impairment and Alzheimer's disease: An investigation of the CERAD-NP test battery. *Fortschritte der Neurologie-Psychiatrie* 73/10. 568–576.
- BAUER, Vladimir – ALERIC, Zorica – JANCIC, Ervin – KNEZEVIC, Bojana – PRPIC, Dubravka – KACAVENTA, Anita 2013. Subjective and perceptual analysis of voice quality and relationship with neurological dysfunction in multiple sclerosis patients. *Clinical Neurology and Neurosurgery* 115/Suppl.1. S17–20.
- BAYLES, Kathryn A. – BOONE, Daniel R. 1982. The potential of language tasks for identifying senile dementia. *Journal of Speech and Hearing Disorders* 47/2. 210–217.
- BAYLES, Kathryn A. – KASZNAK, Alfred W. – TOMODEA, Cheryl K. 1987. *Communication and Cognition in Normal Aging and Dementia*. College-Hill Press, Boston, MA.
- BAYLES, Kathryn A. – TOMODEA, Cheryl K. – TROSSET, Michael W. 1992. Relation of linguistic abilities of Alzheimer's patients to stage of disease. *Brain and Language* 42/4. 454–472.
- BEATTY, William W. – MONSON, Nancy 1990. Semantic priming in multiple sclerosis. *Bulletin of the Psychonomic Society* 28. 397–400.
- BECKER, Robert – GIACOBINI, Ezio 1996. *Alzheimer Disease: From molecular biology to therapy*. Birkhäuser, Boston.
- BENKE, Thomas H. – HOHENSTEIN, Christian – POEWE, Werner – BUTTERWORTH, Brian 2000. Repetitive speech phenomena in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry* 69/3. 319–324.
- BURGHHAUS, Lothar – HILKER, Ruediger – THIEL, Alina – GALLDIKS, Norbert – LEHNHARDT, Fritz-Georg – ZARO-WEBER, Olivier – STURM, Volker – HEISS, Wolf-Dieter 2006. Deep brain stimulation of the subthalamic nucleus reversibly deteriorates stuttering in advanced Parkinson's disease. *Journal of Neural Transmission* 113/5. 625–631.
- CABALLOL, Nuria – MARTÍ, Maria J. – TOLOSA, Eduardo 2007. Cognitive dysfunction and dementia in Parkinson disease. *Movement Disorders: Official Journal of the Movement Disorder Society* 22/S17. S358–366.
- CALNE, Susan M. – KUMAR, Ajit 2008. Young onset Parkinson's disease. Practical management of medical issues. *Parkinsonism & Related Disorders* 14/2. 133–142.
- CAROTENUTO, Antonio – ARCARA, Giorgio – OREFICE, Giuseppe – CERILLO, Ilaria – GIANNINO, Valentina – RASULO, Mario – IODICE, Rosa – BAMBINI, Valentina 2018a. Communication in multiple sclerosis: Pragmatic deficit and its relation with cognition and social cognition. *Archives of Clinical Neuropsychology* 33/2. <https://doi.org/10.1093/arclin/acx061>.
- CAROTENUTO, Antonio – COCOZZA, Sirio – QUARANTELLI, Mario – ARCARA, Giorgio – LANZILLO, Roberta – MORRA, Vincenzo B. – CERILLO, Ilaria – TEDESCHI, Enrico – OREFICE, Giuseppe – BAMBINI, Valentina – BRUNETTI, Arturo – IODICE, Rosa 2018b. Pragmatic abilities in multiple sclerosis: The contribution of the temporo-parietal junction. *Brain and Language* 185. 47–53.
- CLARK, David G. – CHARUVASTRA, Anthony – MILLER, Bruce L. – SHAPIRA, Jill S. – MENDEZ, Mario F. 2005. Fluent versus nonfluent primary progressive aphasia: A comparison of clinical and functional neuroimaging features. *Brain and Language* 94/1. 54–60.

- CROOT, Karen – HODGES, John R. – XUERE, John – PATTERSON, Karalyn 2000. Phonological and articulatory impairment in Alzheimer's disease: A case series. *Brain and Language* 75/2. 277–309.
- CUETOS, Fernando – ARANGO-LASPRILLA, Juan C. – URIBE, Claramónika – VALENCIA, Claudia – LOPERA, Fransisco 2007. Linguistic changes in verbal expression: A preclinical marker of Alzheimer's disease. *Journal of International Neuropsychological Society* 13/3. 433–439.
- CUMMINGS, Jeffrey – MEGA, Michael S. 2003. *Neuropsychiatry and Behavioral Neuroscience*. Oxford University Press, Oxford.
- DE LOOZE, Céline – MOREAU, Noémie – RENIÉ, Laurent – KELLY, Finnian – GHIO, Alain – RICO, Audrey – AUDOIN, Bertrand – VIALLET, François – PELLETIER, Jean – PETRONE, Caterina 2019. Effects of cognitive impairment on prosodic parameters of speech production planning in multiple sclerosis. *Journal of Neuropsychology* 13/1. 1–24.
- DE LETTER, Miet – SANTENS, Patrick – DE BODT, Marc – BOON, Paul – VAN BORSEL, John 2006. Levodopa-induced alterations in speech rate in advanced Parkinson's disease. *Acta Neurologica Belgica* 106/1. 19–22.
- DE RENZI, Ennio – VIGNOLO, Luigi A. 1962. The Token Test: A sensitive test to detect receptive disturbances in Aphasics. *Brain* 85/4. 665–678.
- DOS SANTOS, Vasco – THOMANN, Philipp A. – WÜSTENBERG, Torsten – SEIDL, Ulrich – ESSIG, Marco – SCHRÖDER, Johannes 2011. Morphological cerebral correlates of CERAD test performance in mild cognitive impairment and Alzheimer's disease. *Journal of Alzheimer's Disease* 23/3. 411–420.
- DUDÁS, Róbert – STACHÓ, László – KOTHENCZ, Gabriella – IVÁDY, Rozália – JANKA, Zoltán 2002. The Hungarian version of ACE 6/11/2002. Addenbrooke's Kognitív Vizsgálat. Hungarian translation based on: MATHURANATH, Pavagada S. – NESTOR, Peter J. – BERRIOS, German E. – RAKOWITZ, Wojtek – HODGES, John R. 2000. A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. *Neurology* 55/11. 1613–1620.
- DUONG, Anh – WHITEHEAD, Victor – HANRATTY, Kate – CHERTKOW, Howard 2006. The nature of lexico-semantic processing deficits in mild cognitive impairment. *Neuropsychologia* 44/10. 1928–1935.
- EMERY, Olga V. 2000. Language impairment in dementia of the Alzheimer type: A hierarchical decline? *International Journal of Psychiatric Medicine* 30/2. 145–164.
- FEENAUGHTY, Linda – TJADEN, Kris – BENEDICT, Ralph H. B. – WEINSTOCK-GUTTMAN, Bianca 2013. Speech and pause characteristics in multiple sclerosis: A preliminary study of speakers with high and low neuropsychological test performance. *Clinical Linguistics and Phonetics* 27/2. 134–151.
- FEENAUGHTY, Linda – TJADEN, Kris – WEINSTOCK-GUTTMAN, Bianca – BENEDICT, Ralph H. B. 2018. Separate and combined influence of cognitive impairment and dysarthria on functional communication in multiple sclerosis. *American Journal of Speech-Language Pathology* 27/3. 1051–1065.

- FOLSTEIN, Marshal F. – FOLSTEIN, Susan E. – MCHUGH, Paul R. 1975. Mini mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 12/3. 189–198.
- FORBES, Katarina E. – VENNERI, Annalena – SHANKS, Michael F. 2002. Distinct patterns of spontaneous speech deterioration: A mild predictor of Alzheimer’s disease. *Brain and Cognition* 48/2–3. 356–361.
- FRASER, Kathleen C. – MELTZER, Jed A. – GRAHAM, Naida L. – LEONARD, Carol – HIRST, Graeme – BLACK, Sandra E. – ROCHON, Elizabeth 2014. Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex* 55. 43–60.
- FRIEND, Karen B. – RABIN, Bernard M. – GRONINGER, Lowell – DELUTY, Robert H. – BEVER, Christopher – GRATAN, Lynn 1999. Language functions in patients with multiple sclerosis. *Clinical Neuropsychology* 13/1. 78–94.
- FYNDANIS, Valantis – MANOUILIDOU, Christina – KOUFOU, Eugenia – KARAMPEKIOS, Spyros – TSAPAKIS, Eva-Maria 2013. Agrammatic patterns in Alzheimer’s disease: Evidence from tense, agreement, and aspect. *Aphasiology* 27/2. 178–200.
- GARRARD, Peter – MALONEY, Lisa M. – HODGES, John R. – PATTERSON, Karalyn 2005. The effects of very early Alzheimer’s disease on the characteristics of writing by a renowned author. *Brain* 128/2. 250–260.
- GARRARD, Peter – RENTOUMI, Vassiliki – GESIERICH, Benno – MILLER, Bruce – GORNO-TEMPINI, Maria L. 2014. Machine learning approaches to diagnosis and laterality effects in semantic dementia discourse. *Cortex* 55. 122–129.
- GIFFARD, Bénédicte – DESGRANGES, Béatrice – NORE-MARY, Florence – LALEVÉE, Catherine – DE LA SAYETTE, Vincent – PASQUIER, Florence – EUSTACHE, Francis 2001. The nature of semantic memory deficits in Alzheimer’s disease: New insights from hyperpriming effects. *Brain* 124/8. 1522–1532.
- GIFFARD, Bénédicte – DESGRANGES, Béatrice – NORE-MARY, Florence – LALEVÉE, Catherine – DE LA SAYETTE, Vincent – PASQUIER, Florence – EUSTACHE, Francis 2002. The dynamic time course of semantic memory impairment in Alzheimer’s disease: Clues from hyperpriming and hypoprimeing effects. *Brain* 125/9. 2044–2057.
- GOBERMAN, Alexander M. – BLOMGREN, Michael 2003. Parkinsonian speech disfluencies: Effects of L-dopa-related fluctuations. *Journal of Fluency Disorders* 28/1. 55–70.
- GOBERMAN, Alexander M. – BLOMGREN, Michael – METZGER, Erika 2010. Characteristics of speech disfluency in Parkinson disease. *Journal of Neurolinguistics* 23/5. 470–478.
- GOLBE, Lawrence I. 1991. Young-onset Parkinson’s disease. A clinical review. *Neurology* 41/2. (Part 1). 168–168.
- GORNO-TEMPINI, Maria L. – MILLER, Bruce L. 2013. Primary Progressive Aphasia as a model to study the neurobiology of language. *Brain and Language* 127/2. 105.
- GOSZTOLYA, Gábor – VINCZE, Veronika – TÓTH, László – PÁKÁSKI, Magdolna – KÁLMÁN, János – HOFFMANN, Ildikó 2019. Identifying Mild Cognitive Impairment and mild Alzheimer’ disease based on spontaneous speech using ASR and linguistic features. *Computer Speech and Language* 53. 181–197.

- GRÓSZ, Tamás – BUSA-FEKETE, Róbert – GOSZTOLYA, Gábor – TÓTH, László 2015. Assessing the degree of nativeness and Parkinson's condition using Gaussian processes and deep rectifier neural networks. In MÖLLER, Sebastian – NEY, Hermann – MÖBIUS, Bernd – NÖTH, Elmar – STEIDL, Stefan (eds): *Speech Beyond Speech Towards a Better Understanding of the Most Important Biosignal: Proceedings of the Annual Conference of the International Speech Communication Association (INTERSPEECH 2015)*. Vols. 1–4. International Speech Communication Association (ISCA), Dresden. 919–923.
- HABEK, Mario 2013. Evaluation of brain stem involvement in multiple sclerosis. *Expert Review of Neurotherapeutics* 13/3. 299–311.
- HARTELIUS, Lena – NORD, Lennart 2000. Temporal speech characteristics of individuals with multiple sclerosis and ataxic dysarthria: 'Scanning speech' revisited. *Folia Phoniatrica and Logopedica* 52/2. 228–238.
- HENRY, Julie D. – BEATTY, William W. 2006. Verbal fluency deficits in multiple sclerosis. *Neuropsychologia* 44/7. 1166–1174.
- HIDASÍ Eszter 2010. A Parkinson-kór modern szemlélete. *Lege Artis Medicinae* 20/12. 825–829.
- HO, Aileen K. – IANSEK, Robert – MARIGLIANI, Caterina – BRADSHAW, John L. – GATES, Sandra 1998. Speech impairment in a large sample of patients with Parkinson's disease. *Behavioural Neurology* 11/3. 131–137.
- HOFFMANN Ildikó – SZATLÓCZKI Gréta – KÁLMÁN János 2016. Mit jelez a nyelvi képességek változása Alzheimer-kórban? In KAS Bence (szerk.): „Szavad ne feledd!” *Tanulmányok Bánréti Zoltán tiszteletére*. MTA Nyelvtudományi Intézet, Budapest. 165–178.
- HOFFMANN, Ildikó – NÉMETH, Dezső – DYE, Cristina D. – PÁKÁSKI, Magdolna – IRINYI, Tamás – KÁLMÁN, János 2010. Temporal parameters of spontaneous speech in Alzheimer's disease. *International Journal of Speech-Language Pathology* 12/1. 29–34.
- HODGES, John R. – MILLER, Bruce 2001. The neuropsychology of frontal variant frontotemporal dementia and semantic dementia. Introduction to the special topic papers: Part II. *Neurocase* 7/2. 113–121.
- HSU, Sih C. – JIAO, Yishan – MCAULIFFE, Megan J. – BERISHA, Visar – WU, Ruey M. – LEVY, Erika S. 2017. Acoustic and perceptual speech characteristics of native Mandarin speakers with Parkinson's disease. *The Journal of the Acoustical Society of America* 141/3. EL293–299.
- ILLES, Judy – METTER, E. Jeffrey – HASON, Wayne R. – IRITANI, S. 1988. Language production in Parkinson's disease: Acoustic and linguistic considerations. *Brain and Language* 33/1. 146–160.
- JACOBSON, Barbara H. – JOHNSON, Alex – GRYWALSKI, Cynthia – SILBERGLEIT, Alice – JACOBSON, Gary – BENNINGER, Michael S. – NEWMAN, Craig W. 1997. The Voice Handicap Index (VHI). Development and Validation. *American Journal of Speech-Language Pathology* 6/3. 66–70.
- JARROLD, William – PEINTNER, Bart – WILKINS, David – VERGRYI, Dimitra – RICHEY, Colleen – GORNO-TEMPINI, Maria L. – OGAR, Jennifer 2014. Aided diagnosis of dementia type through computer-based analysis of spontaneous speech. In *Proceedings of CLPsych*.

- Baltimore, Maryland. 27–37. <https://www.aclweb.org/anthology/W14-3204.pdf> (A letöltés ideje: 2021. január 10.)
- JUNCOS-RABADÁN, Onésimo – PEREIRO, Arturo X. – FACAL, David – RODRÍGEZ, Nelly 2010. Research on language in mild cognitive impairment: A review. *Revista de Logopedia, Foniatría y Audiología* 30/2. 73–83.
- KÁLMÁN János – MAGLÓCZKY Erzsébet – JANKA Zoltán 1995. Óra Rajzolás teszt: gyors és egyszerű dementia szűrő módszer. *Psychiatria Hungarica* 10/1. 11–18.
- KÁLMÁN, János – PÁKÁSKI, Magdolna – HOFFMANN, Ildikó – DRÓTOS, Gergely – DARVAS, Gábor – BODA, Krisztina – BENCsik, Tamás – GYIMESI, Alíz – GULYÁS, Zsófia – BÁLINT, Magdolna – SZATLÓCZKI, Gréta – PAPP, Edina 2013. Early mental test – developing a screening test for mild cognitive impairment. *Ideggyógyászati Szemle* 66/1–2. 43–52.
- KAPLAN, Edith – GOODGLASS, Harold – WEINTRAUB, Sandra 1983. *Boston Naming Test*. Lea & Fibiger, Philadelphia, PA.
- KARAN, Biswajit – MAHTO, Kartik – SAHU, Sitanshu S. 2018. Detection of Parkinson disease using variational mode decomposition of speech signal. In *2018 International Conference on Communication and Signal Processing (ICCSP)*. IEEE. 508–512.
- KEMPLER, David 2004. *Neurocognitive Disorders in Aging*. Sage Publications, Thousand Oaks.
- KEMPLER, David – VAN LANCKER SIDTIS, Diana 2002. Effect of speech task on intelligibility in dysarthria: A case study of Parkinson’s disease. *Brain and Language* 80/3. 449–464.
- KERTESZ, Andrew 1982. *Western Aphasia Battery*. Grune & Stratton, Orlando, FL.
- KERTESZ, Andrew 2004. Frontotemporális demencia. *Psychiatria Hungarica* 19/6. 460–478.
- KIS Orsolya – TÓTH Alinka – JAKAB Katalin – KLIVÉNYI Péter 2020. A beszédsebesség vizsgálata Parkinson-kór-, sclerosis multiplex, valamint stroke-eredetű dysarthriák esetében. *Rehabilitáció: A Magyar Rehabilitációs Társaság folyóirata* 30/1. 3–10.
- KOLLER, William C. 1983. Dysfluency (stuttering) in extrapyramidal disease. *Archives of Neurology* 40/3. 175–177.
- KREMER, Berry – GOLDBERG, Paul – ANDREW, Susan E. – THEILMANN, Jane – TELENUS, Hakan – ZEISLER, Jutta – SQUITIERI, Fer – LIN, Biaoyang – BASSETT, Ann – ALMQVIST, Elizabeth – BIRD, Thomas D. – HAYDEN, Michael R. 1994. A worldwide study of the Huntington’s disease mutation. The sensitivity and specificity of measuring CAG repeats. *New England Journal of Medicine* 300/20. 1401–1406.
- KUJALA, Päivi – PORTIN, Raija – RUUTIAINEN, Juhani 1996. Language functions in incipient cognitive decline in multiple sclerosis. *Journal of the Neurological Sciences* 141/1–2. 79–86.
- LAAKSO, Katja – BRUNNEGARD, Karin – HARTELIUS, Lena – AHLSEN, Elisabeth 2000. Assessing high-level language in individuals with multiple sclerosis: A pilot study. *Clinical Linguistics and Phonetics* 14/5. 329–349.
- LASKE, Christoph – SOHRABI, Hamid R. – FROST, Shaun M. – LÓPEZ-DE-IPÍÑA, Karmele – GARRARD, Peter – BUSCEMA, Massimo – DAUWELS, Justin – SOEKADAR, Surjo R. – MUELLER, Stephan – LINNEMANN, Christoph – BRIDENBAUGH, Stephanie A. – KANAGASINGAM,

- Yogesan – MARTINS, Ralph N. – O'BRYANT, Sid E. 2014. Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer's & Dementia* 11/5. 1–18.
- LEDER, Steven B. 1996. Adult onset of stuttering as a presenting sign in a parkinsonian-like syndrome: A case report. *Journal of Communication Disorders* 29/6. 471–478.
- LETHLEAN, Jennifer B. – MURDOCH, Bruce E. 1993. Language problems in multiple sclerosis. *Journal of Medical Speech-Language Pathology* 1/1. 47–50.
- LETHLEAN, Jennifer B. – MURDOCH, Bruce E. 1997. Performance of subjects with multiple sclerosis on test of high-level language. *Aphasiology* 11/1. 39–57.
- LOUIS, Elan D. – WINFIELD, Linda – FAHN, Stanley – FORD, Blair 2001. Speech dysfluency exacerbated by levodopa in Parkinson's disease. *Movement Disorders: Official Journal of the Movement Disorder Society* 16/3. 562–565.
- LOGEMANN, Jeri A. – FISHER, Hilda B. – BOSHES, Benjamin – BLONSKY, E. Richard 1978. Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. *Journal of Speech and Hearing Disorders* 43/1. 47–57.
- LÓPEZ-DE-IPÍÑA, Karmele – ALONSO, Jesus-Bernardino – TRAVIESO, Carlos M. – SOLÉ-CASALS, Jordi – EGUIRAUN, Harkaitz – FAUNDEZ-ZANUY, Marcos – EZEIZA, Aitzol – BARROSO, Nora – ECAY-TORRES, Marcos – MARTINEZ-LAGE, Pablo – MARTINEZ DE LIZARDUI, Unai 2013. On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer Disease diagnosis. *Sensor* 13/5. 6730–6745.
- MANOUILIDOU, Christina – DE ALMEIDA, Roberto G. 2009. Canonicity in argument realization and verb semantic deficits in Alzheimer's disease. In FEATHERSTON, Susanne – WINKLER, Sam (eds): *The fruits of empirical linguistics, Vol. I: Process*. Mouton de Gruyter, Berlin. 123–149.
- MANOUILIDOU, Christina – DE ALMEIDA, Roberto G. – SCHWARTZ, Gordon F. – NAIR, Neelakanta P. V. 2009. Thematic hierarchy violations in Alzheimer's disease: The case of psychological verbs. *Journal of Neurolinguistics* 22/2. 167–186.
- MARKESBERY, William R. 1998. *Neuropathology of dementing disorders*. Arnold Publishers, London.
- MARTÍNEZ-SÁNCHEZ, Francisco F. – MEILÁN, Juan J. G. – GARCÍA-SEVILLA, Julia – CARRO, Juan – ARANA, José M. 2013. Oral reading fluency analysis in patients with Alzheimer disease and asymptomatic control subjects. *Neurología* 28/6. 325–331.
- MARTÍNEZ-SÁNCHEZ, Francisco – MEILÁN, Juan J. G. – CARRO, Juan – GÓMEZ ÍÑIGUEZ, Consolación – MILLIAN-MORELL, Lymarie – PUJANTE VALVERDE, Isabel M. P. – LÓPEZ-ALBURQUERQUE, Tomás – LÓPEZ, Dolores E. 2016. Speech rate in Parkinson's disease: A controlled study. *Neurología (English Edition)* 31/7. 466–472.
- MESULAM, M. Marsel 1982. Slowly progressive aphasia without generalized dementia. *Annals of Neurology* 11/6. 592–598.
- MEILÁN, Juan J. G. – MARTÍNEZ-SÁNCHEZ, Francisco – CARRO, Juan – SÁNCHEZ, José A. – PÉREZ, Enrique 2012. Acoustic markers associated with impairment in language processing in Alzheimer's disease. *Span Journal of Psychology* 15/2. 2081–2090.

- MESSINIS, Lambros – ANTONIADIS, George P. – BIRIS, Argyrios 2001. *Clinical symptomatology and speech impairments in Youth-onset Parkinson's Disease (YOPD) patients*. [http://psychology.gr.tripod.com/articles/youth\\_onset\\_parkinson\\_s\\_disease\\_en\\_ft.pdf](http://psychology.gr.tripod.com/articles/youth_onset_parkinson_s_disease_en_ft.pdf) (A letöltés ideje: 2017. szeptember 29.)
- MILLER, Pamela H. 2011. Dysarthria in multiple sclerosis. *Clinical Bulletin*. 1–12. <https://www.nationalmssociety.org/NationalMSSociety/media/MSNationalFiles/Brochures/Clinical-Bulletin-Dysarthria.pdf> (A letöltés ideje: 2021. január 11.)
- MINAGAR, Alireza – BARNETT, Michael H. – BENEDICT, Ralph H. B. – PELLETIER, Daniel – PIRKO, István – SAHRAIAN, Mohamad A. – FROHMAN, Elliott – ZIVADINOV, Robert 2013. The thalamus and multiple sclerosis. *Neurology* 80/2. 210–219.
- MORRIS, Robin G. 1996. *The cognitive neuropsychology of Alzheimer-type dementia*. Oxford University Press, Oxford.
- MURDOCH, Bruce E. 2010. Surgical approaches to treatment of Parkinson's disease: Implications for speech function. *International Journal of Speech-Language Pathology* 12/5. 375–384.
- NASREDDINE, Ziad S. – PHILLIPS, Natalie A. – BÉDIRIAN, Valérie – CHARBONNEAU, Simon – WHITEHEAD, Victor – COLLIN, Isabelle – CUMMINGS, Jeffrey L. – CHERTKOW, Howard 2005. The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* 53/4. 695–699.
- PELL, Marc D. – MONETTA, Laura 2008. How Parkinson's disease affects non-verbal communication and language processing. *Language and Linguistics Compass* 2/5. 739–759.
- PÁKÁSKI, Magdolna – DRÓTOS, Gergely – JANKA, Zoltán – KÁLMÁN, János 2012. Validation of the Hungarian version of Alzheimer's Disease Assessment Scale – Cognitive Subscale. *Orvosi Hetilap* 153/12. 461–466.
- PENKE, Botond – DATKI, Zsolt – ZARÁNDI, Márta 2003. Neurodegeneratív betegségek kémiai és biokémiai háttere. *Magyar Tudomány* 48/5. 607–614.
- PETERSEN, Roland C. – DOODY, Rachele – KURZ, Alexander – MOHS, Richard C. – MORRIS, John C. – RABINS, Peter V. – ROSSOR, Martin – THAL, Leon – WINBLAD, Bengt 2001. Current concepts in mild cognitive impairment. *Archives of Neurology* 58/12. 1985–1992.
- PETERSEN, Roland – SMITH, Glenn E. – WARING, Stephen C. – IVNIK, Robert J. – TANGALOS, Eric G. – KOKMEN, Emre 1999. Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology* 56/3. 303–308.
- PINANGO, Maria M. 2006. Thematic roles as event structure relations. In BORNKESSEL, Ina – SCHLESEWSKY, Matthias – FRIEDERICI, Angela D. (eds): *Semantic role universals and argument linking: Theoretical, typological, and psycholinguistic perspectives*. Mouton, Berlin. 303–325.
- RAINER, Anton – DRAGATIS, Ioannis – DIETRICH, Paula 2011. Genetics and neuropathology of Huntington's Disease. *International Review of Neurobiology* 98. 325–372.
- RAPP, Alexander – WILD, Barbara 2011. Nonliteral language in Alzheimer Dementia: A review. *Journal of the International Neuropsychological Society* 17/2. 207–218.

- READ, Jennifer – MILLET, Nick – KITSOU, Nikoletta 2018. Is there an order of loss of sounds in speakers with Parkinson's disease? *Clinical Linguistics & Phonetics* 32/11. 997–1011.
- RENAULD, Stéphanie – MOHAMED-SAID, Lyakout – MACOIR, Joël 2016. Language disorders in multiple sclerosis: A systematic review. *Multiple Sclerosis and Related Disorders* 10. 103–111.
- RITCHIE, Karen – ARTERO, Sylvaine – TOUCHON, Jacques 2001. Classification criteria for mild cognitive impairment: A population-based validation study. *Neurology* 56/1. 37–42.
- ROALF, David R. – MOBERG, Paul J. – XIE, Sharon X. – WOLK, David A. – MOELTER, Stephen T. – ARNOLD, Steven E. 2013. Comparative accuracies of two common screening instruments for classification of Alzheimer's disease, mild cognitive impairment, and healthy aging. *Alzheimer's & Dementia* 9/5. 529–537.
- ROARK, Brian – MITCHELL, Margaret – HOSOM, John-Paul – HOLLINGSHEAD, Kristy – KAYE, Jeffrey 2011. Spoken language derived measures for detecting mild cognitive impairment. *IEEE Transactions on Audio Speech and Language Processing* 19/7. 2081–2090.
- RODGERS, Jonathan D. – TJADEN, Kris – FEENAUGHTY, Lynda – WEINSTOCK-GUTTMAN, Bianca – BENEDICT, Ralph H. 2013. Influence of cognitive function on speech and articulation rate in multiple sclerosis. *Journal of the International Neuropsychological Society* 19/2. 173–180.
- ROSEN, Wilma G. – MOHS, Richard C. – DAVIS, Kenneth L. 1984. A new rating scale for Alzheimer's disease. *American Journal of Psychiatry* 141/11. 1356–1364.
- RUBINSZTEIN, David C. – LEGGO, Jayne – COLES, Rhian – ALMQVIST, Elisabeth – BIANCALANA, Valerie – CASSIMAN, Jean-Jacques – CHOTAI, Kokila – CONNARTY, Margaret – CRAUFURD, David – CURTIS, Anne – CURTIS, Diana – DAVIDSON, Mark J. – DIFFER, Anne-Marie – DODE, Catherine – DODGE, Alan – FRONTALI, Marina – RANEN, Neal G. – STINE, O. Colin – SHERR, Meeia – ABBOTT, Margaret H. – FRANZ, Mray L. – GRAHAM, Colin A. – HARPER, Peter S. – HEDREEN, John C. – JACKSON, Anthony – KAPLAN, JeanClaude – LOSEKOOT, Monique – MACMILLAN, John C. – MORRISON, Patrick – TROTTIER, Yvon – NOVELLETTO, Andrea – SIMPSON, Sheila A. – THEILMANN, Jane – WHITTAKER, Joanne L. – FOLSTEIN, Susan E. – ROSS, Christopher A. – HAYDEN, Michael R. 1996. Phenotypic characterization of individuals with 30–40 CAG repeats in the Huntington Disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36–39 repeats. *American Journal of Human Genetics* 59/1. 16–22.
- RUSZ, Jan – CMELJA, Roman – TYKALOVA, Tereza – RUZICKOVA, Hana – KLEMPER, Jiri – MAJEROVA, Veronika – PICMAUSOVA, Jana – ROTH, Jan – RUZICKA, Evzen 2013. Imprecise vowel articulation as a potential early marker of Parkinson's disease: Effect of speaking task. *The Journal of the Acoustical Society of America* 134/3. 2171–2181.
- SATT, Aharon – HOORY, Ron – KÖNIG, Alexandra – AALTEN, Pauline – ROBERT, Philippe 2014. Speech-based automatic and robust detection of very early dementia. In *Interspeech 2014*. doi: 10.13140/2.1.1258.8805.
- SCHRAG, Anette – HOVRIS, Anna – MORLEY, David – QUINN, Niall – JAHANSHAH, Marjan 2003. Young- versus older-onset Parkinson's disease: Impact of disease and psychosocial



- consequences. *Movement Disorders: Official Journal of the Movement Disorder Society* 18/11. 1250–1256.
- SNYDER, Gregory J. – WILLIAMS, Molly G. – GOUGH, Molly E. – BLANCHET, Paul G. 2018. Fluency-enhancing strategies for hypokinetic dysarthria exacerbated by subthalamic nucleus brain stimulation: A case study. *Perspectives of the ASHA Special Interest Groups* 3/4. 4–16.
- SKODDA, Sabine – SCHLEGEL, Uwe 2008. Speech rate and rhythm in Parkinson's disease. *Movement Disorders: Official Journal of the Movement Disorder Society* 23/7. 985–992.
- SKODDA, Sabine – VISSER, Wenke – SCHLEGEL, Uwe 2011. Vowel articulation in Parkinson's disease. *Journal of Voice* 25/4. 467–472.
- SZATLÓCZKI, Gréta – HOFFMANN, Ildikó – VINCZE, Veronika – KÁLMÁN, János – PÁKÁSKI, Magdolna 2015. Speaking in Alzheimer's disease, is that an early sign? Importance of changes in language abilities in Alzheimer's disease. *Frontiers in Aging Neuroscience* 7. 195. <https://doi.org/10.3389/fnagi.2015.00195>.
- SZÉPE, Judit 2009. *Phonoparaphasias*. Black & Black Publishers, Toronto.
- SZIRMAI Imre (szerk.) 2006. *Neurológia*. Medicina, Budapest.
- TARISKA Péter (szerk.) 2002. *Kortünet vagy kórtünet? Mentális zavarok az időskorban*. Medicina, Budapest.
- TALER, Vanessa – PHILLIPS, Natalie A. 2008. Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. *Journal of Clinical and Experimental Neuropsychology* 30/5. 501–556.
- TJADEN, Kris – WATLING, Greg 2011. Speech and pause characteristics associated with voluntary rate reduction in Parkinson's disease and multiple sclerosis. *Journal of Communication Disorders* 44/6. 655–665.
- TÓTH, László – HOFFMANN, Ildikó – GOSZTOLYA, Gábor – VINCZE, Veronika – SZATLÓCZKI, Gréta – BÁNRÉTI, Zoltán – PÁKÁSKI, Magdolna – KÁLMÁN, János 2018. A speech recognition-based solution for the automatic detection of Mild Cognitive Impairment from spontaneous speech. *Current Alzheimer Research* 15/2. 130–138.
- TSANTALI, Eleni – ECONOMIDIS, Dimitris – TSOLOAKI, Magdalini 2013. Could language deficits really differentiate Mild Cognitive Impairment (MCI) from mild Alzheimer's disease. *Archives of Gerontology and Geriatrics* 57/3. 263–270.
- TÚRI Zsolt – NÉMETH Dezső – HOFFMANN Ildikó 2014. Nyelv és emlékezet. In PLÉH Csaba – LUKÁCS Ágnes (szerk.): *Pszicholingvisztika 1–2*. Akadémiai Kiadó, Budapest. 743–776.
- ULLMAN, Michael T. 2001. The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research* 30/1. 37–69.
- VERTESI, Andrea – LEVER, Judith A. – MOLLOY, D. William – SANDERSON, Brett – TUTTLE, Irene – POKORADI, Laura – PRINCIPI, Elaine 2001. Standardized Mini-Mental State Examination. Use and interpretation. *Canadian Family Physician* 47. 2018–2023.

- VINCZE, Veronika – SZATLÓCZKI, Gréta – TÓTH, László – GOSZTOLYA, Gábor – PÁKÁSKI, Magdolna – HOFFMANN, Ildikó – KÁLMÁN, János 2020. Telltale silence: Temporal speech parameters discriminate between prodromal dementia and mild Alzheimer’s disease. *Clinical Linguistics and Phonetics*. 1–16. doi: 10.1080/02699206.2020.1827043.
- WALKER, Harrison C. – PHILLIPS, Daniel E. – BOSWELL, Deborah B. – GUTHRIE, Barton L. – GUTHRIE, Stephanie L. – NICHOLAS, Anthony P. – MONTGOMERY, Erwin B. – WATTS, Ray L. 2009. Relief of acquired stuttering associated with Parkinson’s disease by unilateral left subthalamic brain stimulation. *Journal of Speech, Language, and Hearing Research* 52/6. 1652–1657.
- WALLACE, Gloriajean L. – HOLMES, Stephen 1993. Cognitive-linguistic assessment of individuals with multiple sclerosis. *Archives of Physical Medicine and Rehabilitation* 74/6. 637–643.
- WUTZLER, Alexander – BECKER, Regine – LÄMMLER, Gernot – HAVERKAMP, Wilhelm – STEINHANEN-THIESSEN, Elisabeth 2013. The anticipatory proportion as an indicator of language impairment in early-stage cognitive disorder in the elderly. *Dementia and Geriatric Cognitive Disorders* 36/5–6. 300–309.
- WYLEZINSKA, Marzena – CIFELLI, Alerio – JEZZARD, Peter – PALACE, Jacqueline – ALECCI, Marcello – MATTHEWS, Paul M. 2003. Thalamic neurodegeneration in relapsing-remitting multiple sclerosis. *Neurology* 60/12. 1949–1954.

## KÖSZÖNETNYILVÁNÍTÁS

A tanulmány megírását a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal NKFIH-K-132460 számú pályázata és a Tématerületi Kiválósági Program támogatta.



# Az orofaciális miofunkcionális diszfunkció és a nyelvlökéses nyelés

Havadi-Nagy Marian

ELTE EÖTVÖS LORÁND TUDOMÁNYEGYETEM  
SZEGEDI TUDOMÁNYEGYETEM

Deme Andrea

ELTE EÖTVÖS LORÁND TUDOMÁNYEGYETEM  
MTA–ELTE LENDÜLET LINGVÁLIS ARTIKULÁCIÓ KUTATÓCSOPORT

## 1. Az orofaciális funkciók

**Orofaciális funkciónak** (azaz száji-arcú működéseknél) tekintjük a rágást, a nyelésnek a szájüregben lezajló szakaszát és a beszéd létrehozását. Az emberi törzsféjlődés folyamatában a nyelési funkció kialakulása megelőzte a beszédképzést, vagyis az artikuláció és a fonáció létrehozásának képességét (MÉSZÁROS–HACKI 2013).

Az orofaciális területen az orális mozgásokat, azaz az ún. **orálmotoros** tevékenységet az ajkak, az állkapocs és a nyelv végzik. A mozgások kivitelezésében fontos szerepe van az izomműködés tónusának, az izomerőnek, a mozgás sebességének, a koordinációnak, valamint magának a mozgástartománynak. BEAN (2013) disszociációnak nevezi az egyes gége fölötti vagy másként szupraglottális beszédszervek egymástól függetlenített mozgásának képességét, amelynek az érett működésekben fontos szerepe van. A logopédiai praxisban a disszociáció képességének az áll- és nyelvmozgás, továbbá a nyelv- és ajakmozgás szétválasztásában van jelentősége: e szervek elkülönített működtetése szükséges a nyeléshez, illetve az egyes beszédhangok ejtéséhez is. ARVEDSON és munkatársai (2020) szerint az orálmotoros mozgások elsajátítása, érése és féjlődése alapját képezi a hangképzésnek és a táplálkozásnak, így az anyamellről táplálásnak, a harapásnak és rágásnak egyaránt. Atipikus orálmotoros mintázatnak felel meg az alsó állkapocs és a nyelv előretolása, az ajkak és a nyelv visszafelé húzása, valamint az orrüregben fellépő regurgitáció, azaz a folyadék vagy táplálék visszaáramlása az orrüregben keresztül (MORRIS 1978).

A következőkben a fenti három orofaciális funkciót tekintjük át, elsőként a normál, normatív vagy másként szabályos működés, majd a nem normatív, atipikus vagy szabálytalan működés esetén. Ezek a funkciók szorosan összefüggnek féjlődésükben és működésükben. Általánosan elmondható, hogy a táplálás és táplálkozás során végzett motoros működés (azaz a rágás és a nyelés) már a korai gyermekkorban hozzájárul a szupraglottális, azaz a gége fölötti beszédszervek (elsősorban az ajkak és a nyelv) izomegyensúlyának és mozgáskoordinációjának féjlődéséhez,

és a rágóizomzat megfelelő terhelése lényegében előkészíti a beszédképzés artikulációs működéseit. Látni fogjuk azonban, hogy az egyes rendszerek esetlegesen szabálytalan működése is összefügghet, így például a nyelés zavara együtt járhat beszédképzési problémákkal.

A tanulmányunkban, különösen a szabálytalan nyelést illetően, elméleti okokból az első szerző által javasolt újabb terminológiát használjuk a magyar nyelvű logopédiai szövegekben elterjedt kifejezések helyett. Erre utal a címbeli kettősség is, ahol egy a kurrens nemzetközi szakirodalomban használt, de a magyar logopédiai praxisban még nem széles körben elterjedt terminus, az *orofaciális miofunkcionális diszfunkció* áll párban a magyar logopédiai szövegekben hagyományosan használt *nyelvlökéses nyelés* kifejezéssel hasonló jelentéskörre értett terminusokként. A szakirodalmi tájékozódás elősegítésére természetesen említeni fogjuk a logopédiában korábban elterjedt szakkifejezéseket is az újabb terminusok mellett, valamint a terminológiai kérdéseknek egy külön fejezetet is szentelünk.

## 1.1. A rágás folyamata

Mind a rágás, mind a nyelés a nyelv, a lágy szájpad, a garatüreg és az arc izmainak, továbbá a gége és a légzőrendszernek az összehangolt működését igényli. A rágást a legtöbb leírásban a nyelés első fázisának tekintik, mely során a táplálékot előkészítjük a nyeléshez (vö. pl. MÉSZÁROS–HACKI 2013). Ezzel együtt a szabályos, ún. **érett nyelés folyamatát** négy fázisra szokás osztani: az orális előkészítő fázisra (harapás és rágás), az orális transzportfázisra, a garati (faringális) fázisra és a nyelőcsövi (özofagális) fázisra (MÉSZÁROS–HACKI 2013). A nyelés első két szakasza, vagyis a harapás és rágás (az orális előkészítő fázis), valamint az orális transzportfázis olyan részben reflexszerű, automatizáltan működő tevékenységek, amelyek (legalábbis részben) szándékosan is irányíthatók (SEIKEL–KING–DRUMRIGHT 2010). A garati és a nyelőcsövi szakaszok azonban akarunktól teljesen függetlenül, reflexszerűen zajlanak (MÉSZÁROS–HACKI 2013). A folyamat további, a rágás utáni szakaszairól az *1.2. fejezetben* lesz részletesebben szó. Hangsúlyozzuk, hogy a jelen tanulmányban csak a már szilárd tápanyag feldolgozására is alkalmas érett rágás és nyelés folyamatait tárgyaljuk, a csecsemőkori működéseket nem.

A rágás folyamatában a harapás után az ételdarabot vagy folyadékot a szánkba vesszük, az ajkakat szorosan összezárjuk, hogy a falatot bent tartsuk, majd a még meg nem rágott falatot a fogak őrlefelületéhez irányítjuk, megőröljük és nyállal keverjük el. Ebben a folyamatban a nyelv is aktív szerepet játszik, hiszen az őrlés és keverés folyamatához előre-hátra mozgatja a feldolgozandó ételt, emellett pedig az arcizmok akadályozzák meg azt, hogy az étel az íny és az orcák által közrezárt üregbe jusson (SEIKEL–KING–DRUMRIGHT 2010). A rágás révén létrejött, felaprított, nyelésre előkészített, megfelelő méretű és sűrűségű étel vagy folyadékadag tudományos neve **bolus** (MÉSZÁROS–HACKI 2013) – nyeléskor ezt vagy ún. üres nyelés esetén az összegyűjtött nyáladagot továbbítjuk a gyomor felé.

A rágás motoros működéseiben a következő izmok és idegek játszanak szerepet. A rágóizomzatot az V. agyideg (*nervus trigeminus*) nervus mandibuláris ágcsoportja idegzi

be. A fogsorok ritmikus záródásáért és nyitódásáért, illetve az alsó állcsont előre-hátra irányuló mozgásáért, melyek a falat őrléséhez szükségesek, a rágóizom (*musculus masseter*) és a röpipizom (*musculus pterygoideus*) felelősek a *musculus temporalis* közreműködésével (SZENTÁGOTHAI–RÉTHELYI 2006a, 2006b). A száj körüli izmok közül a trombitásizom (*musculus buccinator*) feladata, hogy a nyelvvel együtt rágás közben a táplálékot a fogak közé továbbítsa (SZENTÁGOTHAI–RÉTHELYI 2006a, 2006b), és a falatot ne engedje az íny és a pofa (vagy orca, latinul *bucca*) közé ékelődni, illetve a bejutott szájartalmat kijuttassa onnan (vissza a fogak közé) (SEIKEL–KING–DRUMRIGHT 2010). Az ízlelés és szaglás, valamint a szájüregi nyomásérzékelés (tapintás), fájdalom- és hőmérséklet-érzékelés szintén fontos szerepet játszanak a rágás és nyelés folyamataiban az egyes reflexek kiváltóiként (vö. SEIKEL–KING–DRUMRIGHT 2010), de ezek részletes ismertetésétől e helyütt eltekintünk.

## 1.2. A nyelés folyamata

Amint arra már fentebb utaltunk, a szabályos nyelés folyamata négy lépésből áll: az **orális előkészítő fázisból** (rágás), az **orális transzportfázisból**, a **faringális fázisból** és az **özofagális fázisból** (MÉSZÁROS–HACKI 2013). A továbbiakban az orális előkészítés utáni lépéseket ismer-tetjük nyelés címén.

A nyelés indítása és lefolyása lehet akaratlagos vagy reflexes, és a tényleges megindulásához az szükséges, hogy legyen valami továbbítható tartalom a szájüregben (FONYÓ 2011). A nyelést az adott fázist meghatározó izomműködések szempontjából száji és garati (*orofaringális*) és nyelő-csővi (*özofagális*) szakaszokra is oszthatjuk (FONYÓ 2011). A feltétlen reflexes légzés és a nyelés szétválasztását az ezeknek egyaránt bemenetül szolgáló, így kettős funkcióval bíró szájüreg és garatüreg végzi a központi idegrendszer irányításával.

Amint azt az *1.1. fejezet* bemutatta, az orális előkészítő szakaszban valósul meg a nyál-gyűjtés, illetve a harapás és rágás, azaz a nyelésre alkalmas méretű falat létrehozása vagy a fo-lyadék kortyokban történő felvétele.

A nyelés ezt követő fázisában, az orális transzportfázisban történik meg a szilárd vagy folyékony állagú bolus és/vagy a nyál hátrajuttatása a garathoz a nyelv segítségével. Szabályos nyelés esetén az orális transzportfázisban a nyelvtest a garat felé végez hullámzó mozgást úgy, hogy eb-ben a következő mozgásegységek jelennek meg: a nyelv hátrahúzódik, nekinyomódik a kemény szájpadnak, majd a lágy szájpadhoz préselődik, a nyelvhat lesüllyed, ezáltal a bolus/nyál hátrébb kerül, míg végül eléri a garati szakaszt (BjÖRK 1960, idézi VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015).

Amikor a szilárd vagy folyékony állagú bolus vagy nyál az első garatívhez jut, kiváltódik a faringális fázist működtető nyelési reflex. A nyelés a garati szakaszban teljesen reflexszerűen megy végbe. E folyamat során a lágy szájpad megemelkedik és bezáródik, a gége szintén meg-emelkedik és előremozdul (valamint a légutak védelme érdekében bezáródik), a garat kiszéle-sedik, és az így létrejött negatív nyomás segítségével szívó hatást fejt ki a bolusra. A nyelési

reflex aktivációjával a légzés is leáll. A gége előremozdulásával a nyelvőcsőbemenet kitér és ellazul, majd ha a falat/nyál a nyelvőcsőbe került, a bemenet ismét bezáródik. A falatnak/nyálnak a nyelvőcsőbe jutásával szinte egy időben a gége is lesüllyed az eredeti helyzetébe, és a légzés is újraindul. Míg az orális transzportfázis mintegy 1–1,5 s, a faringális vagy garati szakasz körülbelül 1 s (SEIKEL–KING–DRUMRIGHT 2010; MÉSZÁROS–HACKI 2013).

A nyelés folyamatának végén a falat/nyál a nyelvőcsőbe jut, ahol egy időarányosan hosszabb (egyes források szerint 4–20 s, lásd MÉSZÁROS–HACKI 2013, mások szerint 10–20 s, lásd SEIKEL–KING–DRUMRIGHT 2010) működés eredményeként a nyelvőcső perisztaltikus (hullámszerű) mozgásainak segítségével eljut a gyomorig (SEIKEL–KING–DRUMRIGHT 2010; MÉSZÁROS–HACKI 2013).

A nyelés anatómiai háttere a feji és nyaki régióban található csontos és porcos vázszerkezet, a harántcsikolt izomzat, valamint a kapcsolódó neurális működés. A neuromuszkuláris működés idegi kontrollját képezik az efferens és afferens motoros pályák, az agy, a középgagy és a kisagy szinapszissai a középgagyban található „nyelésközpontokkal”, valamint az agytörzs „nyelési központjaival” (DODDS–STEWART–LOGEMANN 1990). (Megjegyezzük, hogy mivel a *központ* kifejezés valamely kompakt neuroncsoportra utal, a kurrens szaknyelv a fentebb idézőjelek között szereplő, hagyományosan használt *nyelési központ* terminus helyett a *központi nyelési mintázatgenerátor* terminust használja, vö. FONYÓ 2011).

### 1.3. A beszédképzés

A beszédképzés folyamatában sorrendben a légvételi (**respirációs**), zöngképzési (**fonációs**) és hangképzési (**artikulációs**) szakaszokat különböztetjük meg. A belégzést követő kilégzés során a tüdőből kiáramló levegő a gégebe kerül, ahol megrezegteti a hangjakakat, és létrejön egy komplex, kváziperiodikus hang, a zöngé. A gége fölötti (szupraglottális) üregrendszer vagy más néven toldalékcső rezonátorként működik a beszédképzés folyamatában az alakjának és méretének megfelelő sajátrezonanciákkal. Ez azt jelenti, hogy a toldalékcsőbe sugárzott zöngé egyes összetevői felerősödnek, mások csillapítódnak, azaz az üregrendszer megszűri a gégeben keletkezett zöngéhangot. A toldalékcsővet három üreg alkotja: a garatüreg, a szájüreg és az orrüreg; ezek közül a szájüreg keretét az alsó és a felső állcsont (a *mandibula* és a *maxilla*) képezi (SZENTÁGOTHAI–RÉTHELYI 2006a, 2006b). Mivel a toldalékcső alakja és mérete a mobilis beszéd szervek (az ajkak, az alsó állkapocs, a nyelv és a lágyszájpad) mozgásának, valamint a gége függőleges helyzetének segítségével bizonyos mértékben változtatható, ezért rezonanciatulajdonságai is befolyásolhatók. Az artikuláció folyamatának a lényege pedig éppen ez: a beszélő a különböző toldalékcsovi beállításokkal a végül a szájon keresztül kisugárzott hang hangszínét változtatja a különböző beszédhangok létrehozásához. Az artikulációs működéseket (így az egyes beszédhangokat) általában úgy jellemezzük, hogy megadjuk, hogy a mozgatható, aktív artikulátorok (az ajkak, az állkapocs, a nyelv és a szájpad) közül melyek hol és hogyan közelítik meg a nem mozgatható, passzív artikulátorokat (az alsó és felső fogsort, a fogmedret, a kemény szájpadot és lágyszájpadot), illetve jelezzük, hogy a hangadás során hol távozik a levegő (ti. az orr- és/vagy a szájüregen keresztül). A közeledés/érintkezés

helyére utal a képzési hely terminus, míg a közeledés/érintkezés, azaz az akadályképzés módjára (összekapcsolva a levegőkiáramlás útjával) a képzési mód kifejezés (vö. pl. DEME 2016; MARKÓ 2017).

Abból következően, hogy a beszéd létrehozása a rágáshoz és nyeléshez hasonlóan az orofaciális területen mint szerkezeten belül megy végbe, továbbá a beszéd artikulációs működéseinek kivitelezését is a rágásért és nyelésért is felelős orofaciális motoros működések végzik, sejthető, hogy a rágás és nyelés szabálytalanságai a beszédhangképzés problémáival járhatnak együtt. A logopédiai tapasztalat pedig valóban azt mutatja, hogy a rágásban-nyelésben részt vevő rendszerek esetlegesen szabálytalan működése gyakran jár együtt beszédképzési problémákkal. Erről a jelen tanulmány 3. fejezetében írunk részletesebben.

## 2. A nyelés zavarai

### 2.1. Diszfágia

A rágás és nyelés komplex folyamatok, melyek – mint láttuk – számos részfolyamat és struktúra összehangolt működését igénylik. Ezzel összefüggésben előfordulhat, hogy a folyamat kivitelezése problémába ütközik, diszharmonia lép fel. A klinikumban **diszfágiának** nevezik a nehezített, fájdalmas nyelést, illetve félrenyelést, azaz a táplálék orális előkészítésének, illetve orális, faringális vagy özofagális transzportjának zavarát (MÉSZÁROS – HACKI 2013). A diszfágia kialakulásának számos oka lehetséges: idegrendszeri tényezők (pl. agyi keringési zavar), strukturális elváltozások (daganat, az egyes struktúrák károsodása sugárterápia során vagy hiányuk, ha ezek a daganat miatt műtéti eltávolításra kerültek), gyulladás, illetve lélektani tényezők (a nyeléstől való félelem) (MÉSZÁROS – HACKI 2013).

A (klinikai) logopédia a száj- és garatüregi működéseket érintő orofaringális diszfágiával foglalkozik. A nyelőcsődiszfágia (özofagális diszfágia) kezelése szakorvosi feladat az alapbetegség (pl. reflux) gyógyításával. A nyelőcsődiszfágiát a lokalizált neuromuszkuláris rendelkezésségek vagy obstruktív elváltozások okozzák (KRUGER 2014).

### 2.2. Az orofaciális miofunkcionális diszfunkció és a nyelvlökéses nyelés

#### 2.2.1. Elnevezések

A szabálytalan nyelésre a szakirodalom szerteágazó nevezéktannal utal. A nem normatív vagy szabálytalan nyelést a legfrissebb külföldi források az **orofaciális** ('száji-arci') **miofunkcionális** ('az izomműködést érintő') **diszfunkció** ('zavar') (*orafacial myofunctional disorder, OMD*) összefoglaló terminus alatt említik annak egy lehetséges vetületeként (vö. pl. BILLINGS et al.



2018). A korábbi nemzetközi forrásokban emellett olvashatjuk még az orális miofunkcionális zavar (*oral myofunctional disorder*) kifejezést (vö. BARRETT–HANSON 1974), valamint a *tongue thrust* (BRAUER–HOLT 1965) kifejezést is. A magyar szakirodalomban legnagyobb részben a nyelvlökéses nyelés megnevezéssel találkozunk a szabálytalan nyelés bármely formájára értve (vö. VASSNÉ KOVÁCS – REHÁK – VIZKELETY 2015).

Amint az a fentiekből is sejthető, a használt fogalmak változatossága legalábbis részben annak köszönhető, hogy egyes terminusok szűkebben, mások tágabban határozzák meg a nyelés működésének zavarait. Értelmezésünkben a legtágabb fogalom, mely bármilyen nem szervi eredetű eltérésből fakadó száji-arc (így pl. nyelési) rendellenességet magában foglalhat, az **orofaciális miofunkcionális diszfunkció**, az OMD (ejtsd: [o:emde:]). Az OMD definíció szerint olyan az arc és száji izmok érintő (lényegében szabálytalan) mintázatok neve, melyek megzavarják az arc-száji struktúrák normális növekedését, fejlődését, illetve működését, vagy (szabálytalan működésükkel) „felhívják magukra a figyelmet” (BILLINGS et al. 2018: 1). Mindez tehát azt jelenti, hogy az OMD-ként rövidített terminus nemcsak a nyelés szabálytalanságait foglalja magában, hanem tágabban az arc-száji izmok működésének bármely területét is, így például a nem táplálkozási célú szopást, ajak- vagy szájszopogatást (vö. BILLINGS et al. 2018). Ugyanakkor a hazai szakirodalom külön tárgyalja, és a nyelésfunkció eltérését kiváltó okként kezeli például ezeket az ún. rossz szokásokat, melyek közül a leggyakoribb az ujjszopás (FÁBIÁN–GÁBRIS–TARJÁN 2013).

Az OMD kiemelten a bukkális (orcák) és periorális (ajak körüli) területet, valamint az ajkakat, a palatumot és a fogsorokat érinti más és más mértékben. Akkor, amikor az OMD a nyelési funkciót érinti, **nyelési OMD**-ként hivatkozhatunk rá – így teszünk a tanulmány további részeiben. Abban az esetben pedig, amikor döntően a nyelvmozgás szabálytalansága a meghatározó benne, **lingvális orofaciális miofunkcionális diszfunkcióról** (LOMD, ejtsd: [el:o:emde:]) beszélhetünk, mely a nyelési OMD egyik legnagyobb előfordulási aránnyal bíró alcsoportja (HANSON–MASON 2003). (A nyelési OMD és az LOMD megnevezéseket tehát mi itt már csak a nyelési funkcióval összefüggésben említjük, más zavarokra nem használjuk.)

Az LOMD esetén a nyelv nyelés közben a szokásosnál nagyobb nyomást fejt ki a fogazatra (különösen az elülső fogakra). Ebből kitűnik, hogy ez a fogalom az, amelyet a magyar szakirodalomban olvasható kifejezések közül a nyelvlökéses nyelésnek leginkább megfeleltethetünk. Fontos azonban látni, hogy a hazai fogalomhasználatban gyakran nyelvlökéses nyelésnek neveznek minden nyelési szabálytalanságot, tekintet nélkül arra, hogy mely beszédszervek és milyen mértékben mutatnak abban szabálytalan működést – így ezt a kifejezést találjuk például az ajaklökéssel létrehozott nyelésre használva is (vö. VASSNÉ KOVÁCS – REHÁK – VIZKELETY 2015).

A fentiek értelmében itt is és korábbi munkáinkban is a nyelési OMD, illetve az LOMD kifejezésekkel hivatkozunk a szabálytalan nyelésre (vö. HAVADI-NAGY 2020a, 2020b). LOMD-nek tekintjük a nyelés orális transzportfázisában fellépő zavart, mely során a nyelv a hátrafelé, a garati szakasz felé irányuló mozgás helyett előre, oldalra vagy előre-oldalra irányuló mozgást végez (vö. BRAUER–HOLT 1965; BILLINGS et al. 2018).

Megjegyezzük, hogy a nyelvlökéses nyelésnek az OMD egy lehetséges típusaként való felismerése nem újdonság a szakirodalomban (vö. SCHREY-DERN 2006, idézi FEHÉRNÉ KOVÁCS

2013, aki szintén az orofaciális funkciózavarok körébe sorolja a nyelvlökéses nyelést, organikus okkal), ám a nyelési OMD, illetve az LOMD (lásd lentebb) megnevezések használata a tárgyalat jelenségekre annak minősül.

### 2.2.2. A nyelési OMD lehetséges okai és tünetei

BILLINGS és munkatársai (2018) életkori csoportok szerint ismertetik az OMD (tehát bármely arciszáji izomműködést érintő diszfunkció) lehetséges tüneteit a születés időpontjától az iskolás- és felnőttkorig. A 3 (pl. BILLINGS et al. 2018) vagy máshol 4-4,5 (FÁBIÁN–GÁBRIS–TARJÁN 2013) éves kor alatti korosztály az **élettani** vagy fiziológiás nyelvlökéses nyelés (FÁBIÁN–GÁBRIS–TARJÁN 2013) vagy másként az **infantilis nyelés** (VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015) időszakában van, tehát ebben az életkorban ez a felnőtteknél szabálytalannak minősülő nyelési forma még teljesen természetes. A fent említett okokra hivatkozva az erre a korszakra jellemző nyelést a jelen tanulmány szerzői élettani vagy **fiziológiás nyelési OMD**-nek nevezik. Mivel ebben a korosztályban ez a nyelésítípus még nem tekinthető szabálytalannak, nem beszélhetünk ennek lehetséges okairól sem. Ám már itt is megfigyelhetők lehetnek olyan tünetek vagy hajlamosító tényezők, melyek a később lehetségesen előforduló nyelési OMD-re utalhatnak. Ilyenek lehetnek például a légzés problémái (gátolt orrlégzés, szájlégzés), a koponyával és arccal kapcsolatos (*craniofaciális*) zavarok, morfológiai eltérések (kiemelten pl. harapási rendellenesség), a szenzomotoros diszfunkció, az alacsony szinten fejlett orális és orofaringális izomtónus, a nyelvfelek diszfunkciója (a túl rövid nyelvfelek, azaz az ún. lenőtt nyelv), az egyes kedvezőtlen orális és evési szokások (pl. ujjszopás, cumizás, körömrágás, illetve túl gyors és/vagy elégtelen mennyiségű rágás), illetve egyéb egyéni jellemzők.

Az életkorban soron következő, a 4-4,5 évesnél idősebb korosztályban megjelenő szabálytalan nyelésmintázatokat (a fent bevezetett paradigmába illeszkedően) összefoglalóan **fejlődési nyelési OMD**-nek nevezzük. Ez a terminus azt az esetet fedi le, amikor elmarad az érett, normatív nyelés kialakulása az idegrendszeri érés folyamatában. (Bár nem ezt a terminust használva, de a jelenséget számos forrás leírja, vö. FÁBIÁN–GÁBRIS–TARJÁN 2013; VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015).

Előfordulhat, hogy a már elsajátított, rögzült normatív nyelés válik valamely okok miatt szabálytalanná, akár még felnőttkorban is (VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015). WIEMER (2017) ennek hátterében a kiváltó okok széles körét ismerteti, így az időskorban bekövetkező izomtónust és fogállományt érintő elváltozásokat, a Parkinson-kórt, a demenciát, valamint olyan traumákat és egyéb megbetegedéseket, amelyek módosíthatják az orofaciális terület, benne a nyelésfunkció működését. A szabályos nyelés szabálytalanná válására a szakirodalomban nem találunk megnevezést. A jelen tanulmány szerzői erre a **szerzett nyelési OMD** terminus használatát vezetik be.

A szabálytalan nyeléssel foglalkozó szerzőknél egyöntetűen megjelenik az etiológiai (oki) tényezők feltérképezésére irányuló igény, mely a nyelészavarok vizsgálatában az anamnézis részben szerepel. Amint arról részben már fentebb is tettünk említést, a beszédfejlődést befolyásoló kezeletlen vagy vissza-visszatérő felső légúti betegségek, az allergiák, a gátolt orrlégzés

gyakorta megfigyelt okok az atipikus nyelés kialakulásában. Emellett említendők még az egyes káros orális szokások (VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015), vagy másként a nem táplálási célú szopási viselkedések (*nonnutritive sucking behaviors*, vö. WIEMER 2017), mint például az ujjszopás, ajaknyalogatás, nyelv szopogatás. Ugyanakkor sejthető, hogy a nyelési OMD kialakulásában szerepe van a genetikának, illetve az utánzásnak is, mert előfordul, hogy halmozódik a családban, továbbadódik a generációk között (FÁBIÁN–GÁBRIS–TARJÁN 2013).

A nyelési OMD kialakulását befolyásolhatja a csecsemőkori etetés és a táplálkozás is. Ismeretes, hogy az anyamellről táplálás során a nyelvtest emelkedése és süllyedése hozza létre a szopáshoz szükséges vákuumot (ARVEDSON–BRODSKY–LEFTON-GREIF 2020), és hogy ezzel összefüggésben mind az anyamellről táplálás, mind pedig aztán később a kemény ételek rágása edzi az arci-száji izmokat (VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015). A szomatikus fejlődés alapja az életkornak megfelelő táplálkozás, így jön létre a nyelés szenzomotoros integrációja, a megfelelő testtartás és hangfejlődés, míg zavara esetén orális rendellenességek léphetnek fel (ARVEDSON–BRODSKY–LEFTON-GREIF 2020).

A nyelési OMD kialakulásának lehetséges okai továbbá az alulfejlett orofaciális izomzat, a hiányzó izomegyensúly, valamint az arc, az ajkak és a nyelv morfológiai és izomműködést érintő sajátosságai. Utóbbin például az ajkak és a nyelv renyhe izomműködését, alaki és működésbeli aszimmetriáját, a nyelv méretbeli paramétereit (túl nagy vagy túl kicsi nyelv) értjük, valamint a nyelv nyugalmi helyzetét, mely problémás esetben „alsó” helyzet (azaz a nyelv nem a kemény szájpadnak nyomódik, hanem az alsó állkapocsban fekszik) (WIEMER 2017). A fentiekén túl a szájlégzés is lehet oka a funkcionális nyelészavarnak, jelenléte pedig tünetként egyértelműen a nyelésfunkció szabálytalanságára utal.

A dentális terület deformitása oka és következménye/tünete is lehet a nyelési OMD-nek. Utaltunk már rá, hogy LOMD esetén a nyelv nyelés közben a szokásosnál nagyobb nyomást fejt ki a fogazatra, ebből következően a nyelésfunkció szabálytalansága fogazatmódosító hatású lehet (FÁBIÁN–GÁBRIS–TARJÁN 2013). Ugyanakkor azt is fontos tudni, hogy a nyelési OMD vagy az LOMD nem szükségszerűen jár együtt a szabályostól eltérő fogazattal. Mindemellett pedig a nyelési OMD és az LOMD megjelenhet fogazati szabálytalanságok vagy változások (pl. fog eltávolítása) miatt is, a megváltozott szerkezethez történő adaptációként.

### 2.2.3. A nyelési OMD és az LOMD osztályozási lehetőségei

Ebben a fejezetben a nyelési OMD néhány lehetséges csoportosítási lehetőségét mutatjuk be a teljesség igénye nélkül, különös tekintettel a magyar logopédiai gyakorlatban leginkább ismert és használt besorolásokra.

A hazánkban legelterjedtebb klasszifikáció szerint nyolc csoportba sorolhatók az egyes szabálytalan nyelésmintázatok, aszerint, hogy a nyelv az üregrendszer mely részeihez nyomódik nyelés során (BARRETT–HANSON 1974; VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015). A típusok a következők: 1. mandibuláris nyelés, mely során a nyelv az alsó állcsonthoz nyomódik; 2. bi-maxilláris nyelés, mely során a nyelv mindkét állcsonthoz hozzápréselődik; 3. mélyharapáshoz

kapcsolódó nyelés; 4. nyitott harapáshoz kapcsolódó nyelés; 5. elülső nyelvlökéses nyelés, mely során a nyelv a frontfogak közé ékelődik a felső metszőfogakat előre, az alsókat hátrafelé nyomva; 6. unilaterális (egyoldali) nyelés, mely egyoldali nyitott harapással jár, és amely során a nyelv valamely oldalra (aszimmetrikusan) mozdul el, és az oldalsó fogak között türemkedik ki nyelés közben; 7. bilaterális (kétoldali) nyelés, amely kétoldali nyitott harapással jár, tehát a nyelvhegy nekinyomódik az alsó vagy a felső fogsor lingvális (belső) oldalának, miközben a nyelvtest kitüremkedik a fogsor két oldalán; és végül 8. teljes nyelvlökéses nyelés, melynek esetében a nyelés során a nyelv és az alsó ajak a felső frontfogak belső, szájüregi oldalához préselődik. Ez a sorban utolsó kategória más felosztásokban egy, a nyelvlökéses nyeléstől eltérő típusként jelenik meg: BJÖRK (1960 és ennek nyomán VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015) ugyanis például ezt ajaklökéses nyelésnek nevezi, és azt mondja, ez a nyelvlökéses nyelés mellett lehetséges második szabálytalannyelés-kategória.

Amint arra már utaltunk, a nyelési OMD-knek és kifejezetten az LOMD-nek fogazatmódosító hatása lehet, illetve együtt járhat fogazati elváltozásokkal. Ennek mentén két nagy csoport lehetséges: a **neutrális LOMD**, amely valóban pusztán a nyelés zavara, és amelyben nem találunk fogazati eltérést és beszédhangképzési problémákat sem. BRAUER és HOLT (1965) ezt a nyeléstípust **deformitással nem járó nyelvlökéses nyelésnek** (*nondeforming tongue thrust*) nevezi. A másik lehetséges típus pedig értelemszerűen a deformitással járó nyelvlökéses nyelés (*deforming tongue thrust*), amelyet a fenti nevezéktna illeszkedve **deformitással járó LOMD**-nek nevezhetünk. A deformitással járó nyelvlökéses nyelést BRAUER és HOLT (1965) további három típusra osztja aszerint, hogy a fogazatbeli deformitás mely fogakat érinti. Ilyen módon a szerzők írnak a szájüreg elején elhelyezkedő fogakat érintő (anterior), a szájüreg oldalán lévő fogakat érintő (laterális) és a szájüreg elülső és oldalsó oldalán elhelyezkedő fogakat egyaránt érintő (anterior-laterális) csoportokról.

Egyes további, friss (külföldi) szakirodalmi források az LOMD (mint *tongue thrust swallow*) esetében a következő nyeléstípusokról beszélnek: anterior (elülső) LOMD-ről, mely során a nyelv előrefelé, a metszőfogak között türemkedik ki a szájüregből nyeléskor; premoláris (a kisörlőknél jelentkező) LOMD-ről, amikor a nyelv a kisörlők között préselődik ki; egyoldali moláris (a nagyörlőket érintő), illetve kétoldali moláris (mindkét oldali nagyörlőket érintő) nyelésről; teljes anterior (elülső) és teljes laterális (oldalirányú) változatról (lásd pl. BILLINGS et al. 2018 szakirodalmi áttekintését). A hazai fogszabályozási irodalomban, magyar nyelven olvashatunk még primérnek tekinthető nyelvlökéses nyelésről (ez az infantilis vagy fiziológiai nyelvlökéses nyelés megfelelője), szeptális nyelvlökéses nyelésről és ajakszorításos nyelésről (FÁBIÁN–GÁBRIS–TARJÁN 2013: 194, 214). A szeptális nyelvlökéses nyelés az állcsontok deformitása miatt fellépő, ún. alkalmazkodó, adaptív nyelés, ebben az esetben tehát a módosított szerkezet hat vissza a funkcióra. Az ajakszorításos nyelés elnevezés itt lényegében ugyanarra utal, mint fentebb is, jelesül, hogy az alsó ajak nyelés közben a felső fogsorhoz préselődik (FÁBIÁN–GÁBRIS–TARJÁN 2013).

A jelen tanulmány szerzői a fenti (illetve egyes további, itt nem közölt) osztályozások egy részének egyfajta szintézisét, illetve ennek kiegészítését alkalmazzák további munkáikban

(HAVADI-NAGY 2020b), mely a nyelési OMD nyelvdomináns formáját, az LOMD-t aszerint osztályozza, hogy a nyelv függőleges irányban milyen mozgásokat végez, illetve miként érintkezik a toldalékcsovéval (hasonlóan HANSON–COHEN 1973 osztályozásához, idézi VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015: 29).

Az általunk bevezetett kategóriák a következők. **Maxilláris** (felső állkapcsi) **LOMD**-nek nevezzük azt, ha a nyelvhegy fognyomásosan (addentálisan) érintkezik a felső fogsor lingvális felszínével. A következő csoportban, a **mediális**, azaz **középirányú LOMD**-ben több alcsoportot is megkülönböztetünk aszerint, hogy a nyelv melyik része és miként vesz részt dominánsan a nyelés végrehajtásában. Mediális-klasszikus LOMD esetén a nyelvhegy középen előre irányuló mozgása jellemző (ez az a nyeléstípus, amelyet egyes felosztások nyitott harapáshoz kapcsolódó nyelvlökéses nyelésnek neveznek, vö. BARRETT–HANSON 1974, idézi VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015). Mediális-teljes LOMD esetén a nyelvtest a fogsorok között szétterülve kifurakszik (ezt BARRETT és HANSON 1974 bimaxilláris protrakcióhoz, mélyharapáshoz kapcsolódó nyelvlökéses nyelésnek nevezi; vö. még VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015). Mediális-laterális LOMD-ként hivatkozunk arra, ha a nyelvtest koronális (elülső) része nyelés közben egy- vagy kétoldali irányban oldalra mozdul (vö. egy- vagy kétoldali nyelvlökéses nyelés, BARRETT–HANSON 1974; VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015). Végül elkülönítjük az alsó állcsontra fognyomást gyakorló nyeléstípust **mandibuláris LOMD** néven. Ebben a típusban a nyelvhegy az alsó fogsor lingvális felszínével érintkezik fognyomással, azaz addentálisan (vö. mandibuláris nyelvlökéses nyelés, BARRETT–HANSON 1974; VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015).

Utaltunk már rá, hogy a nem a nyelvnek, hanem más arci-száji működéseknek a szabálytalanságát mutató nyelési OMD-k között említhető még a valamivel alaposabban dokumentált ajaklökéses nyelés (BJÖRK 1960, idézi VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015), vagy más néven az ún. ajakszorítással képzett nyelés (FÁBIÁN–GÁBRIS–TARJÁN 2013). Emellett azonban a logopédiai praxisban találkozni még ajakbeszívással, illetve arcelfúvással képzett nyeléssel is, ezek tekintetében azonban igencsak hiányos a szakirodalom.

#### 2.2.4. A nyelési OMD és az LOMD vizsgálata

A szabályos és szabálytalan nyelés vizsgálata, elkülönítése elsődlegesen szubjektívebbnek tekinthető technikákon, vizuális és auditív megfigyelésen alapul. Szabálytalan nyelés esetén a logopédus szakember felméri a beszédszervek állapotát, valamint működését. Így megvizsgálja a nyelv és az ajkak morfológiáját, helyzetét és mozgását, az orofaciális terület állcsonti részét, beleértve ebbe a harapást is. A diagnosztika része ezen felül a beszédprodukción és a beszédmotoros kontroll működésének vizsgálata, illetve annak ez elemzése, hogy az orálmotoros funkciók hogyan és miként adaptálódtak a megváltozott szerkezethez (a fogazathoz) (ha van ilyen) (PATIL–SANJEEV–PATIL 2017). Más szerzők a funkcionális működések vizsgálatába a légzést, a nyelésfunkciót, illetve a nyelv- és az ajakfunkció felmérését (is) beleértik (FÁBIÁN–GÁBRIS–TARJÁN 2013). Az itt felsoroltakkal kapcsolatban a szakember a következőket figyeli meg.

Nyelési OMD-re utaló tünet lehet az, ha nyelés közben az ajakzár képzését fokozott izom-összehúzódás kíséri a száj körüli területen. Ezt a logopédus a kliens megérkezésének pillanatától kezdve folyamatosan figyeli a spontán nyelés(ek) során, majd a vizsgálat során fókuszáltan is elemzi (ekkor ő kéri üres nyelésre a páciens). Szabályos nyelés esetén az ajkak körüli izmok feszítettsége alacsony, míg szabálytalan nyelés esetén nagyobb, amely megnövekedett feszítettség arra utal, hogy a nyeléshez szükséges vákuum képzése nem a nyelv és a száypad, hanem előrébb, az orcák területén zajlik. Az ajkak terén emellett ajkaszorítás, ajaklökés és ajkzáródási rendellenesség is felmerülhet, ezeket kell a diagnosztának kiemelten figyelnie.

A legpontosabb információt a nyelv szabálytalan működéséről a nyelésben az adja, ha a páciens a vizsgálat során a logopédus szakember megkéri arra, hogy nyitott vagy félig nyitott ajakkal nyeljen – ez esetben ugyanis jobban megfigyelhető a nyelv mozgása. Az ajkak széthúzását és a nyitott ajakkal nyelést segíti az ún. ajkaszorítókampók használata is, melyek a páciens aktív közreműködése (értsd: száj körüli izmainak megfeszítése) nélkül fedik fel a fogsorokat (1. ábra). Ismeretes azonban, hogy az ajkak és a nyelv működése összefüggésben áll a nyelés során, így az ajkak nyitása (azaz a nyitott vagy szétfeszített ajkakkal történő nyelés) a nyelvműködést is befolyásolhatja (PENG et al. 2004). Az ilyen módon lezajló nyelés tehát nem ad teljesen megbízható képet a kliens nyelési folyamatainak általános (a mindennapokban, nem vizsgálati helyzetben jellemző) működéséről.



1. ábra

*Az ajkaszorítókampók használata (a képeket készítette: dr. Heckenast Lili fogszabályozó szakorvos)*

A kórtörténet szempontjából lényegesek a már fentebb is említett ún. parafunkciók, melyek minden olyan inadekvát orálmotoros tevékenységet magukban foglalnak, melyek az orofaciális területen jelennek meg; ilyenek például a leggyakoribbnak tekintett ujjszopás vagy cumizás, vagy akár a körömrágás is. Ezek a jelenségek ugyanis gátló tényezők lehetnek abban, hogy a fiziológiai nyelési OMD az elvárt módon 4-4,5 éves korra átalakuljon érett, szabályos nyeléssé. A parafunkciók tünettana változatos lehet, egy tényező azonban közös bennük: a nyelv ezekben elülső helyzetet vesz fel (HAVADI-NAGY 2020b).

A nyelésfunkció vizsgálatának legnagyobb kihívást jelentő része a nyugalmi nyelvhelyzet vizsgálata, hiszen a vizuális megfigyelés itt nem lehetséges, ahogyan a kliens kikérdezése sem vezet megbízható eredményre. Éppen ezért ennek elemzésére egyes gyakorlatokban felvetődik a nyelvhát felszínének kontrasztanyagossal megfestése, mely a röntgenfelvételen megfigyelhetővé teszi azt (FÁBIÁN–GÁBRIS–TARJÁN 2013).

PATIL és munkatársai (2017) a nyelv, valamint az ajkak morfológiáját, pozícióját és mozgását, valamint az orofaciális szerkezetet és a harapást vizsgálják, továbbá olyan tényezőket, melyek a nyelv elülső állását kényszeríthetik ki (pl. túlfejtett garatmandulák). A nyelési OMD ugyanis felléphet mint adaptációs és kompenzációs megoldás ezek ellensúlyozására. Ezek a szerzők a beszédprodukción, valamint a motoros beszédkontroll vizsgálatát is beemelik a nyelési OMD felmérésébe. Ehhez hasonlóan a nyelvi és beszédzavarok diagnosztikájának és kezelésének gyakorlati útmutatója (ASHA 2004, vö. BILLINGS et al. 2018) szerint vizsgálandó a faciális-dentális szerkezet (tehát a csontozat), a nyelésfunkció és a beszédprodukción is.

BILLINGS és munkatársai (2018) számos szakirodalmi tételt áttekintő, összefoglaló munkájukban a nyelésfunkción vizsgálatában említik a nyugalmi nyelvpozíció meghatározását, a nyelv külső jellemzőit, relatív méretét, a nyelvfék kérdését, valamint a kemény szájjpad felé irányuló nyelvemelés működését. Emellett a diagnosztika részének tekintik az ún. artikulációs teszteket, azaz a beszédhangok ejtésének vizsgálatát.

A beszédészervi állapot és beszédészervi működés vizsgálatáról bővebben tájékozódhat az olvasó például JUHÁSZ (2007), illetve HAVADI-NAGY (2020a, 2020b) munkáiból, a nyelésfunkción vizsgálatának módszertanáról pedig például VASSNÉ KOVÁCS és munkatársainak (2015), illetve HAVADI-NAGY (2018) tanulmányait javasoljuk.

A nyelésfunkción irányuló objektív, műszeres vizsgálatok száma jelenleg még jóval alacsonyabb, mint a logopédiai vizsgálóeljárásokkal végzettké, minden nyelvetterületen, miközben ezeknek a vizsgálatoknak egy része nem is felel meg (vagy legalábbis nem maradéktalanul) a tudományosság kritériumainak (vö. ALLEN et al. 2020). A nyelés számszerűsített adatokkal operáló elemzésére pedig valóban csak elenyésző számban találunk példát (vö. pl. OHKUBO és SCOBIE 2018-ra datálható megvalósíthatósági tanulmányát nyelvultrahang-készülékkel, mely az ultrahang mint vizsgálati technika újszerűségét példázza a nyelést érintő tanulmányokban).

Kifejezetten a nyelési OMD-re irányul PENG és munkatársainak (2004) eszközös vizsgálata, akik ultrahangkészülék segítségével elemezték a nyelés részfázisainak időarányát és a nyelvmozgás tartományát 40 szabályos, érett nyelést és 15 nyelvlökéses nyelést (*tongue thrust swallow*) mutató, 21,8(±6,1) év átlagéletkorú kísérleti személy összevetésében. A kísérlet talán legfőbb tanulsága az, hogy az ott vizsgált paraméterek tekintetében nem a teljes nyelési folyamatra kapott értékek mutatnak eltérést a szabályos és szabálytalan nyelők csoportjai között, hanem az egyes részfázisok paraméterei. Az eredmények szerint a nyelvmozgás tartományában sem általánosságban, sem az egyes fázisokat illetően nem voltak láthatók különbségek. Ugyanakkor az érett és a nyelvlökéses nyelés eltért a szerzők által „kései transzportnak” nevezett (az itt elkülönített ötből a harmadik) részfázis időtartamát tekintve úgy, hogy nyelvlökéses nyelésben ennek a fázisnak az időtartamaránya (a teljes nyelési folyamatra vetítve) hosszabbnak bizonyult. Ezzel együtt ugyanezen fázisban a nyelvmozgás sebessége is magasabb volt érett nyelésben, mint nyelvlökéses

nyelésben, míg az ezt követő (utolsó előtti) „korai záró szakaszban” éppen fordítva, az érett nyelők mutattak lassabb nyelvmozgást. Mindennek következtében végeredményben a nyelési ciklus egészének időtartama nem tért el a két csoport között, hiszen az időben korábbra eső szakasz lassúságát kompenzálta az azt követő szakasz felgyorsulása a nyelvlökéses nyelők csoportjában. Végezetül a kutatás azt is feltárta, hogy a nyelvlökéses nyelést mutató résztvevők nagyobb (csoporton belüli) változatosságot mutattak a nyelvmozgás tartományában és a nyelv gyorsulásában a „korai záró szakaszban”, miközben a „kései transzport” során éppen ez a csoport volt egységesebb a nyelvmozgás sebességét illetően, nem pedig a szabályos nyelők. Ebből az eredményből pedig arra következtettek a szerzők, hogy a nyelvlökéssel nyelők csoportjában a nyelés összidejét kompenzáló, felgyorsult „korai záró szakasz” kevésbé reprodukálható (változatosabban valósul meg), mint akkor, ha ez a fázis lassabb (ahogyan az az érett nyelőkénél volt megfigyelhető).

Magyar nyelvterületen elsőként a jelen tanulmány első szerzője végzett eszközös elemzéseket a közelmúltban esettanulmány-jelleggel (kis létszámú magyar anyanyelvű személy bevonásával). Az említett tanulmányban a nyelésfunkció működését a palatográfia, azaz a nyelvfestés módszerével vizsgálta elsősorban arra fókuszálva, hogy a nyelv mely szájüregi régiókkal lép kontaktusba nyelés során szabályos nyelést és nyelési OMD-t mutató kliensek esetében (HAVADI-NAGY 2020a, 2020b). A palatográfia segítségével a szervek között a nyelés vagy beszédhang-ejtés közben fellépő kontaktus figyelhető meg: a nyelv megfestése után nyelésre vagy egy adott beszédhang ejtésére kérjük a páciens, így a festék lenyomatot hagy a szájpardon. A palatográfia tehát hozzásegít ahhoz, hogy a szabad szemmel nem (vagy nem jól) látható funkciókat (nyelést vagy beszédhang-ejtést) megfigyeljük. A módszerről, valamint az említett vizsgálat eredményeiről az utolsó fejezetben írunk részletesebben.

Megjegyezzük, hogy a fogazat és a nyelés összefüggései miatt a nyelésfunkcióra irányuló diagnosztika csak részben logopédiai, másik részben azonban fogszabályozási feladatkör. A fogszabályozási szakvizsgálat feltárja, hogy a fennálló fogazati eltérések (pl. nem megfelelő fogi pozíciók és/vagy harapási sík) hátterében (vagy ezek mellett) feltételezhető-e atipikus nyelés, vagy sem (FÁBIÁN–GÁBRIS–TARJÁN 2013). A részletező, a szupraglottális beszédszervek állapotára, valamint működésére vonatkozó vizsgálatot a logopédus szakember végzi el, aki a diagnosztika részleteinek elemző értékelése alapján építi fel a nyeléskorrekcióra irányuló logopédiai terápiát.

### 3. A nyelési OMD és a beszédhang-ejtés összefüggései

#### 3.1. A beszédhangok szabályos és szabálytalan ejtése, illetve ezek összefüggései a nyeléssel – a logopédiai szakirodalom szerint

BLOOMER (1971, idézi VASSNÉ KOVÁCS 1998) a csontozatnak (másként a szerkezetnek) és a funkciónak (azaz a motoros működéseknek) az alakulását bemutató ábrája alapján megfogalmazható, hogy a szerző szerint az orofaciális vázszerkezet és a légnyelési funkciók együttes szabályos működése



eredményezi az „ép beszédet”. Ha a struktúra (elsősorban a fogak helyzete) szabálytalanságot mutat, akkor is létrejöhet „ép beszéd” abban az esetben, ha a beszédmozgások adaptálódnak. Amennyiben azonban a funkció is szabálytalan, az egyértelműen „hibás beszéddel” jár együtt. Beszédhanghiba mindemellett azonban szabályos szerkezet esetén is előfordulhat akkor, ha a beszédmozgások (tehát a funkciók) szabálytalanok – erre utal a szakirodalomban a funkcionális ok (etiológia) kifejezés.

A beszédhangok ejtésében a logopédiai szakirodalom artikulációs zavarnak vagy diszláliának (KOVÁCS–REHÁK 2013: 41), illetve **beszédhanghibának** (TAR 2017: 67) nevezi a beszédhangejtés eltérését a normának tekintett mintától (azaz a beszédhangok torzított ejtését, cseréjét vagy kihagyását). Amennyiben a beszédhanghiba szervi (organikus) okokra vezethető vissza, diszGLOSSZIÁNAK (LUCHSINGER–ARNOLD 1965, idézi KOVÁCS–REHÁK 2013, VASSNÉ KOVÁCS 1998), vagy pedig **organikus eredetű beszédhanghibának** (TAR 2017: 68, 104) nevezik. Utaltunk már rá, hogy a szabályostól eltérő nyelvmozgásnak fogazattorzító hatása lehet, amely tehát beszédhanghibával is járhat. Az ilyen jellegű és eredetű beszédhangejtési zavart a logopédiában a fentebb bevezetett fogalmak segítségével dentális diszGLOSSZIÁNAK (vö. pl. VASSNÉ KOVÁCS 1998: 101; KOVÁCS–REHÁK 2013: 42) vagy dentális eredetű beszédhanghibának is nevezhetjük. Ha a beszédhangok ejtése nem szervi, hanem funkcionális okokkal magyarázható (beleértve ebbe a kogníciót, illetve az érzékelő-mozgató – szenzomotoros – rendszer működését is), akkor annak neve funkcionális diszlália (KOVÁCS–REHÁK 2013: 41) vagy **ismeretlen eredetű beszédhanghiba** (TAR 2017: 68) lehet.

Nyelési OMD esetén, hogyha van fogazati vagy más szerkezeti eltérés, akkor a beszédhangok nem szabályos ejtésére organikus eredetű beszédhanghibaként vagy diszGLOSSZIÁKÉNT utalhatunk; ha azonban nem áll fenn szerkezeti eltérés, akkor funkcionális diszláliaként vagy ismeretlen eredetű beszédhanghibaként. Az elsősorban a nyelv szabálytalan működését mutató LOMD esetén általában a nyelvtest elülső részével képzett mássalhangzók lehetnek szabálytalan ejtésűek, azaz az alveoláris, posztalveoláris és palatális zár-, rés- és zár-rés hangok, illetve egyes szonoránsok.

Utaltunk már rá, hogy a diszlália vagy beszédhanghiba azonosításához a logopédia feltetelez egy ún. „artikulációs normát”, azaz a beszédhangoknak egy egységes és ideális megvalósítási mintázatát, amelyről azt gondolja, hogy a magyar nyelv köznyelvi (vagy másként sztenderd) variánsának ejtészváltozata (vö. KOVÁCS–REHÁK 2013: 41). A nyelési OMD-ben magyar nyelvű beszélők esetében jelentkező beszédhanghibák bemutatásához először tehát át kell tekintenünk az itt releváns (értsd: a szakirodalom és a logopédiai gyakorlat alapján nyelési OMD esetén beszédhanghibákkal megvalósulható) magyar köznyelvi/sztenderd beszédhangok szabályos ejtészmintázatait (artikulációs konfigurációját), mely a logopédia számára a viszonyítási alapot, a „normát” jelenti. Az egyes beszédhangok létrehozásához szükséges képzőszervi konfigurációk bemutatását lentebb a logopédiai szakirodalomból idézzük (THOROCZKAY 2016), hiszen ez képezi a logopédiai vizsgálatok alapját, ám ehhez a fonetikai szakirodalomban megtalálható tudományos szempontú leírásokat is társítunk (BOLLA 1995; GÓSY 2004). Mivel az atipikus nyelés esetén jelentkező szabálytalan beszédhangképzés elsősorban a beszédhangok

képzési helyét érinti, így a jelen tanulmányban mi is csak a hangok képzési helyére fókuszálunk. (Megjegyezzük azonban, hogy a képzési mód változásainak kutatása is releváns iránya lenne a szabálytalan nyelés és beszédhangejtés vizsgálatának, melyről a jelenleg rendelkezésre álló logopédiai és kísérletes szakirodalom lényegében egyáltalán nem értekezik, legalábbis a magyar beszédet illetően.) A szakirodalom és a logopédiai tapasztalat alapján a kifejezetten LOMD mellett megjelenő beszédhanghibák általában az alveoláris [t d n l s z ts], a posztalveoláris [ʃ ʒ ʝ] és ritkábban a palatális képzésű [c ɟ ɲ] hangokat érinthetik (az elnevezések és a csoportosítás itt GÓSY 2004 leírásait tükrözik). Bár képzési sajátosságait (a képzés helyét és módját) figyelembe véve a [dʒ] és [dʒ] mint a [ts] és [tʃ] zöngés megfelelői is ezeknek a hangoknak a körébe tartoz(hat)nának, a [dʒ] és [dʒ] beszédhangokról és ezek lehetségesen hibás ejtéséről a logopédia szakirodalma egyáltalán nem tesz említést. (Ez feltehetően azért van így, mert ezek a beszédhangok nagyon ritkák, illetve esetleg azért, mert ezeket a logopédia nem önálló beszédhangoknak, hanem hangkapcsolatoknak tartja, szemben a magyar fonetikai irodalom zömével, vö. pl. BOLLA 1996; GÓSY 2004). Éppen ezért ezeket a hangokat mi sem említjük alább. Ugyanakkor felhívjuk a figyelmet arra, hogy a fenti okból a [dʒ] és [dʒ] beszédhangoknak a beszédhanghibás ejtéséről egyelőre nincsenek ismereteink, sem LOMD-ben, sem más ok mellett, tehát ezeknek a vizsgálata mind logopédiai vizsgálóeljárásokkal, mind műszeres elemzésekkel is figyelmet érdemel a jövőben.

Az LOMD mellett megjelenhető beszédhanghibákban általában értintettnek tartott hangok képzési konfigurációja a részt vevő nyelvi és szájpadi régiók szerint a következőkben adható meg:

- **[t]: apiko-dentialveoláris mássalhangzó** (a nyelvheggyel a fogak és fogmeder területén képezzük) (BOLLA 1996: 166), vagy máshol fogmederhang, azaz alveoláris (GÓSY 2004: 79), melyben a logopédiai szakirodalom szerint a nyelv pereme képezi a zárt a felső fogmedren, valamint kétoldalt, a felső fogsor mentén (THOROCZKAY 2016: 154–420);
- **[d n l r]: apiko-alveoláris mássalhangzók** (a nyelvheggyel a fogmeder területén képezzük) – itt a képzés helyét illetően lényegében egyeznek a fonetikai (vö. BOLLA 1996: 166; GÓSY 2004: 79) és a logopédiai (vö. THOROCZKAY 2016: 154–420) leírások;
- **[s z ts]: predorzo-dentialveoláris mássalhangzók** (a nyelvhat elülső részével a fogak és a fogmeder területén képezzük) (BOLLA 1996: 166), vagy máshol fogmederhangok, azaz alveolárisok (GÓSY 2004: 79), melyekben a logopédiai szakirodalom szerint a nyelv hegye érinti az alsó fogmedret (tehát nem [csak] a felsőt, mint a fonetikai leírások szerint), a nyelvhat feldomborodik, és a nyelv oldalpereme zárszerűen simul a kemény szájpád széléhez és a felső fogsorhoz (THOROCZKAY 2016: 154–420);
- **[ʃ ʒ]: apikopredorzo-alveoprepalatális mássalhangzók** (a nyelvheggyel és a nyelvhat elülső részével a fogmedernél és a szájpád elülső részén képezzük) (BOLLA 1996: 166), vagy máshol hátulsó fogmederhangok, azaz posztalveolárisok (GÓSY 2004: 79), melyekben a logopédiai szakirodalom szerint a nyelv elülső pereme képezi az akadályt a fogmeder mögötti részen, és a nyelv pereme rásimul a kemény szájpád szélére és a „záfogakra” (THOROCZKAY 2016: 154–420);

- **[ʃ]: apiko-alveoprepatális mássalhangzó** (a nyelvheggyel a fogmeder és a szájpad elülső részén képezzük) (BOLLA 1996: 166), vagy máshol hátulsó fogmederhang, azaz posztalveoláris (GÓSY 2004: 79), melyben a logopédiai szakirodalom szerint a nyelv elülső pereme képezi az akadályt a fogmeder mögötti részen, és a nyelv pereme rásimul a kemény szájpad szélére és a „zápfogakra” (THOROCZKAY 2016: 154–420);
- **[c j ɲ]: dorzo-palatális mássalhangzók** (a nyelvháttal a kemény szájpad területén képezzük) (BOLLA 1996: 166), vagy másként elülsőszájpad-hangok, azaz palatálisok (GÓSY 2004: 79), melyben a logopédiai szakirodalom szerint a nyelv hegye felfelé néz, és a nyelv-hát középső része képezi az akadályt a szájpadlás elülső részével, miközben a nyelv oldalpereme zárszerűen simul a kemény szájpad széléhez és a „zápfogakhoz” (THOROCZKAY 2016: 154–420).

Vegyük észre, hogy a képzés egyes mozzanatait illetően találunk eltéréseket a fonetikai és logopédiai leírások, valamint az egyes fonetikai források között is. Az előbbire példa az alsó fogmeder érintése mint jegy a [s z ts] hangok esetében, melyről a beszédhangok tudományos szempontú, fonetikai megfigyelése nyomán nem számolnak be a szakirodalmi források, ám a logopédiában megjelenik, illetve ugyanitt a felső fogmedren képzett akadály, melyről pedig a logopédiai szöveg nem tesz említést, miközben a fonetika tudományában ezt meghatározó jegyként, a képzés helyeként tartjuk számon ezeknek a hangzóknak az esetében. Utóbbira ad példát az alveoláris és dentialveoláris hangzók megkülönböztetése az egyik (vö. BOLLA 1995), és meg nem különböztetése a másik (vö. GÓSY 2004) szövegben, továbbá az [ʃ z ʃ] (rendre) alveoloprepatálisnak, illetve posztalveolárisnak minősítése.

A logopédiai és fonetikai leírások közti ellentmondásoknak nyilvánvalóan következményei vannak a beszédhangok szabályostól eltérő leírását illetően is a fonetikai és logopédiai irodalomban, hiszen mindez azt eredményez(het)i, hogy ahogyan a szabályos képzés leírása, tehát a viszonyítás alapja eltér a két megközelítésben, a szabálytalan ejtésben is más minősítéseket találunk. Amint azt látni fogjuk azonban, az LOMD esetén leírt beszédhanghibák alább tárgyalt jellegzetességei lényegében nem esnek ebbe az ellentmondásos tulajdonság csoportba, így a logopédiai és fonetikai leírások ellentmondásainak kérdésével a továbbiakban nem foglalkozunk.

A beszédhangok különböző fonetikai leírásai közötti ellentmondások pedig azt a kérdést bonyolítják, hogy az objektív leírásokban pontosan mit is kellene kiindulásnak tekintenünk. Így kérdésként merül fel, hogy valódi különbséget jelöl-e a *dentialveoláris* és *alveoláris* megnevezés BOLLA (1995) esetében, és a [t] ejtésében valóban előrébb képződő akadállyal kell-e számolnunk, mint a többi, mindkét forrásban alveolárisnak tekintett beszédhang esetében. Továbbá az is kérdés, hogy a [ʃ z ʃ] esetében a posztalveoláris vagy az alveoloprepatális képzési hely számítható-e referenciának, ugyanakkor az sem világos, hogy ezek egyáltalán eltérnek-e egymástól, és különböző régiókat jelölnek-e a szájüregben. A jelen tanulmányban abból a feltételezésből indulunk ki, hogy a nem átfedésben lévő elnevezések által jelölt jelenségek lényegileg ugyanazok, melyekre a szerzők egyszerűen más terminusokat használnak (vö. alveoloprepatális vagy posztalveoláris), illetve abból, hogy az említett képzéshelyi megkülönböztetés (dentialveoláris vagy alveoláris) szintén csak nevezéktani eltérés, mely nem egy valós fonetikai szembenállást

tükröz. Ezekben az esetekben a szélesebb körben elterjedt terminusokkal (posztalveoláris, illetve alveoláris) utalunk a beszédhangokra a továbbiakban.

Az atipikus nyelés esetén jelentkező szabálytalan beszédhangképzés tehát, mint írtuk, elsősorban a beszédhangok képzési helyét érinti, különösen, ha LOMD-ről van szó. Itt a logopédiai praxisból származó megfigyelések szerint, összefüggésben a nyelés közben megfigyelt működésével, a nyelv a beszédhangok ejtésében is elülsőbb helyzetet vesz fel, tehát a beszédhangok képzési helye is „előretolódik”.

A „nem normatívan” megvalósuló képzési hely lehet interdentális, addentális vagy laterális. **Interdentális ejtés** során a nyelv a két fogsor között préselődik ki (KOVÁCS–REHÁK 2013: 42). **Addentális** (azaz fognyomásos) **ejtés**kor a nyelv a fog belső oldalához nyomódik (KOVÁCS–REHÁK 2013: 42; FRINT 1982 nyomán). **Laterális** (azaz egy- vagy kétoldali oldal-) **képzésben** pedig a nyelv kitér a középvonaltól, és oldalirányú mozgással hozza létre az akadályt (DAMICO–BALL 2019), illetve a levegő egy vagy két oldalon a kisórló fogak irányába távozik (KOVÁCS–REHÁK 2013: 42; FRINT 1982 nyomán). Szintén megjelenhet a képzési hely palatálissá válása egyébként nem palatális képzésű hangokban, a nazális (orrhangú) ejtés egyébként nem nazális ejtésű hangokban, valamint a dorzális képzés egyébként nem dorzális képzésű hangokban (VASSNÉ KOVÁCS 1998).

A képzési hely interdentálissá válása a [s z ts ʃ ʒ ʧ l] (VASSNÉ KOVÁCS 1998) vagy más források szerint a [t], a [d n l r], a [s z ts], a [ʃ ʒ ʧ], valamint a [c ɟ n] hangokra jellemző (VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015) (a hangok csoportosítását a normatív variánsuk ejtési hasonlósága adja, lásd fentebb). Addentálissá váló képzési helyről a [t] felpattanó zárhang és [d n] felpattanó és felnyíló zárhangok, a [s z ts] és [ʃ ʒ ʧ] réshangok és zár-rés hangok (VASSNÉ KOVÁCS 1998), illetve más források szerint a [t] és a [d n l] esetében (VASSNÉ KOVÁCS–REHÁK–VIZKELETY 2015) lehet szó. Laterális (oldalirányú) ejtéssel a [s z ts] és [ʃ ʒ ʧ] hangok, illetve a [c ɟ] és a [t] és [d] jelenhetnek meg. Szintén a [s z ts] és [ʃ ʒ ʧ] megjelenhet palatálisan (a kemény szájpad területén) képezve, továbbá nazális (orrhangú) változatban is. Egyoldali laterális ejtés figyelhető meg a [l] esetében, melyre egyébként dorzális és palatálissá váló produkció is jellemző lehet az LOMD-t kísérő beszédhanghibaként. A dorzális ejtés a [t] és [d] esetében is megjelenhet (VASSNÉ KOVÁCS 1998). Végül torzított ejtést eredményezhet az is, ha a képzési hely alveoláris vagy posztalveoláris helyett (vö. GÓSY 2004) palatális ejtésűvé válik. Ebben az esetben a nyelv a „normatív” megvalósításban tapasztalhatóán hátrébb helyezkedik el, ami miatt a beszédhang ejtése tompább hanghatásúvá válhat (VASSNÉ KOVÁCS 1998: 116).

A szabálytalan nyelésnek és a beszédhanghibáknak a fentiekben ismertetett összefüggései, mint írtuk, a logopédiai gyakorlatból, így elsősorban gyakorlott logopédusok hallási és vizuális alapú vizsgálóeljárásának tapasztalataiból származnak – a magyar beszédhangok szabálytalan ejtésének viszonylatában ugyanis szinte kizárólag ilyen adatokkal rendelkezünk. A beszédhang-ejtés megfigyelése azonban szabad szemmel (még ha kiegészítésként a hallásunkra is hagyatkozunk abban) nehézségekbe ütközik, hiszen a beszédhangok artikulációja legnagyobb részt a szájüregi zárt téren belül zajlik, miközben hallási észlelésünk megbízhatósága is kérdéses az artikulációs mozgások megbecslését illetően az akusztikai szerkezet alapján. A fonetika tudománya az artikulációnak ezt a rejtetten zajló részét eszközökkel, műszerekkel vizsgálja, amire

korábban a röntgenográfia és a palatográfia (illetve lingvográfia) eljárásait használta, ma pedig már a nyelvultrahang, az elektromágneses artikulográfia vagy az elektropalatográfia módszerei is rendelkezésre állnak (bár csak meglehetősen anyagi befektetések árán) (vö. pl. DEME et al. 2019; GRÁCZI et al. 2020; MARKÓ et al. 2020). Ezeknek az eszközöknek a használata mind a logopédiai diagnózis, mind pedig a terápia számára sok fontos hozadékkal szolgálhat(na) a jövőben.

### 3.2. A beszédhangok szabályos és szabálytalan ejtése, illetve ezek összefüggései a nyeléssel – a kísérleti adatok tükrében

Ahogy fentebb már említettük, a szabályos és szabálytalan nyelés műszeres vizsgálatára más nyelvterületen – ha nem is kifejezetten nyelési OMD esetén, de – már vannak példák (lásd pl. ALLEN et al. 2020 összefoglalását; továbbá pl. OHKUBO–SCOBIE 2018), ahogyan a beszédhanghibák feltárására (sőt akár kezelésére) is léteznek már kísérletek, például a nyelvultrahang módszerével (CLELAND–SCOBIE–WRENCH 2015). És bár a kérdésre kevesebb irodalmat találni, de a nyelvhelyzet és nyelvműködés, illetve a beszédhangejtés összefüggéseit is vizsgálták már műszeresen (pl. KRAVANJA et al. 2018). Hazánkban a jelen tanulmány első szerzője igyekszik új eredményekkel gazdagítani a nyelés és az ahhoz kapcsolódó beszédhangejtési problémák objektív, műszeres vizsgálatának szakirodalmát a palatográfia eszközeivel (HAVADI-NAGY 2020b). Tanulmányunk utolsó előtti fejezetében ezeknek a kísérleteknek a módszertanát és eredményeit ismertetjük röviden, majd egy-egy kiragadott példán keresztül ízelítőt adunk abból is, hogy az ott kapott eredmények alapján milyen összefüggések láthatók a nyelés és beszédhangejtés között az eszközös elemzés szerint. Mivel azonban az említett kísérlet módszertanában központi a palatográfia mint kísérleti eljárás, a következőkben elsőként kitérőt teszünk, hogy bemutassuk ezt a technikát.

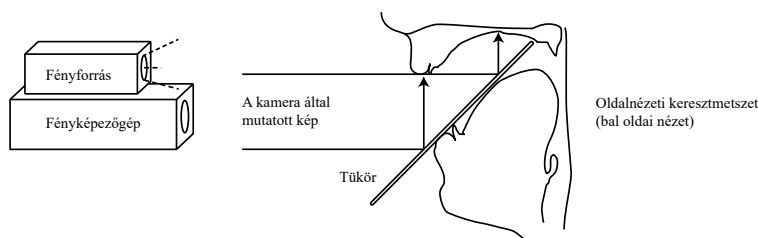
#### 3.2.1. A palatográfia módszere

A **palatográfia** elnevezés a latin *palatum*, 'szájpad' és a görög *γραφειν* ['γρα.φιν] 'írni' kifejezésből ered, és a 'szájpadra rajzolt kép'-re utal, ugyanis a palatográfia módszerével a nyelv és a szájpad (illetve a fogak) kontaktusmintázatai figyelhetők meg a szájpadon például nyelés vagy beszédhangejtés közben. A módszerrel készített regisztrátum neve **palatogram**, amely elsődlegesen a nyelv által a szájpadon hagyott lenyomatra használatos, míg a nyelvnek a fogakkal való érintkezéséből előálló lenyomatot **dentogramnak** nevezhetjük (a latin *dente/dentes* 'fog/'fogak' szóból), a befestékezett szájpadnak a nyelven hagyott lenyomatát pedig **lingvogramnak** (a latin *lingua* 'nyelv' szóból). (Ugyanakkor megjegyezzük, a dentogram kifejezés ritkán használatos, és gyakran a palatogram terminus utal a foga lenyomatokra is.)

A palatogram készítése során az első lépés az, hogy a kutató egyenletesen befesti a kísérleti személy nyelvét valamely ehető, az egészségre nem káros színezőanyaggal, például olívaolaj és szénpor egyenlő arányú keverékével (LADEFOGED 2003; ANDERSON 2008) vagy ételfestékekkel (HAVADI-NAGY 2020b), általában egy vastagabb ecset segítségével. (Fontos megjegyezni, hogy a színező médium összetételéről a kutatonak egyeztetnie kell a kísérleti személyekkel a kísérlet

előtt minden esetben, azért, hogy az esetleges allergiás reakciók lehetősége teljesen kizárható legyen.) Ez után a kutató arra kéri a résztvevőt, hogy (a nyelés vizsgálata esetén) nyeljen egyet, vagy (egy beszédhang vizsgálata esetén) nyelés nélkül ejtse ki a vizsgálandó beszédhangot izoláltan (azaz önmagában, pl. [s]), esetleg hangsorban (azaz egy létező vagy álszóban, pl. [so:] vagy [sa:]), és ezután enyhén hátradöntve a fejét rögtön nyissa ki a száját. A kísérletvezető arra is figyelmezteti a résztvevőt, hogy ne ejtsen ki további beszédhang(ka)t, illetve ne nyeljen (ismét), hogy a nyelv és a szájpad valóban csak egyetlen egyszer (a vizsgált nyelési fázis, illetve a kérdéses beszédhang ejtése során) találkozzon, és így a létrejövő lenyomat is csak ezt az egyetlen érintkezést örökítse meg. Ugyanezen okból például a [s] és [ʃ] ejtésekor jellemző kontaktus vizsgálatára alkalmasak lehetnek a [so:] és [ʃo:] szavak (melyek minimális párok), illetve pl. a [se:] és [ʃe:] álszavak, de nem alkalmasak például a [so:t] és [ʃo:t] szavak, mert ezekben a nyelv közel ugyanazon területe kétszer is érintkezik a szájpadalattal.

A nyelv és a szájpad érintkezése miatt a szájpadon keletkezett lenyomatot (vagy ennek hiányát) a kutató egy kezítükör segítségével figyeli meg, úgy, hogy a tükröt a szájterbe helyezi, illetve erről a lenyomatról azonnal fényképet is készít (lásd 2. ábra). Az adatelemzés során ezeket a fényképeket vizsgáljuk és rendszerezzük (tanácsos azonban az egész folyamatot videóra is felvenni). A folyamat során használt eszközöket, ecsetet, tükröt természetesen minden egyes felvételkészítés előtt alaposan fertőtleníteni kell. (A palatogram készítésének további gyakorlati részleteiről lásd ANDERSON 2008 részletes útmutatóját.)



2. ábra

*A nyelv a szájpadon hagyott lenyomatának lefényképezéséhez használt lehetséges elrendezés sémája: a nyílak azt mutatják, ahogyan (ideális esetben) a kamera számára a szájpad láthatóvá válik (LADEFOGED 2003: 37 és ANDERSON 2008: 5 nyomán)*

Ha a használt színezék viszkozitása megfelelő (tehát nem túl folyós és nem túl ragacsos a festék), akkor a palatogramon elsődlegesen a nyelv lenyomatának mérete és helye figyelhető meg, tehát az, hogy mekkora felületen és hol érintkezett a nyelv a szájpadalattal (vagy a fogakkal) a nyelés vagy beszédhangéjtés közben. Ezek mentén a lenyomatok csoportosíthatók és nominális változósokként elemezhetők (lásd pl. HAVADI-NAGY 2020b), de megfelelő képi minőség esetén számszerű elemzések is lehetségesek (ANDERSON 2008). Bizonyos mértékig a lenyomat színerejtése sejteti a működések jellegét is, mert a halványabb színből nem feszes, hanem inkább lazább érintkezésre vagy esetleg súrlódásra következtethetünk. Mivel azonban a statikus kép a képzés dinamikus módját nem képezi le egyértelműen (illetve mert a színeződés erőssége függ – többek között – a fénykép

minőségétől, illetve a nyelvre felvitt festék mennyiségétől is), a kontaktus jellegére valójában nem következtethetünk megbízhatóan (szemben a kontaktus helyével és az érintkezés felületével). A statikus palatográfiás eljárás modernebb (és költségesebb) változata a dinamikus palatográfiaként is emlegetett elektropalatográfia, mely során a szájba helyezett műanyag műszájpad és az ebben lévő elektródák regisztrálják az érintkezést (bővebben lásd pl. LADEFOGED 2003).

Természetesen információértéke van annak is (tehát adatként kezelendő), ha a szájpadon nem jön létre lenyomat, hiszen ez arra utal, hogy a nyelv nem találkozott a szájpadalattal a nyelés vagy beszédhangejtés során. Megfigyelendő lehet továbbá az alsó fogsorokon keletkezett lenyomat (vagy ennek hiánya) is, hiszen ez arról informál, hogy a nyelv (felső, festékes felszíne) az alsó fogsorhoz nyomódott (vagy a lenyomat hiánya esetén: nem nyomódott a fogsorhoz) az ejtés vagy nyelés közben. Ez különösen fontos lehet például a nyelés vizsgálatokor, hiszen egyes LOMD-típusok esetében éppen abban áll a nyelés szabálytalansága, hogy a nyelv a szájpad helyett a felső és/vagy az alsó fogsorokhoz nyomódik.

Egy-egy palatogram elkészítése viszonylag egyszerű és alacsony költségvetés mellett is kivitelezhető. A használt festékanyag (és a beszélő személy egyéni sajátosságainak) függvényében azonban a vizsgálat folyamata igencsak időigényes lehet, hiszen például az ételfestéket nem lehet azonnal teljes mértékben kiöblíteni a szájüregből. Ezért ha például egy-egy résztvevőtől öt különböző beszédhang ejtését is vizsgálni kívánjuk, akkor legalább öt alkalommal találkozunkunk kell, hogy elkészíthessük az összes regisztrátumot. Ha olyan festékanyagot használunk, ami könnyebben kiöblíthető (ilyen lehet az olaj és szénpor keveréke LADEFOGED 2003 szerint), akkor pedig az egyes ülések ideje hosszabbodik meg (a szájterület öblögetése, illetve az eszközök tisztogatása miatt az egyes tesztingerek ejtése között). Ugyanakkor ebben az esetben sem garantált, hogy minden beszélő nyálkahártyája ugyanannyira könnyen tisztítható (egyes személyek nyálkahártyája jobban, másoké kevésbé szívja magába a színezőanyagot), tehát így is fennáll annak az esélye, hogy a kísérlet csak több ülésben végezhető el.

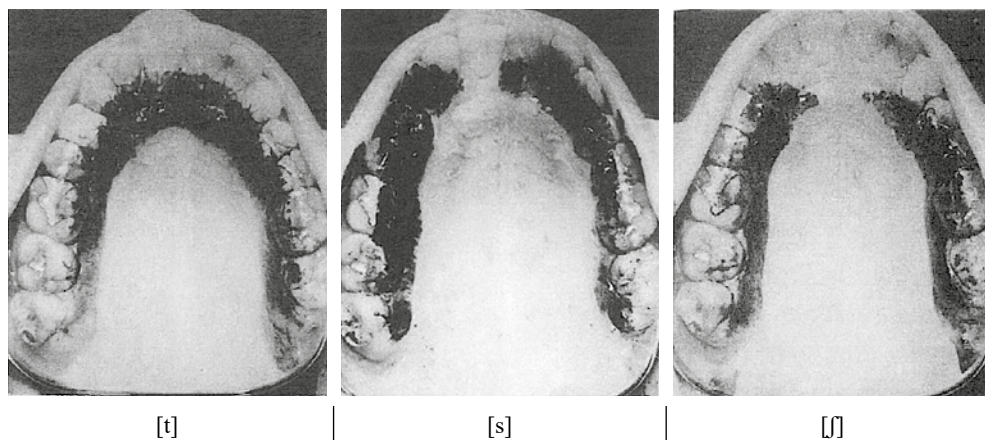
### 3.2.2. Egy palatográfiás vizsgálat tanulságai: a nyelés jellegzetességei és a beszédhanghibák gyakoriságának összefüggései az összes kísérleti személy adatainak tükrében

Ahogy arra már fentebb utaltunk, a nyelés és a beszédhangejtés összefüggéseit magyar nyelvterületen kísérletesen elsőként HAVADI-NAGY Marian (2020b) vizsgálta részben a palatográfia módszerével. Az elemzésben a) a nyelési OMD lingvális változatát (azaz az LOMD-t) és ebben a nyelés orális transzportfázisát, b) a beszédhangejtést, illetve c) ezek gyakoriság szerinti összefüggését igyekezett feltárni viszonylag kis létszámú kísérleti személy bevonásával, eset-tanulmány-jelleggel. Az idézett tanulmányban a nyelés és a beszédhangejtés összefüggéseiről elsősorban a gyakoriság mentén esik szó, a nyelés és ejtés jellegének együtt járásairól nem – az adatok ilyen szempontú elemzése ugyanis még jelenleg is zajlik. A jelen tanulmány zárásaképpen ezekből a részben jelenleg is formálódó eredményekből mutatunk be szemelvényeket abból a célból, hogy a lehetséges összefüggéseket, valamint a választott módszertan alkalmasságát és lehetséges korlátait szemléltessük.

A fentebb idézett kísérletében HAVADI-NAGY (2020b) tehát részben a palatográfia módszerét használta arra, hogy megvizsgálja a nyelés és a beszédhangjejtés során a palatogramokon megfigyelhető kontaktusmintázatokat. Emellett a beszédhangoknak egy nagyobb csoportját a logopédiai gyakorlatban alkalmazott auditív-vizuális megfigyeléssel is elemezte.

Az auditív-vizuális módszerrel végzett adatfelvétel során a szerző logatomok (azaz álszavak) utánmondását kérte öt fonológiai helyzetben: izolált beszédhangjejtésben (pl. [d]), szótagkezdő (pl. [da:], [di:], [du:]), szótagvégi (pl. [a:d], [i:d], [u:d]), intervokális (azaz magánhangzók közötti, pl. [a:da:], [i:di:], [u:du:]) és mássalhangzó-torlódásos (pl. [a:mda:], [i:mdi:], [u:mdu:]) helyzetben. A vizsgált beszédhangok az alveoláris [t], [d], [n], [l], [r], [s], [z], [ʃ], a posztalveoláris [ʒ], [ʒ], [ʒ], valamint a palatális [c], [j] [ɲ] voltak, ezek magánhangzós kontextusa minden esetben az [a: i: u:] volt, mássalhangzós kontextusa pedig a [m]. A beszédhangok ejtéséről videofelvétel készült, az elemzés pedig a logopédiai praxisban alkalmazott, ún. auditív-vizuális módszerrel zajlott. Az auditív-vizuális vizsgálat során a logopédus hallás alapján (és tapasztalataira támaszkodva) állapítja meg a beszédképző szervek állását, illetve a nyelv és a szápad által képzett zár vagy szűkület helyét (a képzés helyét) és jellegét (a képzés módját) úgy, hogy a leírásban a fonetikából ismert képzési jegyek említésén túl szubjektív hangélményi leírásokat is megad(hat).

HAVADI-NAGY (2020b) az alveoláris felpattanó zárhang [t], az alveoláris frikatíva [s] és a posztalveoláris frikatíva [ʒ] mássalhangzókat vizsgálta palatográfiaival, beszélőnként egy-egy izolált ejtésbeli megvalósulás (azaz egy-egy palatogram) alapján. Mivel a magyar logopédiai szakirodalomban nem ismertek palatográfias elemzések, a kísérlet ezen részében a beszédhangok ejtése során kapott lenyomatokat a szerző a BOLLA (1995) által dokumentált ejtési megvalósulások palatogramjaival hasonlította össze vizuálisan (3. ábra), és ez alapján minősítette az ejtést. Az ejtés „normatív” vagy „szabályos” címkét kapott akkor, ha az a BOLLA (1995) által közölt palatogrammal lényegében egyezést mutatott, illetve „nem normatív” vagy „szabálytalan” címkét kapott akkor, ha a palatogramon látható festékeződés a BOLLA (1995) által dokumentált lenyomattól jelentősebben eltért.



3. ábra

A [t], [s] és [ʒ] palatogramjai (BOLLA 1995: 170–173)



HAVADI-NAGY (2020b) a palatográfiás elemzésekhez kék színű ételfestéket használt, mely tartósabb elszíneződést okoz (nem öblíthető ki azonnal teljes mértékben), így a vizsgálat több (összesen négy) ülésben, több elkülönült napon zajlott. Az első alkalommal az üres nyelés során tapasztalható lenyomatokat rögzítette palatográfiával, valamint a beszédhangejtést elemezte a logopédiai praxis auditív-vizuális módszerével. A további alkalmak során pedig egy-egy ülés erejéig az egyes kiemelt beszédhangok palatográfiás vizsgálata zajlott.

A kutatásban tíz kísérleti személy vett részt, melyből öt fő a logopédiai diagnózis szerint valamely nyelési szabálytalanságot mutató tesztszemély volt, öt fő pedig nyelési szabálytalanságot nem mutató, a tesztszemélyekhez nemben, életkorban és testmagasság szerint illesztett kontrollszemély. A kutatásban részt vevőket egy megelőző logopédiai vizsgálat alapján osztotta a szerző szabályos nyelők és LOMD-t mutató szabálytalan nyelők csoportjába. Az átlagéletkor a tesztcsoportban és a kontrollcsoportban is egyaránt 8;2 év (azaz 8 év 2 hónap) volt.

A kísérlet előtt HAVADI-NAGY (2020b) azt a feltételezést fogalmazta meg, hogy a szakirodalomban olvasható leírásokkal egybevégezően a nyelés során létrejövő lenyomat LOMD (azaz a tesztszemélyek) esetében a fogak belső, tehát lingvális (nyelv felőli) felszínén jelenik meg, míg szabályos nyelésfunkció (azaz a kontrollszemélyek) esetében kizárólag a szájpardon. Emellett pedig arra is számított, hogy LOMD-ben nagyobb arányban jelennek meg az auditív-vizuális, illetve a palatográfiás elemzés alapján azonosítható beszédhanghibák, mint szabályos nyelés esetében.

Az eredmények azt mutatták, hogy a nyelés orális transzportfázisában szabályos nyelők esetében valóban jellemző volt a palatumon létrejött (intenzívebb) festékes lenyomat, míg a dentális területen – egy kísérleti személy kivételével – nem volt tapasztalható festékeződés (az egyetlen kivétel esetében a palatumon és a fogakon egyaránt lenyomatot hagyott a megfestett nyelv). Az LOMD-t mutató kísérleti személyek esetében a felső fogsor elülső (lingvális) részén volt tapasztalható festékleNyomat, melyet egyes esetekben a palatum egyes részein (középen és néha oldalt) is megjelenő (halványabb) festékeződés kísért. A szerző feltételezése szerint a halványabb lenyomat arra utal, hogy a nyelv nyelés közben előrenyomult, súrolva a szájpadot – azonban emlékeztetünk arra is, hogy a statikus palatográfiából ilyen jellegű következtetések csak feltételesen vonhatók le, ennél fogva ez a magyarázat is csak feltételezésként értelmezendő.

A beszédhangejtés vizsgálata lényegesen több beszédhanghibát tárt fel a szabálytalan nyelők esetében, mint a szabályos nyelők esetében, tehát az eredmények alátámasztották, hogy a szabályostól eltérő nyelvmozgás funkcionálisan a beszédhangok nem normatív (addentális, interdentális, laterális, illetve az ezekbe a kategóriákba be nem sorolható szabálytalan) ejtését eredményezheti. Ugyanakkor a szerző azt is tapasztalta, hogy szabályos nyelés esetén is előfordulhat – más, a szabálytalan nyeléstől eltérő okra visszavezethető – beszédhanghiba. A kutatás 14 különböző beszédhangra irányult, a beszédhangejtés auditív-vizuális vizsgálatának eredményei alapján az LOMD-t mutató beszélők esetében ezeknek 78,5%-ában volt regisztrálható beszédhanghiba, míg a szabályos nyelők esetében a 14 beszédhang 64,2%-ában. A csoportonként öt beszélő ejtésében előforduló összesen 70 beszédhangot 100%-nak véve a beszédhanghibák előfordulási aránya 40% volt a szabálytalan nyelők és 17,1%-os a szabályos nyelők csoportjában. A beszédhanghibák aránya LOMD esetén 5,6 beszédhanghiba/fő volt, míg szabályos nyelés esetén 2,4 beszédhanghiba/fő.

### 3.2.3. Egy palatográfiás vizsgálat tanulságai: a nyelés és a beszédhanghibák jellege és ezek lehetséges összefüggései három eset leírása alapján

A nyelés és beszédhanghibák jellegének összefüggéseit egy tesztszemély és két kontrollszemély eredményeinek bemutatásán keresztül hasonlítjuk össze. A beszélők pontosabb beazonosítása céljából az idézett cikkben a szerző által használt, a beszélőkre vonatkozó jelöléseket alkalmazzuk (tesztszemély: T5, illetve kontrollszemélyek: K2 és K4).

A 4. ábrán a szabályos nyelés, illetve az LOMD esetén megjelenő festéklenyomatokat mutatjuk be.



4. ábra

*A nyelésfunkció palatográfiás lenyomata: bal oldalon egy szabályos nyelőr (K4) szájpadi festéklenyomatát, a jobb oldalon pedig egy LOMD-t mutató beszélő (T5) szájpadi lenyomatát láthatjuk*

#### 3.2.3.1. Az első szabályos nyelőr

A 8;1 éves kontrollszemély (K4) esetében a felső fogsor előre dőlő, ún. prognátiás szerkezeti eltérést tapasztaltuk, mely a vizsgálat elején felvetette annak a lehetőségét, hogy a nyelésfunkció ennek a személynek az esetében szabálytalanul működik. Ezzel szemben mind a megelőző logopédiai vizsgálat, mind pedig a palatográfia azt támasztotta alá, hogy a beszélő nyelésfunkciója szabályos. A nyelési palatográfia során a szájpád közepén jött létre festékes lenyomat (lásd 4. ábra, bal oldal), míg a dentális terület sem az alsó, sem a felső fogsornál nem festékeződött be. K4 nyelési palatogramja tehát szabályos nyelésre utal, ugyanakkor a beszédhangejtés auditív-vizuális vizsgálata során inkonzekvens unilaterális [l] hangot állapítottunk meg, ami azt jelenti, hogy a beszédhang ejtésekor a beszélő a nyelvét a fogsor bal oldalához nyomta, de nem minden egyes előfordulás esetében.

Az auditív-vizuális és a palatográfiás vizsgálat (5. ábra) ugyancsak megerősítette, hogy K4 ejtésében a [t], a [s] és a [ʃ] beszédhangok képzése egyaránt normatív, vagy másként szabályos.

Ez azt jelenti, hogy (a 4. ábrán illusztrált módon) a beszédhangok kétoldali nyelv-szájpad érintkezéssel jöttek létre úgy, hogy az érintkezés eleje (a szájüreg szájnylás felőli oldalától, azaz a metszőfogaktól nézve) a [t] és [s] esetében az alveoláris (fogmedri) területen, a [ʃ] esetében pedig a posztalveoláris (hátsó-fogmedri) területen volt található.



5. ábra

*A [t] (bal oldalon), a [s] (középen) és a [ʃ] (jobb oldalon) hang ejtésének palatográfiai vizsgálata szabályos nyelés esetén (K4): mindhárom esetben a normának tekintett variánsnak (vö. BOLLA 1995) megfelelő festékleNyomatokat látunk*

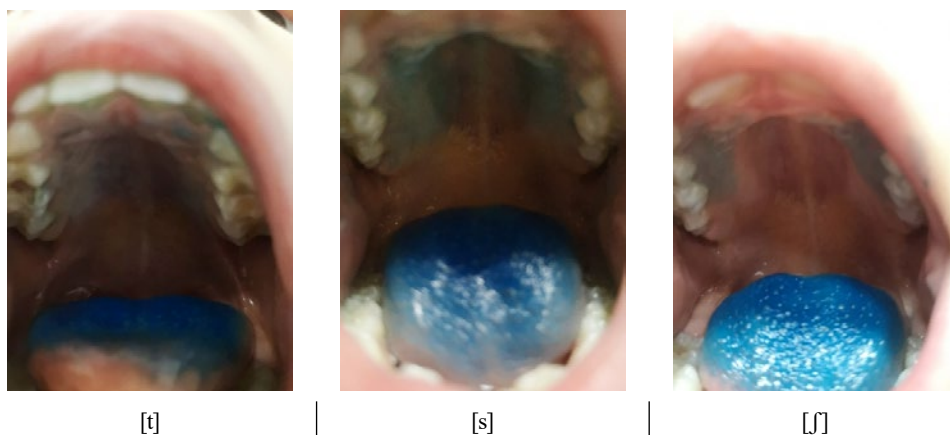
### 3.2.3.2. A második szabályos nyelő

A 8;0 éves kontrollszemély (K2) esetében fordult elő a legkiterjedtebb formában a beszédhangajtás zavara a szabályosan nyelő kísérleti személyek közül: a tizennégyből hat beszédhang esetében találtunk beszédhanghibát. Ez a vizsgálat előtt szintén felvetette a szabálytalan nyelésfunkció lehetőségét, a megelőző logopédiai vizsgálat ugyanakkor kizárta ezt, amit a palatográfia is megerősített: a palatumon festékes lenyomatot találtunk, míg a fogsorokon nem tapasztaltunk festékeződést.

A beszédhangok ejtésének auditív-vizuális vizsgálata során az alveoláris [s], [z] és [ts] esetében nem besorolható, a posztalveoláris [ʃ], [ʒ] és [tʃ] esetében dentális, tehát foghoz nyomott képzést állapítottunk meg.

A 6. ábra bal oldalán a [t] palatogramját figyelhetjük meg K2 ejtésében, melyről az auditív-vizuális vizsgálat során normatív ejtést állapítottunk meg. Látható, hogy ezt a minősítést a palatogramon rögzített kép ugyancsak alátámasztja, hiszen a 4. ábrához hasonlóan itt is az alveoláris területen kezdődő kétoldali lenyomat képe rajzolódik ki. K2 esetében az auditív-vizuális vizsgálat alapján nem besorolható típusú beszédhanghibát találtunk a [s] esetében. A palatogramon

ugyanezen hang esetében a referenciaként használt mintához hasonlóan kétoldali festékeződést találtunk, de a fogmedri ívben látható hézag nagyobbak mutatkoztak, mint a BOLLA (1995) által közölt lenyomaton. A [ts] hang esetében az auditív-vizuális vizsgálat alapján addentális képzést állapítottunk meg, a palatográfia során ugyanakkor a referenciának megfelelő festéklenyomatot tapasztaltunk, ami alapján (legalábbis a palatogram szerint) normatív minősítés adható a hang ejtésének.



6. ábra

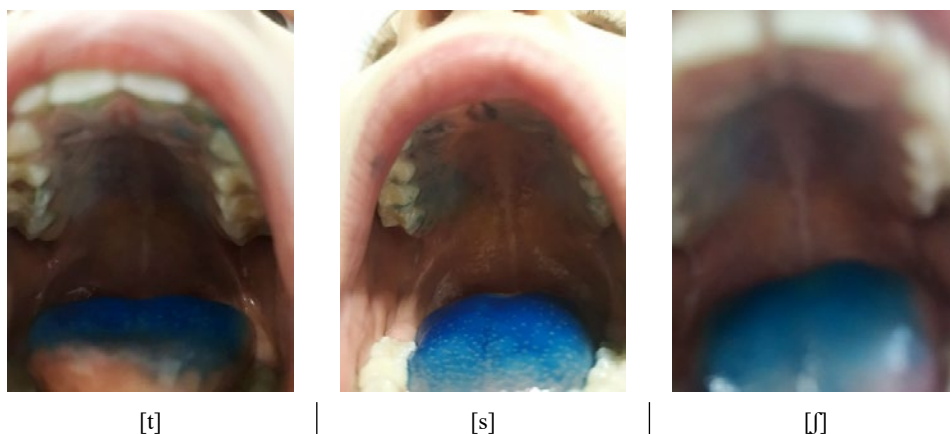
A [t], a [s] és a [ʃ] hang ejtésének palatográfias vizsgálata szabályos nyelés esetén (K2)

### 3.2.3.3. Egy LOMD-t mutató beszélő

A 8;5 éves tesztszemély (T5) nyelésfunkciója a megelőző logopédiai vizsgálatban szabálytalanak bizonyult, amit a palatográfia is megerősített: a szájpadi (halványabb) festékes lenyomat (lásd 4. ábra, jobb oldal) mellett a felső fogsor lingvális felszínén is találtunk festéknyomatot. A 2.2.3. fejezetben bevezetett osztályozási rendszer szerint ezt a nyelésmódot a szabálytalanság jellege alapján a maxilláris (felső állkapcsi) LOMD-csoportba sorolhatjuk.

Az auditív-vizuális vizsgálat szerint a [t], [d], [n] hangok mindegyike vagy addentális, tehát fognyomósos, vagy interdentális, vagyis fogsorok közötti ejtést mutatott. Szórványos jelleggel, csupán egy-egy esetben az alveoláris [l] és [r] hangok unilaterális, illetve laterális képzését tapasztaltuk, ami ebben az esetben azt jelenti, hogy a nyelv a fogsor egyik vagy másik oldalához nyomódva képzett zárt a kérdéses hangok ejtése közben.

A [t] hang palatográfias vizsgálata megerősítette a beszédhangéjtés auditív-vizuális vizsgálata során tett megfigyelést, ugyanis a normatív szájpadi lenyomat helyett a felső fogsor lingvális ívében jelent meg a festékeződés (7. ábra, bal oldal).



7. ábra

A [t], [s] és [ʃ] hang ejtésének palatogramja az LOMD-t mutató kísérleti személy esetén

Az [s] és [ʃ] esetében a beszédhangejtés auditív-vizuális vizsgálata során normatív képzést állapítottunk meg, melyet a palatográfia is alátámasztott, itt ugyanis a korábbiaknak megfelelően a BOLLA (1995) által közölt palatogramokhoz hasonló lenyomatokat láttunk (3. ábra).

#### 3.2.3.4. Az esetbemutatókból levonható következtetések

Az elemzett esetek alapján megfogalmazhatók sejtések a nyelés és a beszédhangejtés szabálytalanságainak lehetséges összefüggéseiről, valamint tehetők észrevételek az auditív-vizuális vizsgálat és az eszközös elemzések lehetőségeiről, illetve érvényességéről. A két szabályos nyelő példája alapján láthattuk, hogy a szabályos nyelés együtt járhat a beszédhangok ejtésének szabályszerűségével, de akár nagyobb számú beszédhanghibával is. Ebből egyértelműen levonható az a következtetés, hogy a nyelés szabályossága nem elégséges feltétele a hangzóejtés szabályosságának. Ráadásul mivel az LOMD-t mutató beszélő ejtésében is találtunk (az auditív-vizuális vizsgálat és a palatográfia szerint is) szabályosan ejtett beszédhangokat az alveolárisok között (ezek voltak a réses képzésű [s] és [ʃ] hangok), az a következtetés is megfogalmazható, hogy a nyelés szabályossága egyúttal nem is szükséges feltétele a beszédhangok szabályos ejtésének, hiszen szabálytalan nyelésben is előfordulhat.

A nyelés és ejtés jellegének összefüggéseivel kapcsolatosan levonható következtetések a kis elemszám és az esettanulmányi jelleg miatt nem általánosíthatók túlságosan, de tesztelendő felvetéseként, hipotézisekként szolgálhatnak a jövőbeli kutatásokhoz. Az egyik ilyen állítás például az lehet, hogy a nyelés és a zárhangok ejtése közben tapasztalható kontaktus a nyelv és a szájüreg boltozata között összefüggést mutat, hiszen a palatográfia segítségével fény derült arra, hogy az LOMD-t mutató beszélő produkciójában mind a nyelés, mind pedig a [t] hang ejtése során a nyelv az elülső felső metszőfogakhoz nyomódott. Kérdésként fogalmazható meg, hogy vajon ez az összefüggés mennyiben általánosítható, azaz igaz-e, hogy a maxilláris LOMD-t mutató beszélők

általában a mássalhangzóéjtésben is addentális képzési helyet produkálnak (az alveoláris vagy más képzési helyek helyett). Szintén kérdés, hogy vajon ez a képzéshelybeli eltolódás csak a zárhangokat érinti-e, avagy előfordulhat-e például a /l/, /r/ magánhangzószerű hangok vagy a nazálisok esetében, illetve akár a réshangokban (amire a jelen kísérleti személy esetében nem láttunk példát). Kérdésként merül fel továbbá az is, hogy csak az alveolárisok esetében figyelhető-e meg a képzési hely addentálissá változása, vagy jellemző-e más (itt nem vizsgált) „elülső” képzésű mássalhangzókra (azaz más alveolárisokra, posztalveolárisokra, esetleg palatálisokra). Végezetül pedig az is felvethető, hogy nemcsak az addentális (fognyomósos) nyelési mintázat tevődik át a beszédhangejtésre (és jelentkezik addentális képzési helyként), hanem más szabálytalanságok is, tehát például a mandibuláris (az alsó fogsorhoz nyomódó nyelvvel történő) nyelés is jelentkezik az alsó fogsorhoz nyomott nyelvvel képzett szabálytalan mássalhangzóéjtéssel és így tovább.

Ami az alkalmazott módszertanokat illeti, láttuk, hogy a logopédiai vizsgálóeljárás és a palatográfia egyes esetekben eltérő eredményeket hozott. Ez felhívja a figyelmet arra, hogy a logopédiai vizsgálat szubjektivitása ellensúlyozható az eszközös elemzéssel, illetve az eszközös elemzés hiányosságai is kiküszöbölhetők (legalábbis részben) a logopédiai vizsgálatból származó benyomásokkal. A jelen vizsgálatban használt palatográfiának mint módszernek jelentős hátránya, hogy abban a nyelv dinamikus mozgásának csak egyetlen statikus (ráadásul kumulatív) lenyomatát kapjuk (tehát ha több érintkezés történik egy kiejtett hangsor alatt, a létrejövő lenyomat minden érintkezést együttesen és utólag szétválaszthatatlanul fog leképezni). Ezzel szemben a tapasztalt logopédus auditív és vizuális feldolgozórendszere a dinamikus éjtésből von le következtetéseket, melyek ennél fogva részletesebbek és akár pontosabbak is lehetnek – legalábbis bizonyos tekintetben. Hangsúlyozandó ugyanakkor, hogy az eszközös vizsgálat kétségtelen előnye a szubjektív észlelettel szemben az, hogy az eszközzel felvett „nyers” adat ellenőrizhető és bármikor (akár más módszerekkel is) újra-elemezhető: a palatogramokról készült fotók újraosztályozhatók (más kutatók által is), illetve azokon a későbbiekben további elemzések, akár például mérések is végezhetőek lehetnek. A logopédiai vizsgálóeljárással felvett adatok viszont már önmagukban, „nyers” formájukban is egy elemzési folyamat termékei, hiszen az ily módon dokumentált megfigyelések általában már az észleltekből levont következtetések. Mindenképpen érdemesnek látszik tehát szorgalmazni az olyan olcsó és hozzáférhető eszközök bevonását akár a logopédiai vizsgálatok hétköznapi gyakorlatába is, mint amilyen a palatográfia, és növelni ezzel a diagnosztikai módszerek megbízhatóságát. Az pedig okvetlenül szükségesnek látszik, hogy a tudományos vizsgálatokban a szubjektívebb logopédiai módszerek mellett megjelenjenek az objektívebb megfigyelési technikák, azokat kiegészítendő.

## 4. Összefoglalás

Tanulmányunkban nagy vonalakban áttekintettük a nyelés folyamatát és lehetséges szabálytalan működését, illetve a kérdés kapcsolódó szakirodalmát, továbbá bemutattuk a szabálytalan nyelés összefüggéseit a beszédhangejtés esetlegesen megjelenő problémáival a magyar nyelv

beszédhangjaira szorítkozva összefoglalásunkban. Az áttekintésből úgy véljük – és reméljük –, kiderült, hogy a kérdéses folyamatok (a nyelés és a beszédképzés), valamint ezek összefüggései komplexnek mondhatók, és ezek további, objektív eszközöket mozgósító vizsgálatának fontossága vitathatatlan.

## FOGALMAK

*orofaciális funkció; rágás; nyelés; az érett nyelés folyamata; bolus; diszfágia; orofaciális miofunkcionális diszfunkció (OMD); nyelési OMD; lingvális orofaciális miofunkcionális diszfunkció (LOMD); élettani vagy fiziológiai nyelési OMD; fejlődési nyelési OMD; szerzett nyelési OMD; neutrális nyelési OMD; deformitással járó nyelési OMD; beszédhanghiba; organikus eredetű beszédhanghiba; ismeretlen eredetű beszédhanghiba; interdentális képzés; addentális képzés; laterális képzés; palatográfia; palatogram*

## IRODALOM

- ALLEN, Jodi E. – CLUNIE, Gemma M. – SLINGER, Claire – HAINES, Jemma – MOSSEY-GASTON, Corinne – ZAGA, Chariss J. – SCOTT, Becky – WALLACE, Sarah – GOVENDER, Roganie 2020. Utility of ultrasound in the assessment of swallowing and laryngeal function: A rapid review and critical appraisal of the literature. *International Journal of Language and Communication Disorders* 56/1. 1–31. <https://doi.org/10.1111/1460-6984.12584>
- AMERICAN SPEECH-LANGUAGE-HEARING ASSOCIATION 2004. *Preferred Practice Patterns for the Profession of Speech-Language Pathology [Preferred Practice Patterns]*. <https://www.asha.org/siteassets/publications/pp2004-00191.pdf> (A letöltés ideje: 2021. január 5.)
- ANDERSON, Victoria B. 2008. Static palatography for language fieldwork. *Language Documentation & Conservation* 2/1. 1–27. <http://hdl.handle.net/10125/1808> (A letöltés ideje: 2021. március 16.)
- ARVEDSON, Joan C. – BRODSKY, Linda – LEFTON-GREIF, Maureen A. 2020. *Pediatric Swallowing and Feeding*. Plural Publishing Inc., San Diego, CA.
- BARRETT, Richard. H. – HANSON, Marvin L. 1974. *Oral Myofunctional Disorders*. Mosby, Saint Louis, MO.
- BEAN, Allison 2013. Oral-motor skills. In VOLKMAR, Fred R. (ed.): *Encyclopedia of Autism Spectrum Disorders*. Springer, New York.
- BILLINGS, Mary – GATTO, Kristie – D’ONOFRIO, Linda – MERKEL-WALSH, Robyn – ARCHAMBAULT, Nicole 2018. *Orofacial Myofunctional Disorders*. <http://iaom.com/wp-content/uploads/2018/10/OMD-Overview-IAOM.pdf> (A letöltés ideje: 2019. március 9.)
- BJÖRK, Arne 1960. Deglutition. In LUNDSTROM, Anders (ed.): *Introduction to Orthodontics*. Ivar Haeggströms Boktryckari Ab., Stockholm.
- BLOOMER, Harlan H. 1971. Speech defects associated with dental abnormalities and malocclusions. In TRAVIS, Lee Edward (ed.): *Handbook of Speech Pathology and Audiology*. Appleton-Century-Crofts, New York.
- BOLLA Kálmán 1995. *Magyar fonetikai atlasz. A szegmentális hangszerkezet elemei*. Nemzeti Tankönyvkiadó, Budapest.

- BRAUER, James S. – HOLT, Townsend V. 1965. Tongue thrust classification. *The Angle Orthodontist* 35/2. 106–112.
- CLELAND, Joanne – SCOBIE, James M. – WRENCH, Alan A. 2015. Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics and Phonetics* 29/8–10. 575–597.
- DAMICO, Jack S. – BALL, Martin J. 2019. *The SAGE Encyclopedia of Human Communication Sciences and Disorders*. SAGE Publications, Thousand Oaks.
- DEME Andrea 2016. *Magánhangzók ejtése és észlelése a szopránéneklésben*. ELTE Eötvös Kiadó, Budapest.
- DEME Andrea – BARTÓK Márton – GRÁCZI Tekla Etelka – CSAPÓ Tamás Gábor – MARKÓ Alexandra 2019. A mondathangsúly hatása a magánhangzók megvalósulásának változosságára. *Nyelvtudományi Közlemények* 115. 199–232.
- DODDS, Wylie J. – STEWART, Edward T. – LOGEMANN, Jeri A. 1990. Physiology and radiology of the normal oral and pharyngeal phases of swallowing. *American Journal of Roentgenology* 154/5. 953–963.
- FÁBIÁN Gábor – GÁBRIS Katalin – TARJÁN Ildikó 2013. *Gyermekfogászat, fogszabályozás és állcsont-ortopédia*. Semmelweis Kiadó, Budapest.
- FEHÉRNÉ KOVÁCS Zsuzsa 2013. A beszéd- és nyelvféjlődés zavarainak logopédiai vonatkozásai és kezelése. In HIRSCHBERG Jenő – HACKI Tamás – MÉSZÁROS Krisztina (szerk.): *Foniátria és társtudományok II*. ELTE Eötvös Kiadó, Budapest. 24–36.
- GRÁCZI Tekla Etelka – CSAPÓ Tamás Gábor – DEME Andrea – JUHÁSZ Kornélia – MARKÓ Alexandra 2020. A réshangok zöngésségével összefüggő nyelvpozíciós jellemzők a megelőző magánhangzóban. *Nyelvtudományi Közlemények* 116. 155–190.
- FONYÓ Attila 2011. *Az orvosi élettan könyve*. Medicina Könyvkiadó Zrt., Budapest.
- GÓSY Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- HANSON, Marvin L. – COHEN, Melvin S. 1973. Effects of form and function on swallowing and the developing dentition. *American Journal of Orthodontics* 64/1. 63–82.
- HANSON, Marvin L. – MASON, Robert M. 2003. *Orofacial Myology: International Perspectives*. Charles C. Thomas Publisher, Springfield, IL.
- HAVADI-NAGY Marian 2018. Intraorális szenzitivitás és orális sztereognózia. A szájérzékelés fejlődésének szerepe a logopédiai gyakorlatban. In KARLOVITZ János T. (szerk.): *Elmélet és gyakorlat a neveléstudományok és szakmódszertanok köréből*. International Research Institute sro, Komárno, Slovakia. <http://www.irisro.org/pedagogia2018januar/34Havadi-NagyMarian.pdf> (A letöltés ideje: 2018. március 10.)
- HAVADI-NAGY Marian 2020a. A szabálytalan nyelésfunkció (myofunkcionális diszfunkció) klasszifikációs lehetőségei. In FÓRIS Ágota – BÖLCSKEI Andrea – NÁDOR Orsolya – SÓLYOM Réka (szerk.): *Nyelv, kultúra, identitás. V. Nyelvpedagógia, nyelvoktatás, nyelvelsajátítás*. Akadémiai Kiadó, Budapest. [https://mersz.hu/dokumentum/m716nyki5nynyny\\_\\_30](https://mersz.hu/dokumentum/m716nyki5nynyny__30) (A letöltés ideje: 2020. március 30.)
- HAVADI-NAGY Marian 2020b. Az atipikus nyelés és a beszédhangejtés vizsgálata palatográfiával gyermekeknél. In VÁRADI Tamás (sorozatszerk.) – LUDÁNYI Zsófia – GRÁCZI Tekla



- Etelka (szerk.): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2020. XIV. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Nyelvtudományi Intézet, Budapest. 32–50. doi: 10.18135/Alknyelvdok.2020.14.3
- JUHÁSZ Ágnes 2007. *Logopédiai vizsgálatok kézikönyve*. Logopédia Kiadó KKT, Budapest.
- KOVÁCS Emőke – REHÁK Gizella 2013. Artikulációs zavar, pöszeség. In HIRSCHBERG Jenő – HACKI Tamás – MÉSZÁROS Krisztina (szerk.): *Foniátria és társtudományok II*. ELTE Eötvös Kiadó, Budapest. 41–48.
- KRAVANJA, Sanda L. – HOCEVAR-BOLTEZAR, Irena – MAROLT MUSIC, Maja – JARC, Ana – VERDENIK, Ivan – OVSENIK, Maja 2018. Tongue posture and its impact on articulation disorders in preschool children with anterior open bite. *Radiology and Oncology* 52/3. 250–256.
- KRUGER, Danielle 2014. Assessing esophageal dysphagia. *Journal of the American Academy of Physician Assistant* 27/5. 23–30.
- LADEFOGED, Peter 2003. *Phonetic Data Analysis*. Blackwell, Oxford.
- LUCHSINGER, Richard – ARNOLD, Gottfried E. 1965. *Voice–Speech–Language*. Wadsworth Publishing, Belmont, CA.
- MARKÓ Alexandra 2017. Hangtan. In IMRÉNYI András – KUGLER Nóra – LADÁNYI Mária – MARKÓ Alexandra – TÁTRAI Szilárd – TOLCSVAI NAGY Gábor: *Nyelvtan*. Osiris Kiadó, Budapest. 75–206.
- MARKÓ Alexandra – CSAPÓ Tamás Gábor – BARTÓK Márton – DEME Andrea – GRÁCZI Tekla Etelka 2020. Magyar gyermekek artikulációs vizsgálatának lehetőségei: miért és hogyan? In BÓNA Judit – KREPSZ Valéria (szerk.): *Nyelvfejlődés csecsemőkortól kamaszkorig*. ELTE Eötvös Kiadó, Budapest. 23–44.
- MÉSZÁROS Krisztina – HACKI Tamás 2013. Nyelés, nyelészavarok. In HIRSCHBERG Jenő – HACKI Tamás – MÉSZÁROS Krisztina (szerk.): *Foniátria és társtudományok II*. ELTE Eötvös Kiadó, Budapest. 201–230.
- MORRIS, Suzanne 1978. Oral motor development: Normal and abnormal. In WILSON, Janet M. (ed.): *Oral-Motor Function and Dysfunction in Children*. Division of Physical Therapy, UNC, Chapel Hill, NC.
- OHKUBO, Mai – SCOBIE, James M. 2018. Tongue shape dynamics in swallowing using sagittal ultrasound. *Dysphagia* 34. 112–118.
- PENG, Chien-Lun – JOST-BRINKMANN, Paul-Georg – YOSHIDA, Noriaki – CHOU, Hsin-Hua – LIN, Che-Tong 2004. Comparison of tongue functions between mature and tongue-thrust swallowing — an ultrasound investigation. *American Journal of Orthodontics and Dentofacial Orthopedics* 125/5. 562–570.
- PATIL, Sayam – SANJEEV, Jakati – PATIL, Rutika 2017. Malocclusion, phonetics & palatography: The link express. *Global Journal of Medical Research* 17/1. 19–24.
- SCHREY-DERN, Dietlinde 2006. *Sprachentwicklungsstörungen. Logopädische Diagnostik und Therapieplanung*. Georg Thieme Verlag, Stuttgart.
- SEIKEL, Anthony J. – KING, Douglas W. – DRUMRIGHT, David G. 2010. *Anatomy & Pshysiology for Speech, Language and Hearing*. 4th edition. Cengage Learning, Delmar.

- SURJÁN László – FRINT Tibor (szerk.) 1982. *A hangképzés és zavarai, beszédzavarok*. Medicina Könyvkiadó, Budapest.
- SZENTÁGOTHAJ János – RÉTHELYI Miklós 2006a. *Funkcionális anatómia 1*. Medicina Könyvkiadó Zrt., Budapest. [https://regi.tankonyvtar.hu/hu/tartalom/tamop425/2011\\_0001\\_524\\_Funkcionalis\\_anatomia\\_1/adatok.html](https://regi.tankonyvtar.hu/hu/tartalom/tamop425/2011_0001_524_Funkcionalis_anatomia_1/adatok.html) (A letöltés ideje: 2019. február 25.)
- SZENTÁGOTHAJ János – RÉTHELYI Miklós 2006b. *Funkcionális anatómia 2*. Medicina Könyvkiadó Zrt., Budapest. [https://regi.tankonyvtar.hu/hu/tartalom/tamop425/2011\\_0001\\_524\\_Funkcionalis\\_anatomia\\_2/adatok.html](https://regi.tankonyvtar.hu/hu/tartalom/tamop425/2011_0001_524_Funkcionalis_anatomia_2/adatok.html) (A letöltés ideje: 2019. február 25.)
- TAR Éva 2017. *Fonológiai fejlődés, variabilitás, beszédhanghibák*. ELTE Eötvös Kiadó, Budapest.
- THOROCZKAY Miklósné 2016. *Beszédhangzók fejlesztése*. Logopédia Kiadó, Budapest.
- VASSNÉ KOVÁCS Emőke (szerk.) 1998. *Logopédiai jegyzet 1*. Nemzeti Tankönyvkiadó, Budapest.
- VASSNÉ KOVÁCS Emőke – REHÁK Gizella – VIZKELETY Tamás 2015. *A nyelvlökéses nyelés*. Eötvös József Kiadó, Budapest.
- WIEMER, Heather 2017. *Orofacial Myofunctional Disorders*. [https://www.ebscohost.com/assets-sample-content/Orofacial\\_Myofunctional\\_Disorders.pdf](https://www.ebscohost.com/assets-sample-content/Orofacial_Myofunctional_Disorders.pdf) (A letöltés ideje: 2019. február 25.)

## KÖSZÖNETNYILVÁNÍTÁS

A tanulmány a Bolyai János Kutatási Ösztöndíj, a Tématerületi Kiválósági Program és az Innovációs és Technológiai Minisztérium ÚNKP-20-5 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.



# Idegennyelv-elsajátítás és akcentus

Juhász Kornélia

ELTE EÖTVÖS LORÁND TUDOMÁNYEGYETEM

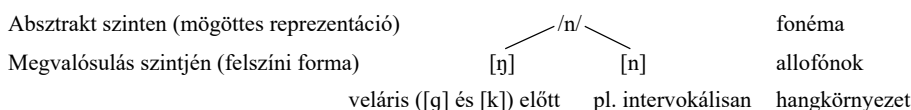
## 1. Bevezetés

Globalizálódó világunkban az **idegen nyelv** (célnyelv, L2) elsajátítása létfontosságúnak számít az ember életében. Kommunikációnk feltétele, hogy gond nélkül megértsük egymást, ez **anyanyelvi** (L1) környezetben nem jelent problémát. Azonban ha egy nyelvtanuló a célnyelvén szólal meg, az anyanyelvitől jellegzetesen eltérő kiejtése sok esetben befolyást gyakorol a megítélésére, akadályozhatja a kommunikációt, és gyakran előítéleteket szülhet (FLEGE 1988). Az előbb említett kiejtésbeli sajátosságokat, amelyek az L2 normatív kiejtéséhez képest eltérően jelennek meg az L2 nyelvközösségéhez nem tartozó, de annak nyelvén megnyilatkozó személy beszédében, **akcentusnak** nevezzük (KASSAI 1995). A jelen tanulmány az akcentus szegmentális és érintőlegesen prozódiai/szupraszegmentális jegyeit mutatja be, azaz elsősorban a beszédhangokkal és fonémákkal (**szegmentális szerkezet**) foglalkozik, azonban néhány **prozódiai/szupraszegmentális** (azaz hanglejtéshez, hangsúlyhoz, tónusokhoz kapcsolódó) jelenségről is szó esik.

Ha a beszédhangokkal kapcsolatos kiejtési pontatlanságokra szeretnénk magyarázatot találni, akkor azt mondhatjuk, hogy az L2-től való eltérés a legtöbb esetben olyan beszédhangoknál következik be, amelyek vagy teljesen hiányoznak az L1-ből, vagy az L1-től eltérően valósulnak meg az L2-ben (JAMES 1986). Azonban önmagában az összes L2-beszédhang helyes ejtése sem garantálja, hogy nincs akcentusunk. Például az angolul tanulók számára nyilvánvaló, hogy szemben a magyarral, ahol mindig a szó első szótagjára esik a szóhangsúly, az angolban a hangsúlyos szótag helye egyrészt nem megjósolható, másrészt jelentésmegkülönböztető szereppel is bírhat. Ugyanígy gondot okozhat a tonális nyelvek elsajátításakor, mint amilyen a kínai vagy a vietnami, hogy minden egyes szótaghoz különböző dallammenetek járulhatnak, így a szótagok helyes ejtéséhez „énekelnünk” is kell, és a tónus megválasztásának is jelentésmódosító hatása van.

A szegmentumokat (beszédhangokat) eltérő szempontok szerint vizsgáló tudományágak a fonetika és a fonológia. A **fonetika** esetében a beszédhangok és prozódiai jellemzők fiziológiai és fizikai természete áll a középpontban, illetve a **beszédhangok** változatos megvalósulása is komoly szerephez jut a kapcsolódó empirikus kutatásokban. A fonetika változatos beszédhang-megvalósulásával szemben a **fonológia** egymástól jól elkülönülő, absztrakt beszédhangtípusokkal, azaz fonémákkal foglalkozik, valamint ezek rendszerével, a rendszerbe szerveződésük tudattalan szabályait vizsgálva. Ilyen szabályok például a (hang)környezet

függvényében fellépő változások, illetve a beszédhangok összefűzéséhez (szekvenálásához) és eloszlásához (disztribúciójához) kötött szabályok (HAYES 2009). A fonológia központi szereplője a **fonéma**, amely egy elvont, mögöttes nyelvi egység, mentális reprezentáció, ami magában foglalja az adott beszédhang minden változatát. A fonéma környezetétől függően eltérően változhat meg, ezeket a felszíni megvalósulásokat **allofón**oknak nevezzük. Az IPA (nemzetközi fonetikai ábécé) szerint a beszédhangokat szögletes zárójel jelöli (pl. [ŋ]), míg a fonémákat ferde zárójel jelzi (pl. /n/). A fonéma–allofón kapcsolatra, azaz az /n/ fonémára és egy-egy allofónjára az 1. ábra mutat be példát.

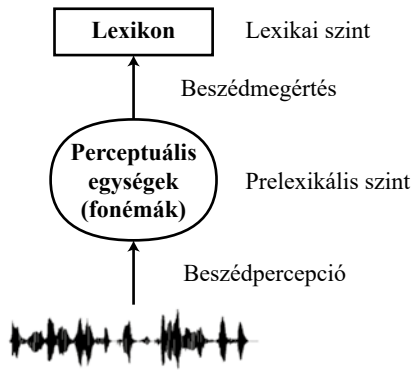


1. ábra

*Az /n/ fonéma különböző kontextusú allofónjai és megvalósulásának szintjei*

Az 1. ábrán bemutatott allofónok közötti eltérést úgy tudjuk megfigyelni, ha megnézzük, a nyelvünk mely része és hol ér a szájpadhoz, amikor a *fon* [fon], illetve a *hang* [hɒŋg] szavak /n/-jét ejtjük ki. Annak dacára, hogy a *hang* esetében a nyelv egy hátsóbb része sokkal hátrébb érinti a szájpadot, mint a *fon* esetében, ahol a nyelv elülső része a fogmederhez (fogínyhez) ér, mégis mindkét változatot /n/-ként azonosítjuk. Azonban, ha az angol nyelvet nézzük, akkor például a *sin* [sɪn] és a *sing* [sɪŋ] szavak már minimálpárt alkotnak, azaz ebben az esetben az /n/ és az /ŋ/ két különböző fonéma, mivel jelentésmegkülönböztető szereppel rendelkeznek (McMAHON 2002). A fonetika és a fonológia közötti különbséget úgy tudjuk nagyon egyszerűen megfogalmazni, hogy míg a fonetikában a színárnyalatokhoz hasonlóan fokozatos (graduális) az átmenet az egyes beszédhangok között, ezzel szemben a fonológiában a kategóriahatárok tisztán elkülöníthetőek.

Az (idegen)nyelv-elsajátítást elsősorban a beszédmegértés és a beszédprodukciónak a teszik lehetővé. Az emberi beszéd feldolgozását, azaz azt a folyamatot, hogy a beszéd információ-tartalmát kognitív folyamatokkal dekódoljuk, kétszintű folyamatként tudjuk értelmezni, melynek két szintje a beszédészlelés (beszédpercepció), illetve a beszédmegértés. A **beszédpercepció** prelexikális, azaz relatíve alacsonyabb rendű, de nyelvspecifikus folyamat, amely során a hallgató a folyamatos akusztikai kontinuum értékeit képes nyelvi egységekké dekódolni. Ezzel szemben a **beszédmegértés** már magasabb, lexikai szintet képvisel, amely során szemantikai és pragmatikai folyamatok közreműködésével képesek vagyunk az akusztikai jeleket a mentális lexikon szójelentéseinek, lexikai reprezentációinak megfeleltetni (McQUEEN 2004). A **beszédfeldolgozás** kétszintű modelljét a 2. ábra mutatja be. Az utóbbi, beszédmegértési folyamattal a jelen tanulmány nem foglalkozik, hiszen az akcentus kialakulása elsősorban a beszédpercepcióhoz kötődik.



2. ábra

*A beszédfeldolgozás kétszintű modellje, azaz a beszédészlelés és megértés folyamatai (ESCUDERO 2005: 29 nyomán)*

A **beszédprodukción** közben a beszélő az elméjében megformázza és nyelvileg kódolja a hallgatónak szánt üzenetet, majd beszéd szervei segítségével hallható jelként üzenetet bocsát ki (azaz beszédhangokat, szavakat, mondatokat ejt ki: beszél) (LEVELT 1989).

A tanulmány a következő idegennyelv-elsajátítással és akcentussal kapcsolatos témákat érinti: általános történeti áttekintés keretein belül tárgyaljuk az L2-elsajátítás és az akcentuskutatás főbb állomásait és néhány további fontos fogalmat. Majd a percepció és a produkció kapcsolata kerül középpontba, pontosabban arra a kérdésre keressük a választ, hogy miért a percepció szolgált az L2-beli beszédhangok elsajátítását vizsgáló modellek alapjául. Ezután röviden összefoglalunk két, az L2-re vonatkozó, beszédhang-elsajátítást magyarázó modellt, illetve a tanulmány végén bemutatunk néhány ezt befolyásoló tényezőt is.

## 2. Történeti áttekintés

Az L2 kiejtésével, az akcentussal kapcsolatos vizsgálatokat történeti szempontból ECKMAN (2012) szerint két nagyobb korszakra bonthatjuk, mely korszakok határát az *interlanguage*, avagy „köztes nyelv” fogalomhoz kötjük. E fogalom meghatározása előtt azonban először az akcentuskutatás kezdeti fázisát mutatom be. A kezdeti elgondolás, azaz POLIVANOV (1931) és TRUBETZKOY (1939) szerint az anyanyelv hatással van a célnyelv elsajátítására, azaz az L2-beszédhangok az L1 hangrendszerén keresztül kerülnek feldolgozásra. Ez azt is jelenti, hogy az L2-beszédhangok inadekvát produkciója percepcióalapú, ahol az L1 rendszere fonológiai szűrőként működik, és ezáltal kategorizáljuk percepciósan az L2-beszédhangokat. Más szóval az akcentus nem a produkció terméke, hanem már a percepció szintjén is jelen van, ez a „perceptuális akcentus” (STRANGE 1995). Azt, hogy

a nyelvtanulók az L1 fonológiai rendszere alapján dolgozzák fel az L2-beszédhangokat, a transzfer fogalmával, avagy a nyelvek közötti interferenciával magyarázták (ODLIN 1989). A **transzfer** az ember általános tanulási folyamatainak alapja, azaz ha új tanulási szituációval kell megküzdenünk, akkor az addig elsajátított mintázatokat, módszereket elővéve próbáljuk leküzdeni az akadályokat. Éppen ezért az L2 elsajátításakor is a korábban már elsajátított nyelvi rendszer, azaz az anyanyelv mintázatait vesszük elő, és próbáljuk alkalmazni az új nyelv elsajátítására (MAJOR 2001). Ez azonban – mint látni fogjuk – az L1 és az L2 közötti különbségekből fakadóan nem minden esetben célravezető módszer. Többek között azért sem, mert ha az L1 transzferált mintázatai eltérnek az L2 mintázataitól, akkor **negatív transzfer**, avagy **interferencia** lép fel a két nyelv között, ami hibákat eredményez, és hátráltatja az L2 helyes elsajátítását. Ellenkező esetben, ha az L1-ből az L2-be áttanszferált rendszer megegyezik, akkor **pozitív transzferről** beszélünk, azaz az L2 adott jellemzőjének elsajátítása gond nélkül és gyorsan megy végbe (BARDOVI–HARLIG–SPOUSE 2017).

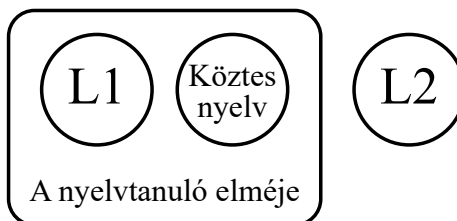
Mint fentebb említettük, az akcentus alapját az a jelenség szolgáltatja, hogy az L1 befolyásolja az L2 elsajátítását. Ezért a köztes nyelv fogalmának megjelenése előtt a nyelvészek az L1 és az L2 hangrendszereit vetették össze a kontrasztív analízis keretein belül. A **kontrasztív analízis** alapja, hogy minél nagyobb a különbség az L1 és az L2 hangrendszere között, annál nehezebb elsajátítani az L2 beszédhangjait. A kontrasztív analízisben fonológiai elemzéssel hasonlították össze az L1-et és az L2-t, azaz vetették össze a fonémák és allofónok eloszlását és megjelenési szabályszerűségeit (ECKMAN 2004). Az L2-beszédhangok ejtési nehézségeit és hibáit az allofónok közötti eltérő eloszlással magyarázták. LADO (1957) szerint az a legnehezebben elsajátítható jelenség, ha a nyelvtanulónak az L1 allofónjait különböző L2-kategóriákba kell sorolni. Erre példaként hozhatjuk fel azt, hogy a spanyol beszélők számára problémát okoz az angol /ð/ és /d/ megkülönböztetése, hiszen a spanyolban mind a két beszédhang a /d/ fonéma kategóriájába sorolandó, de ide csatolhatjuk a fentebb angol /n/ és /ŋ/ megkülönböztetését magyar anyanyelvűek számára is. Később a különböző fonológiai eltérések szituációit nyolcszintes hierarchikus rendszerbe sorolták, ahol az L1-ben jelen nem lévő L2-allofón elsajátítása számított a legnehezebbnek. A kontrasztív analízis által alkotott ejtési nehézségekkel kapcsolatos predikciók általánosságban jól működtek, azonban mégsem tudtak minden jelenséget magyarázni velük (ECKMAN 2012).

Ezt a problémát TRUBETZKOY (1939) oldotta fel a **tipológiai jelöltség** elméletével, melynek alapjául az az akkoriban népszerű elgondolás szolgált, hogy a nyelvi jegyeket bináris opozíciókban (pl. zöngés–zöngétlen szembenállás, +zöngés/–zöngés) értelmezték. E bináris rendszerben gondolkodva jelölt és jelöletlen párokat tudunk meghatározni. A jelöletlen elem valamilyen módon egyszerűbb, „könnyebb”, természetesebb, nem csak az adott nyelven belül, hanem általánosságban a nyelvek között is mindig szélesebb körben jelenik meg. A jelölt elem ezzel szemben bonyolultabb, és lényegesen ritkábban jelenik meg. Pontosabban definiálva egy X struktúra (pl. beszédhang vagy képzési jegy) akkor számít jelöltnek egy másik Y struktúrához képest, ha minden nyelvben, amelyben megtalálható X, van Y is; de nem minden nyelvben, amiben van Y, lesz szükségszerűen jelen az X (ECKMAN 2004: 529). Például a legtöbb nyelvben jelen van az ajakkerekítés nélküli /i/ és /e/, azonban ennél jóval kevesebb nyelvben állnak szemben ezek a beszédhangok az ajakkerekítéses /y/ és /ø/ párjukkal. Illetve szótagszerkezet szempontjából: ha egy nyelvben jelen van CVC (mássalhangzó + magánhangzó + mássalhangzó) szerkezetű szótag, akkor kell lennie CV

(mássalhangzó + magánhangzó) szerkezetűnek is (MAJOR 2001). A tipológiai jelöltséget a kontrasztív analízis kiterjesztéseként szerették volna értelmezni, ezáltal definiálva a nehézség fokát, azon az elven, hogy az L2-beszédhang nehezebben elsajátítható, ha nemcsak fonológiailag tér el az L1-től, de emellett még jelöltebb is nála (ECKMAN 2012; PICKERING 2012).

Később a fonológiai rendszerek közötti eltérésekre koncentráló kontrasztív analízis helyett inkább a nyelvek közötti hasonlóság és általános elvek kerültek középpontba (ECKMAN 2012). A generatív nyelvészet atyja, Noam CHOMSKY szerint a nyelv mint olyan (CHOMSKY terminusával az I-nyelv, ahol az I az intencionálist, individuálist és internálist jelöli), az ember elméjében létező „belső” nyelvi tudásként, azaz a nyelvi szabályok összességéként van jelen, amit grammatikának hívunk (CHOMSKY 1986). A nyelvek közötti hasonlóságok legfőbb forrása az **univerzális grammatika** (UG), amelynek minden ember veleszületetten a birtokában van, és ami kezdetben egy konkrét, nyelvektől független, alulspecifikált rendszer, melynek a paramétereit a nyelvelsajátítás folyamata „hangolja be” az adott, elsajátítandó nyelv tulajdonságai szerint (CHOMSKY 1981). ALBERTI és MEDVE (2006) szerint az univerzális grammatikát tekinthetjük a törtszámok közös nevezőjének is, ahol a törtszámok a nyelvek maguk, és a bennük mélyen megbúvó hasonlóságokban lelhetjük meg az UG működését. Az UG általános elve például, hogy a nyelvek rendelkeznek magánhangzókkal és igékkel, azonban ezek a magánhangzók és igék minden nyelvben eltérnek. Az UG-n kívül azonban másfajta, átfogóbb nyelvi univerzálék is jelen vannak (és meghatározzák az L2 elsajátítását). Például a nyelvekben a CV (mássalhangzó + magánhangzó) szerkezetű szótag számít jelöletlennek, azaz ez a legalapvetőbb és legegyszerűbben elsajátítható, ebből kiindulva az L2-t tanuló könnyebben elsajátítja a mássalhangzókat szótagkezdeti helyzetben, mint a szótagzárlati pozícióban (pl. a jelöltebb, CVC szerkezetű szótagok esetében) (MAJOR 2001; CARLISLE 2001).

Az *interlanguage*, avagy a **köztes nyelv** egy egyénekenként eltérő belső mentális állapot, rendszer, amely lehetővé teszi az L2 percepcióját és produkcióját (NEMSER 1971; SELINKER 1972). Ebből fakadóan a köztes nyelv kizárólag a nyelvtanuló elméjében van jelen, és a köztes nyelvben megkonstruált tudás és szabályrendszer nemcsak nyelvtanulónként tér el (éppen úgy, ahogy a nyelvtanuló által befogadott L2-ingerek is személyenként eltérnek), hanem különbözik az anyanyelvi beszélők L2-jétől is, hiszen a transzfer révén a nyelvtanuló egy egész más szemszögből közelíti meg az L2 rendszerét és szabályait. Más szóval a köztes nyelv a nyelvtanuló sajátos verziója vagy speciális nézőpontja az L2-ről, ami egyfajta idiolektusnak is tekinthető (COOK 2006). A köztes nyelv és az L2 viszonyát a 3. ábra szemlélteti.



3. ábra

Az L1, a köztes nyelv és az L2 sematikus viszonyai (COOK 2008: 15 nyomán)



### 3. A percepció és a produkció, valamint a beszédhang-elsajátítás közötti kapcsolat

Mint ahogy korábban már említettük, a beszédpercepció és produkció alapvető követelmény a nyelv-elsajátításhoz, azonban melyik felelős az akcentus kialakulásáért? KLEIN (1986) szerint háromféleképpen magyarázhatjuk az L2-tanulók pontatlan kiejtését: egyfelől az állhat a háttérben, hogy nem hallják a különbséget az L1 és az L2 beszédhangjai között (észlelési probléma), másfelől az is lehet, hogy nem tudják kiejteni az L2-beszédhangot (produkciós probléma), harmadrészt az is megeshet, hogy e két körülmény együtt jelentkezik. A nyelvelsajátításban, legyen az az anyanyelvé vagy a célnyelvé, a percepció látszólag időben mindig megelőzi a produkciót, hiszen azt mondhatnánk, hogy a percepció révén leszünk birtokában az egyes hangok tulajdonságainak, és ezáltal kapunk visszacsatolást a beszédhangok helyes ejtéséről (KUHLMANN 1980). Számos kutatás célozta meg a percepció és a produkció kapcsolatának és függőségi viszonyának vizsgálatát és feltárását, azonban a téma szerteágazó szakirodalmi ellenére – az ellentmondó eredmények miatt – még sincs egyértelműen definiálva, hogy a percepció és produkció milyen kapcsolatban van egymással.

Először nézzünk néhány példát arra, hogy a percepció megelőzi a produkciót az L2-elsajátításban, amiből azt feltételezhetjük, hogy az L2-tanulók kiejtési nehézségei mögött elsősorban perceptuális problémák állnak. BORDEN, GERBER és MILSARK (1983) szerint a kísérletben részt vevő koreai anyanyelvűek a számukra ugyanazon fonémakategóriához tartozó két allofónt, azaz az angol /r/ és /l/ beszédhangokat észlelésükben hatékonyan megkülönböztették, azonban ejtésükben nem. FLEGE, MACKAY és MEADOR (1999) olasz anyanyelvűek esetében az angol beszédhangok percepciójának és produkciójának vizsgálata révén arra a következtetésre jutott, hogy a produkció helyessége a beszédhangok percepciójának pontosságától függ, azaz a percepció fejlődés korlátozza a produkció előmenetelét. Ez utóbbi következtetést JIA és munkatársai (2006) munkája is megerősítette mandarin anyanyelvű, amerikai angoltanuló nyelvtanulók esetében. Emellett ők még olyan tendenciákat is megfigyeltek, miszerint a percepció fejlődését a produkció csak késve éri utol. GRASSEGER (1991) a percepció vezérelte produkció kapcsán arra is következtetett, hogy a perceptuális tesztekkel „megjósolhatjuk”, hogy a nyelvtanulóknak mely L2-beszédhangok ejtése fog nehézséget és kihívást okozni. NEUFELD (1988) az előbb említett jelenséget, miszerint a nyelvtanulók perceptuálisan sokkal jobb eredményeket értek el, mint produkciósan, **fonológiai aszimmetriának** nevezte. Ez a fonológiai aszimmetria, azaz hogy a helyes percepció megkülönböztetés dacára miért marad el az L2-beszédhangok helyes produkciója, a produkcióban fellelhető problémával is magyarázható. A produkcióhoz szükséges egyfelől egy fonetikai és fonológiai reprezentáció, amely tartalmazza az információt az észlelési célról, másfelől a motorikus procedúrát, amely definiálja, hogyan kell ezt a percepciót produkciósan elérni (FLEGE 1986). A gyerekektől eltérően a felnőttek már nehezebben képesek ezeket a számukra új, L2-beli artikulációs mintázatokat kialakítani, illetve a régi mintázatokat átfőrnálni, továbbá nehézségekbe ütközhet az észlelt L2-beszédhang jegyeinek artikulációs mintázatokká való átalakítása is (KALIKOW–SWETS 1972; VALDMAN 1976).

Vannak azonban olyan eredmények is, amelyek cáfolják a percepció elsőbbségét a produkcióhoz képest. Ilyen például SHELTON és STRANGE (1982) szintén japán anyanyelvűekkel, a /r/ és

/l/ percepciós diszkriminációjával és produkciójával kapcsolatban végzett vizsgálata, amelynek eredményei szerint a helyes produkció korábban mutatkozott, mint a percepciós megkülönböztetés. Azonban az ehhez hasonló eredményeket ESCUDERO (2005) és ISBELL (2016) módszertani hibákra és problémákra vezetik vissza. A jelen tanulmányban mindezek alapján elsődlegesnek tartjuk a percepciót a produkcióhoz képest, és ezzel párhuzamosan meghatározó szerepűnek gondoljuk a percepciót az akcentus tekintetében. Látjuk majd azt is, hogy az 5. fejezetben tárgyalt L2-beszédhang-elsajátítási modellek a perceptuális nehézségekből kiindulva magyarázzák az akcentusos ejtést. Továbbá az akcentust úgy határozhatjuk meg, mint a nyelvtanuló beszédének hangzása és az L1 kiejtési normája között észlelt különbséget, ezért az akcentus vizsgálatában a percepciós tesztekét szokták előnyben részesíteni (MUNRO–DERWING 1998: 160).

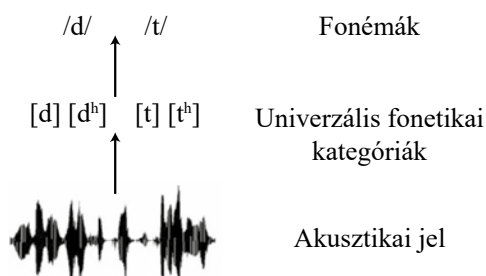
Az anyanyelv-elsajátítás egyik első lépésének tekinthető, hogy a csecsemők képesek észlelni az L1 fonetikai kategóriáinak hasonlóságait és különbségeit (KUHL 1980). Ez a képesség a gyermek növekedésével együtt fejlődik, más szóval a gyerekek fokozatosan, a környezetükben fellelhető ingerek alapján hangolják be a percepciójukat (ASLIN–PISONI 1980). Azonban ehhez a finomhangolási folyamathoz megfelelő és megfelelően változatos ingerek is kellene, amelyek kiemelik az L1 beszédhangjainak főbb akusztikai jellemzőit és a megkülönböztetéshez szükséges akusztikai/felismerési kulcsokat. (Ilyen például a dajkanyelv is, lásd a jelen kötetben KOHÁRI [2021] tanulmányát). **Akusztikai/felismerési kulcsok**nak azokat az akusztikailag releváns jegyeket nevezzük, amelyek alapján meg tudjuk különböztetni a beszédhangokat egymástól (KUHL 2000).

A megfelelő ingerek révén az akusztikai/felismerési kulcsok nyilvánvalóvá válása a nyelvelsajátítás fontos eleme, amely percepciós leképezési és kategóriaalkotási folyamatot BOERSMA, ESCUDERO és HAYES (2003) a *Fokozatos Tanulás* algoritmussal modellezték. A kategóriaformálás alapvetően az egyes akusztikai ingerek gyakoriságán és azok eloszlásán nyugszik, tulajdonképpen ahogy egy adott akusztikai kontinuumon (például magánhangzók esetében a nyelvállásfokkal összefüggést mutató  $F_1$ -skálán) haranggörbeszerűen kirajzolódik az értékek eloszlása. A haranggörbe csúcsánál találjuk a leggyakoribb értékek csoportjait (azaz a kategóriák legrepresentatívabb variánsait), valamint a görbe lapultságából következtetni tudunk arra, hogy a kategóriák mennyire engednek meg variabilitást az értékek között. Az algoritmus működését úgy képzelhetjük el, mintha egy dartsverseny viszonylag jól teljesítő játékosait figyeljük, de magát a céltáblát nem láthatjuk, mégis a lövéseik alapján kell több céltábla helyzetét és középpontját meghatároznunk. Mivel jól teljesítő játékosokról beszélünk, a dobások nem teljesen véletlenszerű eredménnyel valósulnak meg, hanem minden egyes új dobással több és pontosabb információt nyerhetünk a céltáblák helyzetéről. A céltáblák középpontjánál feltételezzük a legtöbb találatot (ezek a találatok a legpontosabb akusztikai megvalósulások), míg a céltáblák peremén már egyre kevesebb a találat, hiszen ezekben az esetekben pontatlan a dobás (avagy a beszédhangok megvalósulása sem teljesen sztenderd). A nyelvelsajátításban is ugyanilyen fokozatosan „hangolódik be” egy-egy beszédhang-kategória, és ezen az elven rendeljük a rendhagyó megvalósulásokat is a meglévő kategóriákba. ESCUDERO (2005) szerint e gyakoriságon és eloszláson nyugvó folyamaton alapul mind az L1, mind az L2 beszédhang-kategóriáinak kialakítása. Azonban az általa vázolt teljes percepciós

folyamat többszintes, és optimalitáselméleti korlátok (PRINCE–SMOLENSKY 1993/2004) szabályozzák (ESCUADERO 2005).

De vajon meddig vagyunk képesek erre a kategória-finomhangolásra, és meddig vagyunk fogékonyak az új beszédhangok elsajátítására? Vajon van biológiai határa az akcentusmentes észlelés és ejtés elsajátításának? LENNEBERG (1967) **kritikus periódus hipotézise** szerint kétéves kortól serdülőkorig a nyelvelsajátítás természetesen és különösebb energiabefektetés nélkül megy végbe, azonban a serdülőkor (pontosabban az agyféltekék közötti funkciómegoszlás, a lateralizáció) lezárultával az emberek biológiailag már nem fogékonyak, és ezért nem lehetséges az L2-t akcentusmentesen elsajátítani. A kritikus periódusra, illetve annak életkori, avagy biológiai határait számos más kutatás is tett javaslatokat (pl. PATKOWSKY 1990; SCOVEL 1988). Azonban a kritikus periódus hipotézissel ellentmondó eredményeket mutat fel FLEGE (1995), aki kutatásával azt erősíti meg, hogy a beszédhangok percepciója és produkciója az egész életciklus alatt adaptívan képes változni, ugyanúgy, ahogyan új beszédhang-kategóriákat is létre tudunk hozni a már meglévő kategóriák megváltoztatásával. Ezért nem állíthatjuk, hogy a serdülőkor lezárultával „hirtelen” eltűnik az ember képessége az L2-hangzók elsajátítására. Ehelyett inkább lineáris kapcsolatot figyelhetünk meg, ahol az idegennyelvi akcentus egyre jellemzőbb a kor előrehaladtával (FLEGE 1999), hiszen az anyanyelvi fonémakategóriák a növekvő számú megvalósulás miatt egyre körvonalazottabbá és merevebbé válnak (SCOVEL 1981). A körvonalazottabb és merevebb kategóriák működését egy másik szempontból vizsgálva meg, az L2-beszédhangok észlelését egy L1-alapú torzításként is megfogalmazhatjuk, melyben az L1 beszédhangjai „mágnesként” húzzák magukhoz az L2-beszédhangokat, és csökkentik a perceptuális távolságot az L2-beszédhang és a „mágneses” anyanyelvi beszédhang között (KUHL 1991; IVERSON–KUHL 1995).

Habár a beszédpercepcióban az akusztikai jel beszédhangként való kategorizálása közvetlen műveletnek tűnik, BROWN (1998) szerint ezt a folyamatot mégis két szintre kell bontanunk. Először is a beszédhangok percepciója esetében az akusztikai jelet általános akusztikai jellemzői alapján univerzális (nem nyelvspecifikus) kategóriákba soroljuk (pl. zárhang vagy elől képzett magánhangzó), ami minden ember nyelvelsajátítási folyamatában automatikusan működik. Majd ezeket az univerzális kategóriákat egy felsőbb, absztrakt szinten feleltetjük meg az illeszkedő nyelvspecifikus fonémakategóriák reprezentációjának (4. ábra).



4. ábra

*A beszédészlelés kétszintű modellje (BROWN 1998: 148 nyomán)*

A beszédpercepció kétszintes rendszerét MIYAWAKI és munkatársai (1975) kutatásainak eredményeiben is megfigyelhetjük, ahol japán anyanyelvűekkel végeztek diszkriminációs vizsgálatokat. A japánok számára a [r] és a [l] ugyanazon fonéma kontextusfüggetlen változatainak, szabad alternánsainak tekintendő, így a /ra/ és a /la/ angol szótagok megkülönböztetésekor nehézségeket feltételezhetünk. Amikor a kísérlet japán résztvevőinek a /ra/ és a /la/ szótagokat pusztán egy akusztikai jellemzőjük ( $F_3$ -átmenet) alapján kellett megkülönböztetniük, akkor jól teljesítettek, ha azonban beszédkontextusban (szótagokba ágyazva) kellett megkülönböztetni a két elemet, akkor már a nyelvspecifikus tudás nemcsak kihívások elé állította, hanem akadályozta is a diszkriminációt. (Ugyanezt a nem nyelvspecifikus akusztikai természetű feldolgozást relatíve rövid ingerközi idővel is el lehet érni: ha a két inger közötti időintervallum megfelelően rövid, akkor „nem jut idő” a nyelvspecifikus kategorizálásra [WERKER–LOGAN 1985].)

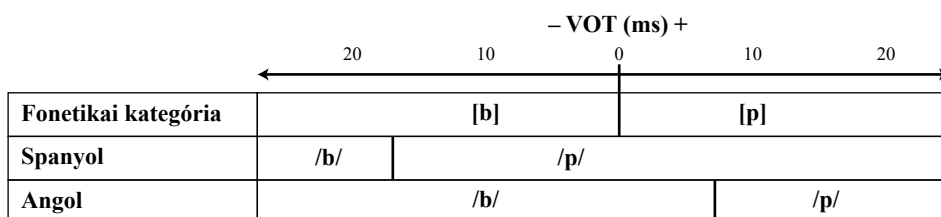
Összegezve, a nyelvi tapasztalat gyarapodásával az univerzális akusztikai diszkriminációból egy nyelvspecifikus szűrőn átszűrt diszkrimináció jön létre, ami az L1-kategóriákkal és -kategóriahatárokkal van behangolva, és elsődlegesen az L1-ben megtalálható szembenállásokat tartja fókuszban. Ha egy új nyelvet kezdünk tanulni, ott a szembenállások és beszédhang-kategóriák sok esetben az anyanyelvtől eltérően valósulnak meg, illetve emellett a szótagszerkezeti és szupraszegmentális tényezők sem mindig egyeznek. Mivel a felnőttek „már jól megszokott” és kikristályosodott L1-feldolgozási mechanizmusokkal rendelkeznek, és a transzfer révén ezt a megszokott L1-rendszert kell „újrahangolniuk”, ezért nehézségekbe ütközik az L2-re jellemző új feldolgozási procedúrák kiépülése. Tömören ezzel a jelenséggel magyarázhatjuk az akcentus kialakulását.

## 4. Transzferjelenségek

A következőkben a nyelvek közötti eltérésekből fakadó szegmentális és szupraszegmentális (transzfer)jelenségekkel, valamint ezek percepciójával és produkciójával foglalkozunk, valamint ezekre mutatunk be egy-egy jellegzetes példát. Ezekben a példákban az L1 transzfere meghatározó, hiszen kihat mind a szegmentális, mind a szupraszegmentális szintre (FLEGE 1988; MAJOR 2001; ECKMAN 2004).

Segmentális szinten a mássalhangzóknál, főleg a zárhangoknál és zár-rés hangoknál viszonylag egyszerű empirikus mérésekkel pontosan meghatározni, hogy az adott zárhang fonetikailag milyen zöngésségi kategóriába sorolható, vagyis hogy zöngés vagy zöngétlen. E beszédhangok esetében **zöngelkedési időt** (*voice onset time*, **VOT**) mérnek (LISKER–ABRAMSON 1964). A VOT erősen összefügg azzal, hogy a zárhangok és zár-rés hangok képzése során a beszédképző szervek akadályt, zárat képeznek a toldalékcsovében. A /b/ és a /p/ hang esetében ezt a zárat az ajkak képezik, és a zár feloldásakor hallhatjuk a beszédhangokra jellemző „pukkanást”. A zöngelkedési idő mérésekor a zárfeloldást tekintjük referenciapontnak, és azt nézzük meg, hogy a hangszalagok

rezgése a zárfeloldás előtt vagy után történik meg. Eszerint két fonetikai kategóriát tudunk megkülönböztetni: amennyiben a zárfeloldás előtt már rezegnek a hangszalagok (–VOT, azaz negatív a VOT), akkor zöngés hangról; ha a zárfeloldáshoz képest csak relatíve később indul meg a zöngé (+VOT, azaz pozitív a VOT), akkor zöngétlen beszédhangról beszélhetünk. Azonban a nyelvek általánosságban nem a zárfelpattanást tekintik referenciapontnak és „kategóriahatárnak” a bennük zöngésként és zöngétlenként számontartott beszédhangok között. Például LISKER és ABRAMSON (1970) az előbb említett /b/ és /p/ bilabiális zárhangok percepcióját vizsgálta angol és spanyol anyanyelvűeknél. A diszkriminációs és identifikációs kísérletben a résztvevőknek a VOT-kontinuumon különböző pozitív, illetve negatív VOT-értékű mesterséges stimulusokat kellett /b/-nek vagy /p/-nek kategorizálniuk (vö. 5. ábra).



5. ábra

*A spanyol és az angol zöngés-zöngétlen bilabiális zárhangok perceptuális kategóriahatára a VOT (ms) alapján (LISKER–ABRAMSON 1970 és ESCUDERO 2005: 18 nyomán)*

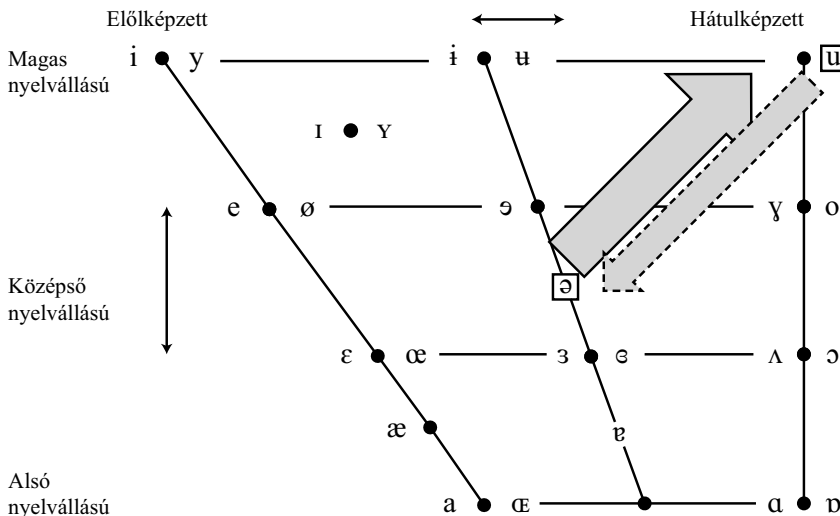
Mint ahogy az 5. ábrán látható, az akusztikai információkat az anyanyelvi minta alapján soroljuk fonémakategóriákba, azaz az akusztikailag megegyező inger más-más fonémareprezentációnak felelhet meg eltérő anyanyelvű személyek esetében. Például a spanyol anyanyelvűek a –5 ms-os VOT-értéket, azaz fonetikai szempontból [p] beszédhangot már zöngétlen /p/-ként észlelik, míg az angol anyanyelvűek számára ez még éppen a „zöngés” /b/ fonémakategóriájába sorolandó inger.

Amit még a kategóriahatárok észlelésével kapcsolatban tudni érdemes, az az, hogy a mássalhangzók esetében (melyek produkciójára sokkal kevésbé van hatással a beszédhang-környezet) **katégorikus észlelést** feltételezünk, azaz például egy VOT-kontinuumban az angol anyanyelvűek zöngés-zöngétlen kategóriái jól elkülönülnek, nincs igazán átmenet az észlelésben. Ez azt jelenti, hogy az akusztikai kontinuumot nem fokozatosan változóan észleljük (mint amilyen valójában), hanem a fonémahatárnál az észlelés hirtelen zöngésből zöngétlenbe „csap át”. (Szemben a magánhangzók észlelésével, ahol a sokkal változatosabb és kontextus-függőbb megvalósulások miatt sokkal kevésbé nyilvánvalóak a beszédhang-kategóriák közötti határok [FRY et al. 1963; PISONI 1973].) A katégorikus észleléssel kapcsolatban még azt érdemes megemlíteni, hogy sokkal könnyebb olyan ingereket megkülönböztetni egymástól, amelyek a kategóriahatár eltérő két oldalán helyezkednek el, mint két olyan ingert, amelyek ugyanazon kategórián belül (IVERSON–KUHL 2000). Például ha angol anyanyelvűekkel készítünk diszkriminációs kísérletet, akiknek a fonémahatáruk +5 VOT ms-nál van, akkor ők könnyebben és nagyobb eséllyel megkülönböztetik a –10 VOT ms-os ingert a +20 VOT ms-os ingertől, mint

a számukra ugyanazon kategóriába tartozó +20 VOT ms-os ingert a +50 VOT ms-os ingertől, annak dacára, hogy a két-két inger azonos „távolságra” van egymástól a VOT-kontinuumon. Ha viszont ugyanezeket az ingereket a spanyol anyanyelvűek szemszögéből nézzük, akiknél ugyanaz a határ -17 VOT ms-ra tehető, nekik egyik ingerpár megkülönböztetése sem lesz egyszerű, mivel a két-két ingerpár mind ugyanazon fonémakategóriába esik számukra. Ezáltal például az L1 és az L2 VOT-kategóriáinak viszonyai alapján következtethetünk arra, hogy az L2-tanulók milyen percepciósi nehézségekkel fognak küzdeni.

Az L1-ben és az L2-ben eltérő kategóriahatárokhöz kapcsolódó eredményeket mutat be L2-magánhangzók percepcióját illetően ROCHET (1995) kísérlete is, habár az alábbi esetben egy addig a nyelvtanulók számára ismeretlen, pontosabban a nyelvekben nem önálló fonémaként létező beszédhangot kellett kategorizálni. Az angol és a portugál nyelvben csak két felső nyelvállású magánhangzó-fonéma van jelen, az /i/ és az /u/. A franciában azonban felső nyelvállású az /y/ is, melyet a portugálok /i/-ként, míg az angolok /u/-ként kategorizáltak. Ebben az esetben ugyanazt a jelenséget figyelhetjük meg, mint a VOT-értékeknél, csak itt a hátul képzettséget számszerűsítő második formáns ( $F_2$ ) értékének „felosztása” tér el a portugálok és az angolok között a nyelvspecifikus észlelésükből fakadóan, és ezért sorolják eltérő fonémakategóriákba a francia /y/-t.

A magánhangzók percepcióját érintően meghatározó modellt a POLKA- és BOHN-féle (2011) *Natural Referent Vowel Framework*, azaz a magánhangzók által meghatározott természetes referenciakeret az észlelésben. A kutatás egy olyan perceptuális aszimmetriajelenséget vizsgált, mely során különbséget találtak a gyermekek és a felnőttek között a magánhangzó-kategóriák észlelésében. A magánhangzókat képzési tulajdonságaik szerint egy kétdimenziós trapéz alakjában tudjuk szemléltetni, amelynek a dimenzióit vízszintesen az elől és hátul képzetség, függőlegesen pedig a nyelvállásfok jellemzi (6. ábra).

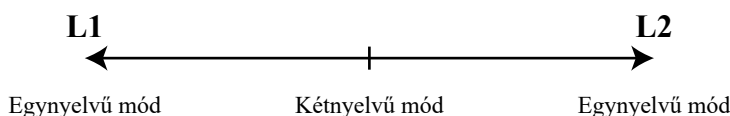


6. ábra

A magánhangzóter két-dimenziós ábrája (WIELING–MARGARETHA–NERBONNE 2011: 116 nyomán)

POLKA és BOHN kísérlete szerint, ha egy magánhangzó-kontinuumot észlelünk, amely kontinuum akusztikailag és artikulációs szempontból is a magánhangzótér periferiájától a magánhangzótér közepe felé halad, úgy sokkal nehezebb észlelni a változást, mint fordított esetben, ha a magánhangzótér közepétől indulva, a periferiális magánhangzók felé haladva kellene kategorizálni. A 6. ábrán is látható példákkal szemléltetve, ha a magánhangzótér közepén elhelyezkedő semleges svától [ə] lépkedünk a trapéz periferiáján található [u] beszédhang felé, akkor sokkal nyilvánvalóbban észleljük a kontinuumbeli változást, mint fordítva, ha az [u]-tól haladnánk a trapéz közepe felé a sváig [ə]. Ez alapján POLKA és BOHN azt feltételezte, hogy a magánhangzótér szélén elhelyezkedő magánhangzók percepciósi horgonyként szolgálnak a nyelvsajátításban, és habár az L1-kategóriák kialakulásával és az életkor előrehaladtával ezek a perceptuális aszimmetriák elhalványulnak, azonban például felnőttkorban új, L2-beli szegmentumokkal, kontrasztokkal találkozva újra felszínre kerülnek, és jelentős szereppel bírnak az L2-beszédhangok észlelésében. POLKA és BOHN (2011) e perceptuális aszimmetriára hivatkozik SEBASTIAÁN-GALLÉS, ECHEVERRIA és BOSCH (2005) kutatásának percepciósi eredményeit illetően is. Ebben a kísérletben katalán–spanyol kétnyelvűeknek kellett hibásan ejtett szavakat detektálniuk, mely szavakban (a csak katalánban jelen lévő) /ε/ és (a katalánban és a spanyolban is jelen lévő) /e/ magánhangzókat cserélték fel egymással. Amennyiben az akusztikai ingerben a helyes /e/-t cserélték le a magánhangzótér középső részéhez közelebb eső /ε/-re, akkor a kétnyelvűek kevésbé észlelték a hibás ejtést, viszont fordított irányban, amikor a magánhangzótérben „beljebb” elhelyezkedő /ε/-t cserélték a jobban periferiára eső /e/-re, akkor a kísérleti személyek szinte mindig detektálták az eltérést.

Arra vonatkoztatva, hogy milyen környezeti tényezők befolyásolják az akcentusos percepció, illetve produkció kialakulását (például a korábban említett VOT-kategóriahatárokat illetően) érdemes néhány szót ejtenünk a nyelvi módokról. GROSJEAN (2001) **nyelvi mód hipotézise** (*Language Mode Hypothesis*) szerint a kétnyelvű személy (itt: aki rendszeresen használja mind az L1-et, mind az L2-jét) adott időpillanatban a környező pszichológiai és nyelvi tényezőktől függően, eltérő szinteken aktiválhatja az L1-et és az L2-t az elméjében (GROSJEAN 2001: 3). Más szóval a nyelvi módot egyfajta kontinuumként foghatjuk fel, mely két abszolút végponttal rendelkezik, ezek az egynyelvű módok (legyenek bár az L1-re vagy az L2-re vonatkoztatva), és a két végpont között helyezkedik el a kétnyelvű mód (7. ábra). Az egynyelvű módokban csak az egyik (az L1 vagy az L2) aktiválódik, és a másik nyelv deaktivált állapotban van jelen. A kétnyelvű mód esetében a bilingvis mindkét nyelve aktív, azonban az alapnyelv, amely a kommunikáció fő nyelve (itt: L1), aktívabb az L2-nél.



7. ábra

A GROSJEAN (2001) által javasolt nyelvi módok (ESCUADERO 2005: 119 nyomán)

A nyelvi módok, azaz a két nyelvben eltérő aktivációs szintek a külső, környezeti nyelvi ingerek függvényében változnak, és befolyásolják a percepciót. Ilyen környezeti inger lehet a kísérletvezető által használt nyelv(változat), a kísérlet anyaga, illetve a megjelenített ortográfia, a kísérlet célja (ha sejthető), valamint a helyzet formalitása stb. Például CARAMAZZA és munkatársai (1973) kutatásában angol–francia kétnyelvűeknél szintetikus VOT-mintákkal vizsgálták a zárhangok zöngéességét. A megfelelő nyelvi aktivitást egy kezdeti produkciós feladattal, valamint a kísérletvezető személyével és az instrukció nyelvvel próbálták kontrollálni, azaz az angol ingerek esetében angolul, a francia ingerek esetében franciául beszéltek a kísérleti személyekhez. Arra számítottak, hogy a bilingvisek percepciós kategóriahatára mindkét nyelv esetében az anyanyelvi kategóriahatárral fog egyezni, azonban a két nyelv együttes aktivációja révén az angol és a francia kategóriahatárhoz viszonyítva mégis köztes eredményeket kaptak. Ezt az eredményt ELMAN, DIEHL és BUCHWALD (1977) kutatásában azzal magyarázta, hogy a szintetikus és jelentéssel nem bíró ingerek nem alkalmasak az egyelvű mód fenntartására. ELMAN és munkatársai (1977) a saját kutatásukban a két nyelv együttes aktivációját úgy küszöbölték ki, hogy az előbb említett szintetikus és jelentéssel nem bíró ingerek helyett természetes, önálló jelentéssel rendelkező szavakat használtak. Másfelől a kísérleti személyek minden inger előtt „orientáló” mondatokat is hallhattak az éppen vizsgált nyelven. Meg kell jegyeznünk, hogy a jövőbeli (idegen) nyelvi kutatások szempontjából is lényeges figyelembe venni, hogy a kísérletek közben alkalmazott nyelv(változat) hatással van az észlelésre, és adott esetben a produkcióra is, ezáltal befolyásolhatja a kísérlet eredményét. Emellett, amennyiben lehetőség van rá, törekedjünk természetes és jelentéssel bíró kísérleti anyagot vagy ingereket összeállítani (ELMAN–DIEHL–BUCHWALD 1977).

Egy másik jellegzetes jelenség, ami nyelvspecifikusan elferdítheti a percepciót, az az adott nyelvre jellemző szótagszerkezet és **fonotaktika** (azaz a beszédhangok sorba rendezésének lehetséges variációi). A szótagszerkezetek esetében a legtöbb nyelvben univerzálisan megjelenő CV (azaz mássalhangzó + magánhangzó szerkezetű nyílt szótag) tekinthető alapvetőnek. Ezt a CV szerkezetet további mássalhangzókkal bővítve egyre jelöltebb (és ezzel egyre nehezebben elsajátítható) szótagokat kapunk (CARLISLE 2001). Ezen szótagok esetében gyakran jelentkezik az L1-transzfer hatása, azaz egy magánhangzó betoldása (MAJOR 2001). Az univerzálisnak számító CV szerkezetű nyílt szótag a japán nyelvben is meghatározó. Az L1-beli szótagszerkezet hatását DUPOUX és munkatársai (2001) következő kísérletével is szemléltethetjük: a japán nyelvben a \*/mikdo/ és \*/sokdo/ helytelen hangsorok, azonban a /mikado/ és /sokudo/ valóban létező szavak. (A \* a helytelen, nem létező formákat jelöli.) A percepciós tesztben a japán anyanyelvűek a \*/kd/ mássalhangzó-torlódásokat mindkét esetben egy /u/ betoldásával, azaz /mikudo/-ként és /sokudo/-ként észlelték. Azaz ebből egyrészt arra következtethetünk, hogy az anyanyelvi szótagszerkezet és fonotaktika valóban hatással van az észlelésre, másfelől a beszédészlelés nem képeződik le közvetlenül a lexikai szintre (mivel akkor a \*/mikudo/ helyett egy tényleges japán szót, azaz a /mikado/-t észleltek volna). BERENT és munkatársai (2007) ezt a jelenséget a következő percepciós folyamattal magyarázták: miután az akusztikai jelet az univerzális fonetikai kategóriákba soroltuk, a nyelvspecifikus fonológiai reprezentáció szintjén fonotaktikai ellenőrzés



történik, hogy az észlelt hangsor helyesnek tekinthető-e. Amennyiben a hangsor helytelennek bizonyul, a fonológiai reprezentációban az anyanyelvi mintának megfelelő hangsor alakul ki. Azaz míg a hangsorban ténylegesen jelen lévő beszédhang-szekvenciák már fonetikai szinten meghatározásra kerülnek, addig a betoldott magánhangzók, amelyek a fonotaktikai javítás eredményei, már a fonológiai, nyelvspecifikus szinten lesznek kódolva. Ugyanígy, az L1-beli szótag-szerkezet révén például a spanyol nyelvben a \*/sC/ (azaz /s/ + mássalhangzó) szó(tag)kezdet helytelennek számít, ezért a spanyol anyanyelvűek az L1-beli fonotaktikai szabály megsértését egy /e/ betoldásával oldják fel, azaz \*/sC/ → /esC/. Ami azt jelenti, hogy az angol *snoob* /snob/ szót /esnob/-ként ejtik (GUEVARA–RUKOZ 2018). Ugyanígy a magyarul tanuló japánok a *meg kell szokni* [mæk:el sokəni] kifejezésben a japán nyelvben nem létező \*/kn/ mássalhangzó-torlódást egy svá [ə] beszúrásával oldották fel. Az olasz, valamint a mandarin kínai anyanyelvűek számára a szótagvégi mássalhangzó jólformáltság tekintetében ellentmond az anyanyelvük fonotaktikai szabályainak, ezért ők a szó végére toldanak be egy svát, pl. az olasz ejtésben az *élek* ['e:lekə] és az *akkor* ['ɒ:k:orə] (HUSZTHY 2019: 149), valamint a kínaiak esetében az angol *hot* /hɒt/ és *good* /ɡud/ szavak /hɒtə/, illetve /ɡudə/ ejtésénél (HAN 2013).

Ugyanakkor a mássalhangzó-torlódásokat nemcsak egy magánhangzó betoldásával lehet kiküszöbölni, hanem egy-egy mássalhangzó törlésével is. Például a vietnami nyelvre nem jellemző a mássalhangzó-torlódás, ezért az angol *first* [fɜ:st] helyett [fɜ:t], illetve a *grow* [ɡrɔʷ] helyett [ɡoʷ]-t produkálnak. Ugyanakkor megeshet, hogy a mássalhangzók szótag végén egytől egyig törlésre kerülnek, pl. *last* [læst] helyett [læ]-t ejtenek (SATO 1984).

A fonotaktikai szabályszerűségek mellett a más természetű fonológiai szabályok transzferét is meg kell említenünk mint akcentust kialakító tényezőt, habár ez a jelenség nem közvetlenül áll kapcsolatban a percepcióval (ECKMAN 1981; FLEGE 1986; MAJOR 2001). Nézzük meg a magyar és az angol nyelv példáját, a zöngésségi hasonulást tekintve, amely mindkét nyelv hangzásában fellelhető. Ha a magyar nyelvben két zörejmássalhangzó követi egymást, amelyek zöngésségi minőségükben eltérnek (itt az alábbi példában zöngétlent követ zöngés), akkor regresszív (visszafelé ható) zöngésségi hasonulás történik, azaz a mássalhangzó-kapcsolat második tagja határozza meg az előtte álló mássalhangzó zöngésségi minőségét. Egyszerűen a példában zöngésedés történik. Ezzel szemben az angolban, ha zöngés mássalhangzó követ zöngétlent, akkor a magyarral mind irányában, mind minőségében ellentétes a zöngésségi hasonulás: progresszív, azaz előrefelé ható zöngétlenedés történik. Tehát míg az angolok a *matchbox* [mætʃbɒks] szót zöngétlenedett [b] -vel mondják, addig ezt a szót magyarul vagy „magyarosan” [medzɒks]-nak, azaz zöngés [dʒ]-vel ejtjük (BALOGNÉ BÉRCES–SZENTGYÖRGYI 2006). Megjegyzendő, hogy ebben a konkrét esetben a szemléltetésre szánt idegen szavunkat már a magyar nyelv részeként is elfogadjuk, ezért természetesnek mondható, hogy a magyar fonológiai szabályok szerint ejtjük ki. Ha azonban ezeket az anyanyelvi mintázatokat átvisszük az idegen nyelvi kiejtésbe, az is okozhat akcentusos ejtést (ECKMAN 1981).

Ha feltételezzük, hogy képesek vagyunk a célnyelvi minta szerint ejteni az L2 beszédhangjait, valamint megfelelően követjük az L2 fonológiai és fonotaktikai szabályait, még így sem lehetünk biztosak abban, hogy nem beszélünk akcentusosan. Hiszen a nyelvek szuprasegmentális, avagy prozódiai jellemzőik terén is eltérnek egymástól, és a prozódiai jegyek

transzfere is jellemző az idegen nyelv elsajátítására (FLEGE 1986; ECKMAN 2004). Ilyen prozódiai jellemző például az alaphangmagasság ( $f_0$ ) változása is. **Intonációs**, avagy **atonális nyelvek** esetében (mint amilyen a magyar is) az alaphangmagasság változása hosszabb, közlés- vagy mondatszintű egységekhez társul. **Tonális nyelvek** esetében azonban a szavakhozszótagokhoz rendelt eltérő dallamkontúroknak jelentésmegkülönböztető szerepük van (például a kínai nyelvben). Az utóbbi, tónusos nyelvek elsajátításával kapcsolatban, például a mandarin kínaiban, a dallamkontúr mintázatának azonosítása és produkciója is okozhat problémákat (LI et al. 2020), és emellett a nyelvtanulók az alaphangmagasság regiszterében is eltérhetnek a célnyelvi ejtéstől. Például a mandarin beszélők sokkal mélyebb és magasabb alaphangtartományokat használnak a tónusok produkciójára, azaz tágabb hangterjedelemben valósítják meg a tónusokat, mint az angolok (SHEN 1989). Azonban az atonális nyelvek között is eltérő lehet az intonáció jellege, ami felismerhetővé teheti a beszélő akcentusát. HUSZTHY (2019) tanulmánya szerint míg a magyar anyanyelvűek számára az angol anyanyelvűek magyar beszéde intonáció szempontjából meglehetősen monotonnak érződött, addig az olasz beszélők dallamíveit pedig a magyarnál sokkal nagyobbban és erőteljesebben észlelték.

Egy másik szupraszegmentális tényező, amelyben a nyelvek eltérhetnek egymástól, a hangsúly. Míg a magyarban kötöten mindig a szavak első szótagja a hangsúlyos, addig az angol nyelvben (a magyartól eltérően) a több szótagú szavakban a hangsúly helye nem állandó. Ezért interferencia adódhat abból, ha az angol *importálni* igét [ɪm'pɔ:(r)t], aminek a második szótagjára esik a hangsúly, a magyar kötött szóhangsúly szerint [ˈɪmpɔ:(r)t] -nak mondjuk, ami az *import* szó főnévi jelentésére utal (a hangsúlyt a felső függőleges vonás [ˈ] jelöli) (PIUKOVICS–ÜSTÖKI 2019). Ritmikailag a nyelvek között szótag-időzítésű és hangsúly-időzítésű nyelveket különböztethetünk meg (valamint moraidőzítésűeket, mint a japán nyelv, de specifikussága miatt erről itt most nem ejtünk szót.) A **szótag-időzítésű nyelvekben**, mint a magyarban vagy a spanyolban, a hangsúlytól függetlenül minden szótag majdnem ugyanolyan hosszúságú. Ezzel szemben a **hangsúly-időzítésű nyelveknél**, mint amilyen az angol is, a hangsúlyos szótagok időtartamukban hosszabbak, a hangsúlytalanoknak pedig nemcsak az időtartamuk rövidebb, hanem magánhangzó-rövidülés is történik, ezért a ritmusuk is eltér a szótag-időzítésű nyelvekétől, hiszen itt a hangsúlyos szótagok adják az ütemet (MAJOR 2001). A nyelvek eltérő ritmusára hatással van az is, hogy a beszédhangok időtartamát jellemzi-e szembenállás, azaz megkülönböztetünk-e fonológiailag rövid és hosszú magánhangzókat és/vagy mássalhangzókat. A magyar nyelvre ilyen szempontból mind a magánhangzók, mind a mássalhangzók esetében jellemző a rövid-hosszú kontraszt (pl. *kor* [kor] és *kór* [ko:r], illetve *had* [hɒd] és *hadd* [hɒd:]). Az angolban viszont az esetleges rövid-hosszú variációkból nem következik jelentésbeli eltérés, hiszen azok csak allofonikus variánsok (pl. *see* [si:] > *seed* [si:d] > *seat* [si:t]) (JAKIELSKI–GILDERSLEEVE–NEUMANN 2018). Összegezve, ha az anyanyelvi ritmust és hangsúlyozást használjuk, valamint a beszédhangok időtartamának ejtésekor nem vesszük figyelembe az L2 fonológiai szabályszerűségeit, úgy könnyen eltérhetünk az adott nyelv normáitól, és válhat akcentusossá az ejtésünk. A szupraszegmentális tényezők jellegzetes voltát, illetve azt, hogy ezek mennyire áruklodók lehetnek az akcentusos beszédet illetően, MUNRO (1995) kísérletével tudjuk szemléltetni,

ahol angol anyanyelvűek és mandarin anyanyelvűek angol beszédét alul átérésztő szűrővel, szegmentális szempontból érthetetlenné torzították, csak a dallam, hangsúly és tempó jellemzőit hagyva meg a kísérleti anyagban. Azonban ezen szupraszegmentális jegyek alapján is felismerhető maradt az akcentusos beszéd, jelezve, hogy a beszédhangok mellett a prozódia szerepe sem elhanyagolható az L2 elsajátításában.

## 5. Modellek az L2 beszédhangjainak elsajátításával kapcsolatban

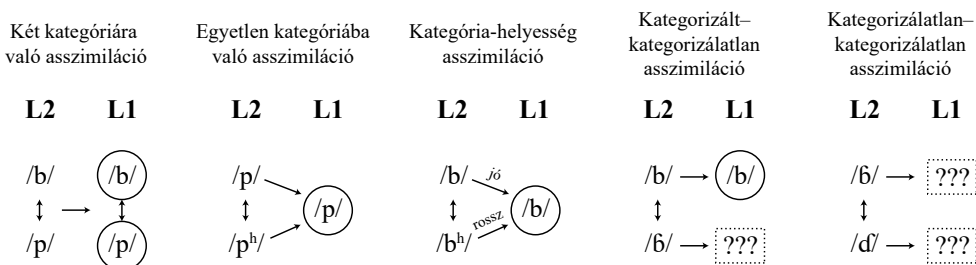
Ebben a részben két modell főbb téziseivel ismerkedünk meg, amelyek az L2 beszédhang-elsajátítási folyamatait vizsgálják a nyelvtanulás különböző fázisaiban. Ez a két modell FLEGE (1995) beszédelsajátítási modellje (*Speech Learning Model*), valamint BEST (1995) perceptuális asszimilációs modellje (*Perceptual Assimilation Model*). Mindkét bemutatott paradigma inkább fonetikailag, mint fonológiailag közelíti meg az L2-beszédhangok elsajátítását, azonban amíg FLEGE modellje az L1- és L2-beszédhangokat önállóan kezeli, és azt vizsgálja, hogy ezek fonetikailag mennyire hasonlítanak egymásra vagy térnek el, addig BEST modellje a beszédhangok közötti szembenállásokra összpontosít. A két modellben közös azonban, hogy egyik sem kérdőjelezi meg az L1 hatását az L2-szegmentumok percepciójára, azaz mindkettő a nyelvspecifikus észlelésből indul ki.

FLEGE (1995) **beszédelsajátítási modellje** (*Speech Learning Model*, a továbbiakban: SLM) az L2-beszédhangok percepcióját helyezi a középpontba, illetve a fonetikai jegyek megkülönböztetését az akusztikumban. A nyelvtanulás kezdeti szakaszában az L2-beszédhangok percepciójakor a nyelvtanuló az L2-beszédhangot a legközelebbi L1-beli pozicionális allofónhoz (azaz anyanyelvi beszédhanghoz) kapcsolja. Így e szerint a modell szerint az L1- és L2-beszédhangok közötti percepció hasonlóság és megfeleltetés pusztán fonetikai alapokon nyugszik, a fonémák szintjéhez nem kötődik, mivel annál kevésbé absztrakt szinten megy végbe (FLEGE 1995: 239). FLEGE (1987a, 1987b, 1995) az L2-beszédhangokat fonetikai szempontból három kategóriába sorolja: az L1-hez viszonyítva lehetnek újak, hasonlóak vagy azonosak. Amennyiben az L2-beszédhang az L1-beli megfelelőjével fonetikailag azonos, akkor a nyelvtanuló nem ütközik problémákba az elsajátítás során, azonban az új és a hasonló L2-beszédhangok esetében már nehézségekbe ütközhet. Az új L2-beszédhangok nem rendelkeznek közvetlen ekvivalenssel az L1-ben, azaz FLEGE (1987b: 48) szerint akusztikailag különböznek az L1-hangzóktól. Ugyanitt az új beszédhang meghatározására a fent bemutatott /y/ esetét hozza példának az angol anyanyelvűek számára. A hasonló L2-beszédhangok fonetikai szempontból különböznek az L1-től, azonban nem olyan mértékben, hogy ez az akusztikai különbség a nyelvtanuló számára kontrasztív legyen, hiszen az anyanyelvi észlelés kiszűri azokat az akusztikai jellemzőket, amelyek segítenék a diszkriminációt. Ilyen hasonló beszédhangnak tekinthető szintén a beszédhang-percepciónál tárgyalt árnyalatnyilag különböző

angol és spanyol /b/ és /p/. Azt a jelenséget, amikor az akusztikai különbségek dacára a nyelvtanuló az anyanyelvi megfelelőjével helyettesíti az L2-beszédhangot, **ekvivalensként való osztályozásnak** (*equivalence classification*) nevezzük (FLEGE 1987a, 1987b, 1988, 1995). A hasonló hangok problémás percepciójából, avagy ekvivalensként való osztályozásából következően FLEGE (1995) úgy véli, hogy az új L2-beszédhangokat sokkal könnyebb elsajátítani, mint az L1-hez hasonlókat. Ez a megállapítás párhuzamba állítható azzal a jelenséggel, amelyet RANSCHBURG Pál (1914/1988) **homogén gátlás** néven írt le: minél homogénebb ingerek érik egyszerre az ember tudatát, annál egységesebbnek és egyezőbbnek érezzük ezeket az ingereket, ezért kevésbé észleljük a közöttük fellelhető eltéréseket. A homogén gátlásként is leírható jelenséget FLEGE úgy fogalmazza meg, hogy minél kisebb az L1 és az L2 beszédhang közötti észlelt fonetikai távolság, azaz minél hasonlóbbak ezek a beszédhangok, annál jobban érvényesül a gátló hatás, ami megakadályozza, hogy észleljük a közöttük lévő különbséget; illetve minél távolabb vannak egymástól, annál valószínűbb, hogy új kategória jön létre. Azonban FLEGE nem ad pontos meghatározást arra vonatkozóan, hogy hogyan lehet mérni ezt a beszédhangok közötti észlelt fonetikai távolságot, valamint az L2-beszédhangok újként vagy hasonlóként való kategorizációja sem mindig tűnik egyértelműnek (ESCUDERO 2005). Ugyanakkor ez a modell nemcsak az L2-elsajátítás kezdeti, hanem a lezáró szakaszával is foglalkozik. Mint már említettük, FLEGE nevéhez fűződik egy, a kritikus periódus hipotézisnek ellentmondó kutatás is, miszerint az L2-elsajátítás és az új beszédhang-kategóriák létrehozása az egész életciklus alatt lehetséges (FLEGE 1995: 233), azonban a kategóriák kikristályosodásával (azaz a korrallal és a nyelvi tapasztalattal) egyre nehezebb ellenállni az ekvivalensként való osztályozásnak és az új beszédhang-kategóriák kialakításának.

BEST (1994, 1995) **perceptuális asszimilációs modellje** (*Perceptual Assimilation Model*, a továbbiakban: PAM) az L2-elsajátítás kezdeti szakaszát helyezi a középpontba, azon belül is a beszédhangok közötti – elsősorban nem anyanyelvi – szembenállásra összpontosítva. Ez a modell azért is számít különlegesnek, mert BEST szerint a felnőttek az artikulációs gesztusokból, azaz a képzési jegyek összességéből nyernek közvetlen információt a beszédhangokról, és az észlelés során nem alakul ki mentális reprezentáció (vö. BROWMAN–GOLDSTEIN [1989] artikulációs fonológiai elméletével). FLEGE SLM-jéhez hasonlóan ebben a modellben sem kérdőjeleződik meg a nyelvspecifikus észlelés, ami a percepció tanulás eredménye. Az anyanyelv-elsajátítás során a csecsemők kiszűrik az L1-ben kontrasztív, azaz jelentésmegkülönböztető jegyeket, ezáltal egyszerűbbé válik az észlelésük, azonban később, az L2 elsajátításakor a nem natív szembenállásokat is az észlelés ezen tulajdonsága gátolja. Ebből a folyamatból következően asszimilálják a nyelvtanulók az L1-től eltérő L2-beszédhangokat anyanyelvi beszédhangokkal. Az asszimiláció a PAM szerint perceptuális hasonlóságból fakad, amit a nyelvtanulók elsősorban artikulációs, és csak másodsorban akusztikai információ alapján nyernek ki. Ugyanakkor, mint azt már az előbb említettük, a PAM nem beszédhang és beszédhang közötti artikulációs és/vagy akusztikai természetű hasonlóságokkal foglalkozik, hanem az adott nyelvben jelen lévő szembenállásokkal. Ezért az asszimiláció típusait is az alapján határozza meg, hogy az L2-beli kontraszt mennyiben feleltethető meg valamilyen L1-beli szembenállásnak, és ezáltal valószínűsíthető-e a nem natív opozíció diszkriminációja és az új beszédhang-kategória létrehozása. A diszkrimináció

abban az esetben lesz a legpontosabb, ha két kategóriába való asszimiláció történik, azaz az L2-szembenállás natív szembenállásnak feleltethető meg. Ennél kevésbé sikeres diszkrimináció jósolható, ha a nyelvtanulók a szembenálló beszédhangokat egyetlen L1-kategóriába sorolják (egyetlen kategóriába való asszimiláció), de ez csak abban az esetben történhet, ha a két inger ugyanannyira jó vagy rossz megvalósulása az adott kategóriának. Ha a két inger egyike az adott L1-kategória jó, míg a másik kevésbé jó verziója, akkor kategória-helyesség szerinti asszimiláció történik, amire BEST az előbbieknél is pontatlanabb diszkriminációt előfeltételez. Illetve akkor, ha a beszédhangpárnak csak az egyik tagja, vagy esetleg egyik tagja sem asszimilálható egyetlen natív beszédhang-kategóriába sem, akkor az előbbi esetben kiváló, míg az utóbbi esetben az ingerektől függően kiváló vagy nagyon gyenge megkülönböztetést feltételez (kategorizált–kategorizálatlan és kategorizálatlan–kategorizálatlan asszimiláció). Az asszimilációtípusokat a 8. ábra mutatja be.



8. ábra

A PAM különböző asszimilációtípusai az L1 és az L2 viszonylatában (BEST 1995 nyomán)

Ugyanakkor a PAM az SLM-moddellel ellentétben az új beszédhang-kategória kialakítását tekinti nehezebb folyamatnak a nyelvtanuló számára. A nyelvi tapasztalat növekedésével azonban a tanulók képessé válnak több részre osztani az L1-kategóriákat, hogy ezáltal az L2-nek megfelelő, új kategóriákat hozzanak létre.

A tanulmány terjedelmi keretei miatt csak e két modell részletesebb bemutatására került sor, azonban további modelleket is meg kell említenünk. Az előbbieken bemutatott két fonetikai szemléletű modellhez hasonló KUHL (2000) *Native Magnet Model*-je, aki a már említett L1-beszédhangok „mágneses vonzerejével” magyarázza az L2-beszédhangok akcentusos természetét. A fonetikai szemlélettel szemben három másik jelentős modell fonológiailag közelíti meg az L2-beszédhangok percepcióját, ilyen MAJOR (2001) *Onthogeny Phylogeny Model*-je, valamint BROWN (1998, 2000) *Phonological Interference Model*-je. Az előbbi a generatív nyelvészetből merítkezve egy rendkívül általános és egyszerű, de dinamikus folyamatként mutatja be a közties nyelvben lejátszódó átalakulást az L2-ingerek hatására. Az utóbbi pedig a generatív nyelvészetnek egy nemlineáris irányzatából származó fogalmon, a jegygeometrián alapszik, melyben a beszédhangok képzési jegyei egy sajátos, belső hierarchikus struktúrát alkotnak, és e struktúrából következtetve lehet hipotéziseket felállítani az L2-beszédhangok elsajátítására

vonatkozóan. Illetve nem utolsósorban meg kell említenünk ESCUDERO (2005) *Second Language Linguistic Perception Model*-jét is, ami alapjaiban sok szempontból az összes eddigi modellt egybeforrasztja, azonban a modell alapjait az optimalitáselmélet (PRINCE–SMOLENSKY 1993/2004) szolgáltatja. Összességében egyik modell sem vonja kétségbe az L1 hatását az L2-beszédhangok percepciójára.

## 6. Egyéb tényezők az akcentusos beszéd kialakulásában

Az életkor mint meghatározó tényező szerepét már korábban említettük az L2 akcentusmentes ejtésében. Habár azt mondhatjuk, hogy az L2-beszédhangok elsajátításának nincs konkrét életkori határa, azonban a kor előrehaladtával és az L1-kategóriák megszilárdulásával ez a folyamat egyre nagyobb akadályokba ütközik. Ezt alátámasztva OYAMA (1982) kísérletében olasz anyanyelvű angolul tanulókat vizsgált, akiket az USA-ba érkezésükori életkoruk alapján három csoportra osztott. Akik 6–10 évesen kerültek L2-környezetbe, azok összességében kevésbé akcentusos ejtést produkáltak, mint a 11–15 évesen érkezettek, a csoportok közül pedig azoknak a beszéde volt a legakcentusosabb, akik 16–20 éves korban érkeztek az USA-ba. Összegezve, habár jól meghatározható biológiai életkori korlát nincs az L2-beszédhangok akcentusmentes elsajátítására, FLEGE (1995) eredményeihez hasonlóan itt is azt láthatjuk, hogy az életkor előrehaladása negatív hatással látszik lenni erre a folyamatra (PICKERING 2012).

Az L2 elsajátítását az is jelentősen befolyásolja, hogy a nyelvtanuló felé milyenek az elvárások, valamint az is, hogy van-e választási lehetőség a kommunikáció médiumát illetően. BURLING (1981) a Svédországba egy évre kivándorló gyerekek és felnőttek L2-előmenetelét összevetve azt a tendenciát fogalmazta meg, hogy a gyerekek a felnőtteknél lényegesen többet fejlődtek. Ugyanakkor ez abból is fakadt, hogy míg a felnőttek munka közben gyakran a könnyebbség kedvéért svédről angolra váltottak, addig a gyerekeknek az iskolában svédül kellett kifejezniük magukat. Más szóval az L2-környezetben töltött idő mellett meghatározó tényező, hogy az adott idő alatt mennyi L2-inger érte a nyelvtanulót.

Az L2-ingerek mennyisége mellett az ingerek változatossága is megfontolandó szempont az L2-beszédhangok helyes elsajátításában (HARDISON 2012). LOGAN, LIVELY és PISONI (1991) felnőtt japán anyanyelvűeket tréningezett az angol /r/ és /l/ azonosítására. A nyelvtanulókat két csoportba osztották: az egyik csoport 5 különböző beszélő ingereivel gyakorolhatott, azonban a célhangokból az ingerek csak szóeleji, illetve szón belüli megvalósulást tartalmaztak. Ezzel ellentétben a másik csoport mindössze egy beszélő ingereit hallhatta, azonban a célhangokból a szóeleji és szón belüli mellett még szóvégi megvalósulásokkal is találkozhatott. Ilyen formán a kísérlet célja az volt, hogy összehasonlítsák, vajon a szó(tag)beli pozíció szempontjából vagy a beszélők szempontjából változatosabb inger vezet hatékonyabb azonosításhoz. Az első esetben, ahol a nyelvtanulók több anyanyelvi beszélő /r/- és /l/-megvalósulásával találkozhattak, szignifikáns fejlődés

mutakozott a kiinduló teljesítményhez képest, azonban a második csoport esetében – ahol csak egy beszélő ingereivel gyakorolhattak, de többféle szótagbeli pozícióban – nem találtak szignifikáns fejlődést. Azaz konklúzióként levonhatjuk, hogy ha több különböző beszélő ingerei állnak rendelkezésre, akkor a nyelvtanulók hatékonyabban tudják általánosítani az új L2-beszédhangok közötti különbségeket. Azonban azt is meg kell jegyeznünk, hogy ha a változatosság mértéke negatív vagy pozitív irányban túlszárnyalja az optimálist, azaz ha a nyelvtanulót túl gazdag vagy éppen túl szegény ingerkörnyezet veszi körbe, az gátolhatja az L2-beszédhangok pontos elsajátítását (LEATHER 1990).

A beszédhangok percepciójában ugyanakkor nemcsak az akusztikai inger, hanem az artikuláció is nagy szerepet játszik, lásd például a McGurk-hatást (MCGURK–MACDONALD 1976), melyben az artikuláció vizuálisan torzíja az észlelt hang minőségét<sup>1</sup>. Ezért az L2-beli beszédhangok tanulásakor a nyelvtanulónak az akusztikai inger mellett az artikulációs mintázat is sokat segíthet az elsajátításban. Például HARDISON (2003) angolul tanuló japánokat és koreaiakat két csoportban vetett alá audiovizuális, illetve csak auditív tréningnek, a számukra nehézségeket okozó /r/ és /l/ tekintetében. Alapvetően mindkét tréningtípus fejlődést idézett elő a nyelvtanulók teljesítményében, azonban az audiovizuális tréning a kizárólag auditívnál jóval hatékonyabbnak mutatkozott. Ugyanakkor azt is meg kell jegyeznünk, hogy HAZAN és munkatársai (2005) szerint az audiovizuális tréningek csak akkor hatékonyak, ha a kérdéses képzési jegy szemmel jól látható a produkcióban (például ajak- vagy foghangoknál), egyéb esetekben nem.

Mindemellett az L1 használata is befolyásolja, hogy a nyelvtanuló beszéde mennyire bizonyul akcentusosnak. FLEGE, FRIEDA és NOZAWA (1997) szerint azok az olasz anyanyelvű angolul tanulók, akik saját bevallásuk szerint többet használták az anyanyelvüket, relatíve erősebb akcentussal beszéltek az angolt, mint az olaszt ritkábban használó társaik (az egyező L2-környezetbe való érkezés ellenére). MACKAY, MEADOR és FLEGE (2001) eredményei szerint azok a „korai” L2-tanulók, akik többet használták az anyanyelvüket az L2-környezetben, kevésbé teljesítettek jól az angol mássalhangzók azonosításában, mint azok az L2-tanulók, akik kevesebbet használták az L1-üket.

Az L2 elsajátításában és az akcentus kialakulásában meghatározó faktorok továbbá az egyéni különbségek is. Az akcentusmentes, anyanyelvi szintű L2-elsajátítás esetében általában nyelvi tehetségre vagy nyelvi készségre szokás hivatkozni. CARROLL (1993) a nyelvi készséget négy faktor mentén határozta meg, ezek közül az akcentus szempontjából talán a legfontosabb a fonémakódoló képesség, azaz, hogy az ismeretlen akusztikai ingereket a nyelvtanuló hogyan és mennyire képes feldolgozni. A másik lényeges faktor az induktív nyelvtanulási készség, ami az általános szabályszerűségek megfigyelésén alapszik, azaz a nyelvtanulónak képesnek kell lennie a példákban rendszerszerű következtetéseket levonni. A nyelvi készség mellett a motiváció sem elhanyagolható tényező az L2 és a helyes ejtés elsajátításában. A motiváció fakadhat többek között integrációs szándékból (az adott közösséghez szeretnénk tartozni)

---

<sup>1</sup> A McGurk-hatást (MCGURK–MACDONALD 1976), illetve az artikuláció vizuális ingereinek percepcióra gyakorolt hatását lásd az alábbi linken: <https://bit.ly/3viliwR> (A letöltés dátuma: 2021. február 13).

vagy instrumentális célból (az L2 anyanyelvszerű ismerete révén előnyökhöz szeretnének jutni) is. Emellett még az egyéni különbségeket illetően a szocioökonómiai státusból, illetve a tanulmányi háttérből következő eltérő lehetőségek is befolyásolhatják az akcentusos ejtést (SKEHAN 2012; USHIODA – DÖRNYEI 2012).

## FOGALMAK

*idegen nyelv; anyanyelv; akcentus; fonetika; fonológia; beszédhang; allofón; fonéma; beszédfeldolgozás; beszédpercepció; beszédmegértés; beszédprodukción; transzfer; negatív transzfer (interferencia); pozitív transzfer; kontrasztív analízis; tipológiai jelöltség; univerzális grammatika (UG); köztes nyelv (interlanguage); akusztikai/felismerési kulcs; homogén gátlás; szegmentális szerkezet; prozódiai/szupraszegmentális szerkezet; kritikus periódus hipotézis; zöngekezdesi idő (voice onset time, VOT); kategorikus észlelés; nyelvi mód; beszédelsajátítási modell; perceptuális asszimilációs modell; fonotaktika; intonációs/atonális nyelvek; tonális nyelvek; hangsúly-időzítésű nyelvek; szótag-időzítésű nyelvek; ekvivalensként való osztályozás*

## IRODALOM

- ALBERTI Gábor – MEDVE Anna 2006. *Generatív grammatikai gyakorlókönyv*. HEFOP Bölcsész Konzorcium, Budapest.
- ASLIN, Richard – PISONI, David B. 1980. Some developmental processes in speech perception. In YENI-KOMSHIAN, Grace H. – KAVANAGH, James F. – FERGUSON, Charles A. (eds): *Child Phonology – Perception*. Academic Press, New York. 67–96.
- BALOGNÉ BÉRCES, Katalin – SZENTGYÖRGYI, Szilárd (szerk.) 2006. *The Pronunciation of English*. Bölcsész Konzorcium, Budapest.
- BARDOVI-HARLIG, Kathleen – SPROUSE, Rex A. 2017. Negative versus positive transfer. In LIONTAS, John I. – DELLICARPINI, Margo (eds): *The TESOL Encyclopedia of English Language Teaching*. John Wiley and Sons Inc., Hoboken. 1–6.
- BERENT, Iris – STERIADE, Donca – LENNERTZ, Tracy – VAKNIN, Vered 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104/3. 591–630.
- BEST, Catherine T. 1994. The emergence of native-language phonological influences in infants: A perceptual assimilation model. In GOODMAN, James C. – NUSBAUM, Howard C. (eds): *The Development of Speech Perception: The transition from speech sounds to spoken words*. MIT Press, Cambridge, MA. 167–224.
- BEST, Catherine T. 1995. A direct realist view of cross-language speech perception. In STRANGE, Winifred (ed.): *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research*. York Press, Timonium. 171–203.
- BOERSMA, Paul – ESCUDERO, Paola – HAYES, Rachel 2003. Learning abstract phonological from auditory phonetic categories: An integrated model for the acquisition of language-specific sound categories. In SOLÉ, Maria-Josep – RECASENS, Daniel – ROMERO, Joaquin (eds):



- Proceedings of the 15th International Congress of Phonetic Sciences*. Causal Productions, Barcelona. 1013–1016.
- BORDEN, Gloria – GERBER, Adele – MILSARK, Gary 1983. Production and perception of the /r-/l/ contrast in Korean adults learning English. *Language Learning* 33/4. 499–526.
- BROWMAN, Catherina P. – GOLDSTEIN, Louis 1989. Articulatory gestures as phonological units. *Phonology* 6/2. 201–251.
- BROWN, Cynthia 1998. The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research* 14/2. 136–193.
- BROWN, Cynthia 2000. The interrelation between speech perception and phonological acquisition from infant to adult. In ARCHIBALD, John (ed.): *Second Language Acquisition and Linguistic Theory*. Blackwell, Oxford.
- BURLING, Robbins 1981. Social constraints on adult language learning. In WINITZ, Harris (ed.): *Native Language and Foreign Language Acquisition*. Annals of the New York Academy of Sciences 379. New York. 279–290.
- CARAMAZZA, Alfonso G. – YENI-KOMSHIAN, Grace H. – CARBONE, Ettore – ZURIF, Edgar B. 1973. The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America* 54/2. 421–428.
- CARLISLE, Robert S. 2001. Syllable structure universals and second language acquisition. *International Journal of English Studies* 1/1. 1–19.
- CARROLL, John B. 1993. *Human Cognitive Abilities*. Cambridge University Press, Cambridge.
- CHOMSKY, Noam 1981. *Lectures on Government and Binding*. Foris, Dordrecht.
- CHOMSKY, Noam 1986. *Knowledge of Language: Its nature, origin, and use*. Praeger, New York.
- COOK, Vivian 2006. Interlanguage, multi-competence and the problem of the 'second' language. *Rivista di Psicolinguistica Applicata* 6/3. 1–14.
- COOK, Vivian 2008. *Second Language Learning and Language Teaching*. Arnold, London.
- DUPOUX, Emmanuel – PALLIER, Christophe – KAKEHI, Kazuhiko – MEHLER, Jacques 2001. New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes* 16/5–6. 491–505.
- ECKMAN, Fred R. 1981. On the naturalness of interlanguage phonological rules. *Language Learning* 31/1. 195–216.
- ECKMAN, Fred R. 2004. From phonemic differences to constraint rankings: Research on second language phonology. *Studies in Second Language Acquisition* 26/4. 513–549.
- ECKMAN, Fred R. 2012. Second language phonology. In GASS, Susan M. – MACKEY, Alison (eds): *The Routledge Handbook of Second Language Acquisition*. Routledge, New York. 91–105.
- ELMAN, Jeffrey L. – DIEHL, Randy L. – BUCHWALD, Susan E. 1977. Perceptual switching in bilinguals. *The Journal of the Acoustical Society of America* 62. 971–977.

- ESCUDERO, Paola 2005. *Linguistic perception and second language acquisition: Explaining the attainment of Optimal Phonological Categorization*. PhD thesis. Utrecht University, Utrecht.
- FLEGE, James E. 1986. The production and perception of foreign language speech sounds. In WINITZ, Harris (ed.): *Human Communication and Its Disorders*. Vol. 2. Ablex Publishing, Norwood. 155–171.
- FLEGE, James E. 1987a. Effects of equivalence classification on the production of foreign language speech sounds. In JAMES, Allan – LEATHER, James (eds): *Sound Patterns in Second Language Acquisition*. Foris, Dordrecht. 9–39.
- FLEGE, James E. 1987b. The production of ”new” and ”similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15. 47–65.
- FLEGE, James E. 1988. The production and perception of foreign language speech sounds. In WINITZ, Harris (ed.): *Human Communication and Its Disorders: A review*. Ablex Publishing, Norwood. 224–401.
- FLEGE, James E. 1995. Second language speech learning: theory, findings, and problems. In STRANGE, Winifred (ed.): *Speech Perception and Linguistic Experience: Theoretical and methodological issues in cross-language speech research*. York Press, Timonium. 233–277.
- FLEGE, James E. 1999. Age of learning and second language speech. In BIRDSONG, David (ed.): *Second Language Acquisition and The Critical Period Hypothesis*. Lawrence Earlbaum, Mahwah. 101–131.
- FLEGE, James E. – FRIEDA, Elaina M. – NOZAWA, Takeshi 1997. Amount of native-language (L1) use affects pronunciation of an L2. *Journal of Phonetics* 25. 169–186.
- FLEGE, James E. – MACKAY, Ian R. A. – MEADOR, Diane 1999. Native Italian speakers’ perception and production of English vowels. *Journal of the Acoustical Society of America* 106/5. 2973–2987.
- FRY, Dennis B. – ABRAMSON, Arthur S. – EIMAS, Peter D. – LIBERMAN, Alvin M. 1963. The identification and discrimination of synthetic vowels. *Language and Speech* 5/4. 171–189.
- GRASSEGER, Hans 1991. Perception and production of Italian plosives by Austrian learners. In *Actes du XIIIème Congrès International des Sciences Phonétiques*. August 19–24., 1991. Université de Provence, Aix-en-Provence. 290–293.
- GROSJEAN, François 2001. The bilingual’s language modes. In NICOL, Janet L. (ed.): *Explaining Linguistics. One Mind, Two Languages: Bilingual Language Processing*. Blackwell, Oxford. 1–22.
- GUEVARA-RUKOZ, Adriana 2018. *Decoding perceptual vowel epenthesis: Experiments & modelling*. PhD thesis. Ecole Normale Supérieure, Párizs.
- HAN, Feifei 2013. Pronunciation problems of Chinese learners of English. *ORTESOL Journal* 30. 26–30.

- HARDISON, Debra M. 2003. Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics* 24/4. 495–522.
- HARDISON, Debra M. 2012. Second language speech perception: A cross-disciplinary perspective on challenges and accomplishments. In GASS, Susan M. – MACKEY, Alison (eds): *The Routledge Handbook of Second Language Acquisition*. Routledge, New York. 349–365.
- HAYES, Bruce 2009. *Introductory Phonology*. Wiley–Blackwell, Malden.
- HAZAN, Valerie – SENNEMA, Anke – IBA, Midori – FAULKNER, Andrew 2005. Effect of audio-visual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication* 47/3. 360–378.
- HUSZTHY Bálint 2019. Az idegen akcentus elméleti előnyei. *Általános Nyelvészeti Tanulmányok* 31. 137–165.
- ISELL, Daniel 2016. The perception-production link in L2 phonology. *MSU Working Papers in SLS* 7. 52–67.
- IVERSON, Paul – KUHL, Patricia K. 1995. Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America* 97/1. 553–562.
- IVERSON, Paul – KUHL, Patricia K. 2000. Perceptual magnet and phoneme boundary effects in speech perception: Do they arise from a common mechanism? *Perception and Psychophysics* 62. 874–886.
- JAKIELSKI, Kathy J. – GILDERSLEEVE-NEUMANN, Christina E. 2018. *Phonetic Science for Clinical Practice*. Plural Publishing, San Diego.
- JAMES, Allan R. 1986. Phonetic transfer and phonological explanation: Some theoretical and methodological issues. In KELLERMAN, Eric – SHARWOOD, Michael (eds): *Cross-Language Influence in Second Language Acquisition*. Pergamon Press, Oxford. 134–149.
- JIA, Gisela – STRANGE, Winifred – WU, Yanhong – COLLADO, Julissa – GUAN, Qi 2006. Perception and production of English vowels by Mandarin speakers: Age-related difference vary with amount of L2 exposure. *Journal of the Acoustical Society of America* 119/2. 1118–1130.
- KALIKOW, Daniel N. – SWETS, John A. 1972. Experiments with computer-controlled displays in second language learning. *IEEE Transactions on Audio and Electroacoustics* 20/1. 23–28.
- KASSAI Iлона 1995. Pszicho-szociolingvisztikai jegyzetek az akcentusról. *Általános Nyelvészeti Tanulmányok* 18. 103–115.
- KLEIN, Wolfgang 1986. *Second language acquisition*. Cambridge University Press, Cambridge.
- KOHÁRI Anna 2021. Hogyan beszélünk a gyerekekhez? In MARKÓ Alexandra (szerk.): *Tanulmányok a beszéttudomány alkalmazásainak köréből*. ELTE Eötvös Kiadó, Budapest. 175–199.

- KUHL, Patricia K. 1980. Perceptual constancy for speech-sound categories in early infancy. In YENI-KOMSHIAN, Grace H. – KAVANAGH, James F. – FERGUSON, Charles A. (eds): *Child Phonology – Perception*. Academic Press, New York. 41–66.
- KUHL, Patricia K. 1991. Human adults and human infants show a 'perceptual magnetic effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics* 50/2. 93–107.
- KUHL, Patricia K. 2000. A new view of language acquisition. *Proceedings of the National Academy of Sciences USA* 97/22. 11850–11857.
- LADO, Robert 1957. *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press, Ann Arbor.
- LEATHER, James 1990. Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In LEATHER, James – JAMES, Allan (eds): *New Sounds*. University of Amsterdam, Amsterdam. 72–97.
- LENNEBERG, Eric H. 1967. *The biological foundations of language*. Wiley and Sons, New York.
- LEVELT, Willem J. M. 1989. *Speaking. From Intention to Articulation*. The MIT Press, Cambridge.
- LI, Xiaoshi – LUO, Qian – LIU, Jie – RYU, Catherine 2020. Empirical Studies on L2 Mandarin Chinese Production: What Can We Learn from Them? *Researching and Teaching Chinese As a Foreign Language* 3/1. 23–49.
- LISKER, Leigh – ABRAMSON, Arthur 1964. A cross-language study of voicing in initial stops. *Acoustical Measurements* 20/3. 384–422.
- LISKER, Leigh – ABRAMSON, Arthur 1970. Discriminability along the voicing continuum: crosslanguage tests. In HÁLA, Bohuslav – ROMPORTL, Milan – JANOTA, Přemysl (eds): *Proceedings of the 6th International Congress of Phonetic Sciences*. Academia, Prague. 569–753.
- LOGAN, John S. – LIVELY, Scott E. – PISONI, David B. 1991. Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America* 89/2. 874–886.
- MACKAY, Ian R. A. – MEADOR, Diane – FLEGE, James E. 2001. The identification of English consonants by native speakers of Italian. *Phonetica* 58/1–2. 103–125.
- MAJOR, Roy 2001. *The Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*. Routledge, New York.
- MCMAHON, April 2002. *An Introduction to English Phonology*. Oxford University Press, Oxford.
- MCGURK, Harry – MACDONALD, John 1976. Hearing lips and seeing voices. *Nature* 264. 746–748.
- MCQUEEN, James M. 2004. Speech perception. In LAMBERTS, Koen – GOLDSTONE, Rob (eds): *The Handbook of Cognition*. Sage Publications, London. 255–275.

- MIYAWAKI, Kuniko – STRANGE, Winifred – VERBRUGGE, Robert R. – LIBERMAN, Alvin M. – JENKINS, James J. – FUJIMURA, Osamu 1975. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics* 18. 331–340.
- MUNRO, Murray J. 1995. Nonsegmental factors in foreign accent. *Studies in Second Language Acquisition* 17/1. 17–34.
- MUNRO, Murray J. – DERWING, Tracey M. 1998. The effects of speech rate on the comprehensibility of native and foreign-accented speech. *Language Learning* 48. 159–182.
- NEMSER, William 1971. Approximative systems of foreign language learners. *International Review of Applied Linguistics* 9/2. 115–123.
- NEUFELD, Gerald G. 1988. Phonological asymmetry in second language learning and performance. *Language Learning* 38/4. 531–559.
- ODLIN, Terence 1989. *Language Transfer*. Cambridge University Press, Cambridge.
- OYAMA, Susan 1982. The sensitive period for the acquisition of a nonnative phonological system. In KRASHEN, Stephen – SCARCELLA, Robin C. – LONG, Michael H. (eds): *Child–Adult Differences in Second Language Acquisition*. Newbury House, Rowley. 20–38.
- PATKOWSKY, Mark 1990. Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics* 11/1. 73–89.
- PICKERING, Lucy 2012. Second language speech production. In GASS, Susan M. – MACKEY, Alison (eds): *The Routledge Handbook of Second Language Acquisition*. Routledge, New York. 335–348.
- PISONI, David B. 1973. Auditory and memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics* 13. 253–260.
- PIUKOVICS Ágnes – ÜSTÖKI Klaudia 2019. Az angol szóhangsúlymintázatok észlelésének és produkciójának tényezői magyar nyelvtanulók esetében: egy próbakísérlet tanulsága. In É. Kiss Katalin – HEGEDŰS Attila – PINTÉR Lilla (szerk.): *Nyelvelmélet és kontaktológia*. PPKE BTK, Piliscsaba. 237–256.
- POLIVANOV, Evguenij D. 1931. La perception des sons d’une langue étrangère. *Travaux du Cercle Linguistique de Prague* 4. 79–96.
- POLKA, Linda – BOHN, Ocke-Schwen 2011. Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics* 39/4. 467–478.
- PRINCE, Alan – SMOLENSKY, Paul 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Technical Report, Rutgers University Center for Cognitive Science. Blackwell, Oxford.
- RANSCHBURG Pál 1914/1988. A homogén gátlás. *Pszichológia* 8. 280–296.
- ROCHET, Bernard L. 1995. Perception and production of second-language speech sounds by adults. In STRANGE, Winifred (ed.): *Speech Perception and Linguistic Experience*. York Press, Timonium. 379–410.
- SATO, Charlene J. 1984. Phonological processes in second language acquisition: Another look at interlanguage syllable structure. *Language Learning* 34/4. 43–57.

- SCOVEL, Thomas 1981. The recognition of foreign accents in English and its implications for psycholinguistic theories of language acquisition. In SAVARD, Jean G. – LAFORGE, Lorne (eds): *Proceedings of the 5th Congress of AILA*. University of Laval Press, Laval. 389–401.
- SCOVEL, Thomas 1988. *A Time to Speak: A Psycholinguistic Enquiry into the Critical Period for Human Speech*. Newbury House, Rowley.
- SEBASTIÁN-GALLÉS, Nuria – ECHEVERRÍA, Sagrario – BOSCH, Laura 2005. The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language* 52/2. 240–255.
- SELINKER, Larry 1972. Interlanguage. *International Review of Applied Linguistics* 10/1–4. 209–231.
- SHELDON, Amy – STRANGE, Winifred 1982. The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics* 3/3. 243–261.
- SHEN, Xiaonan S. 1989. Toward a register approach in teaching Chinese: The problems with rising tone. *Journal of the Chinese Language Teachers Association* 24/3. 27–47.
- SKEHAN, Peter 2012. Language aptitude. In GASS, Susan M. – MACKEY, Alison (eds): *The Routledge Handbook of Second Language Acquisition*. Routledge, New York. 381–395.
- STRANGE, Winifred 1995. Cross-language study of speech perception: A historical review. In STRANGE, Winifred (ed.): *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. York Press, Timonium. 3–45.
- TRUBETZKOY, Nikolai S. 1939. *Principles of Phonology*. Klincksieck, Paris.
- USHIODA, Ema – DÖRNYEI, Zoltán 2012. Motivation. In GASS, Susan M. – MACKEY, Alison (eds): *The Routledge Handbook of Second Language Acquisition*. Routledge, New York. 395–409.
- VALDMAN, Albert 1976. *Introduction to French Phonology and Morphology*. Newbury House, Rowley.
- WERKER, Janet F. – LOGAN, John S. 1985. Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* 37. 35–44.
- WIELING, Martijn – MARGARETHA, Eliza – NERBONNE, John 2011. Inducing phonetic distances from dialect variation. *Computational Linguistics in the Netherlands Journal* 1. 109–118.



# Hogyan beszélünk a gyerekekhez?

Kohári Anna

NYELVTUDOMÁNYI KUTATÓKÖZPONT

## 1. Mi is az a dajkanyelv?

Mindennapi élményünk, hogy nem ugyanúgy beszélünk, attól függően, hogy kihez intézzük éppen a szavainkat. Ezt a jelenséget, amely során a beszélő a hallgatóhoz vagy hallgatókhoz igazítja a beszédmódját, **beszédalkalmazkodásnak** nevezzük (PICKERING–GARROD 2006; BERNHOLD–GILES 2020). A felnőttek például hajlamosak megváltoztatni beszédüket, amint meglátnak egy csecsemőt vagy egy kisgyereket. Ez a regiszter, amelyben a felnőttek a gyerekekhez intézik szavaikat, **dajkanyelvként** ismert a szakirodalomban. A dajkanyelvi beszédmódnak számtalan nyelvészeti és fonetikai tulajdonságát is feltárták már, amelyekben eltér a felnőttekhez szóló beszédétől. A szülők például hajlamosak egyszerűbb szintaxist és szókincset használni, mikor gyerekeikhez beszélnek (SAINT-GEORGES et al. 2013). Ugyanakkor a beszédük hangzását, fonetikai jellemzőit is megváltoztatják, például jóval lassabban beszélnek a gyerekekhez, mint a felnőttekhez (KO 2012; NARAYAN–MCDERMOTT 2016). A jelen írás első részében a dajkanyelvi beszédmód ezen akusztikai tulajdonságait tekintjük át részletesen. De vajon mi a motivációja a felnőtteknek, hogy a szülő-gyerekek interakcióban másképp beszéljenek, mint a felnőttekkel történő társalgásokban, és vajon a dajkanyelvi beszéd akusztikai sajátosságai hogyan segítik a gyermekek nyelvelsajátítását? A második részben ezekre a kérdésekre keressük a választ.

## 2. A dajkanyelvi beszédmód akusztikai tulajdonságai

### 2.1. A dajkanyelvi beszédmód temporális szerveződése

A beszédnek, így a dajkanyelvi beszédmódnak is egyik meghatározó tulajdonsága az időbeli vagy másképpen **temporális szerveződése**. Az édesanyák például hajlamosak lassabban beszélni gyerekeikhez, mint a felnőttekhez. Ezt a jelenséget rengeteg nyelvben kimutatták már, többek között az amerikai angol, a holland, a japán, a koreai, a német, a tagalog és a tamil nyelvben (FERNALD–SIMON 1984; CRISTIÀ 2010; MARTIN et al. 2016; NARAYAN–MCDERMOTT 2016; HAN et al. 2018). Úgy tűnik tehát, hogy a dajkanyelvi beszéd lassabb tempója univerzálisnak tekinthető. De pontosan hogyan is lassítják a felnőttek a beszédüket a gyerekekhez



szóló beszédben? Ennek a kérdésnek a megválaszolásához meg kell különböztetnünk a tempó kétféle mérési módját. Az egyik esetben csak azoknak a beszédszakaszoknak az időtartamát vesszük figyelembe, amely alatt a beszélő ténylegesen mondott valamit. Megszámoljuk a beszédegységek (pl. beszédhangok vagy szótagok) számát, és elosztjuk a beszédszakaszok tényleges időtartamával. Ennek a műveletnek a kimenetét nevezzük **artikulációs tempónak**. A **beszédtempó** esetében viszont a beszédszakaszok időtartamához hozzászámítjuk a szünetek időtartamát is, és a beszédegységek számát ezzel az összesített értékkel osztjuk el (GÓSY 2004; KOHÁRI 2018). A dajkanyelvi beszédnek mind az artikulációs tempója, mind a beszédtempója lassabb a felnőttekhez szóló beszédhez képest, a szünetek időtartamát pedig hosszabbnak találták a gyerekekhez szóló beszédben (GRIESER–KUHLE 1988).

Egy nemrégiben megjelent tanulmány felvetette ugyanakkor a kérdést, hogy az artikulációs tempó vajon tényleg alacsonyabb-e a gyerekekhez szóló beszéd minden részében, vagy csak egyes szótagok vagy beszédhangok nyúlnak meg jobban bizonyos helyeken (MARTIN et al. 2016). A dajkanyelvről ugyanis ismert, hogy kisebb szótagszámú, szünettől szünetig tartó szakaszokra, kisebb prozódiai egységekre tagolódik, mint a felnőttekhez szóló beszéd (FERNALD et al. 1989; RÉGER 2002). A prozódiai egységek végén lévő hangzók, szótagok tipikusan hosszabbak, mint az egységek belsejében lévő megfelelőik, ezt a jelenséget hívjuk **frázisvégi nyúlásnak** (l. részletesen KOHÁRI 2018). Amennyiben a dajkanyelvben gyakoribb a frázisvég, gyakrabban jelentkezik a frázisvégi nyújtás az utolsó beszédhangokon, illetve szótagokon, azaz gyakrabban fordulnak elő időtartamában hosszabb szótagok a beszédfolyamban. Ez pedig magyarázhatja a dajkanyelv lassabb tempóját, hiszen a tempómérés során az összes szótag vagy beszédhang időtartamát figyelembe vesszük, tekintet nélkül arra, hogy frázisvégen állnak, vagy sem. MARTIN és munkatársai (2016) japán spontán beszédet vizsgálva azt találták, hogy a frázisbelseji szótagok nem hosszabbak dajkanyelvben a felnőttekhez szóló beszédhez képest. Eredményeik alapján arra a következtetésre jutottak, hogy a dajkanyelvben megjelenő rövidebb prozódiai egységek, azaz a gyakoribb tagolás jelenti elsősorban a két regiszter közti különbséget. Ők ugyanakkor spontán beszéden vizsgálták. KO és SODERSTROM (2013) a frázisbelseji szavakon is tapasztalt nyújtást gyerekekhez szóló felolvasásban. Attól függően, hogy milyen beszédmódot, anyagot vizsgáltak, eltérő eredményeket kaptak. A helyzetet tovább bonyolítja, hogy a funkciószavak (pl. névelők, segédigék) magánhangzóinak időtartamában nem találtak eltérést a két regiszterben (SWANSON–LEONARD–GANDOUR 1992). Az adatok tehát ellentmondásosak abban a tekintetben, hogy az édesanyák hogyan produkálják a lassabb tempót a dajkanyelvben a felnőttekhez szóló beszédhez képest. Habár nincs konszenzus arról, hogy a dajkanyelvi beszédmód időzítése pontosan hogyan szerveződik, azt egyértelműen állíthatjuk, hogy a felnőttek mérhetően lassabban beszélnek a gyerekekhez, mint a felnőttekhez.

Felmerül a kérdés, hogy vajon a dajkanyelv lassabb globális tempója a gyerekek hány éves koráig van jelen. Azt találták, hogy a dajkanyelv tempója a gyerek kétéves kora előtt már közelíti a felnőttekhez szóló beszéd tempójához (NARAYAN–MCDERMOTT 2016). Ugyanakkor hozzá kell tennünk, hogy a dajkanyelvi tempó felgyorsulása nem lineáris a gyerek korának előrehaladtával (KO 2012). Tehát nem igaz az, hogy minél idősebb a gyerek, az édesanya

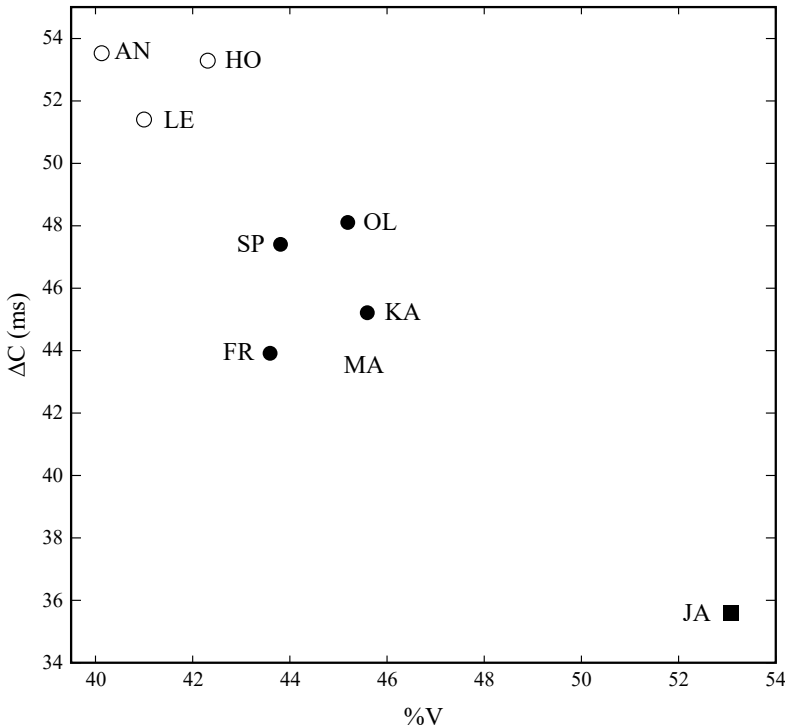
annál gyorsabban beszélne hozzá, ehelyett inkább tempóingadozások figyelhetők meg, de tendenciózusan az édesanyák egyre gyorsabban beszélnek gyerekekükhöz az idő múlásával. Fontos továbbá megjegyeznünk azt is, hogy a dajkanyelvben megjelenő tempóbeli lassítás mértéke nyelvfüggő, és az egyének között is igen nagy változatosság tapasztalható (KO 2012; NARAYAN–MCDERMOTT 2016), ugyanakkor a dajkanyelvnek egy olyan sajátosságáról van szó, amely a baba születésétől kezdve nagyjából kétéves koráig mérhető (LEE et al. 2014; NARAYAN–MCDERMOTT 2016).

A dajkanyelv temporális sajátságait a magyar nyelvben is vizsgálták már. Azt találták, hogy az édesanyák tendenciózusan lassabban beszélnek 4 és 8 hónapos korú csecsemőjükhez, mint a felnőttekhez. Újszülött gyerekek esetében viszont nem találtak különbséget a két regiszter között magyar nyelven, miközben például az ausztrál angolra és a thai nyelvre igen (THANAVISUTH–LUKSANEYANAWIN 1998; LEE et al. 2014). Az újszülöttekkel végzett vizsgálat esetében a sajátos kórházi környezet, illetve a körülmények, az édesanya kimerültsége is előidézhette a kapott eredményeket, ugyanakkor meg kell említenünk, hogy ezek a mérések automatikus módszerekkel készültek, amelyek kisebb eltérések esetén kevésbé tekinthetők hatékonyak. Nemcsak a tempót, hanem a frázisvégi nyúlást is megvizsgálták magyar anyanyelvű édesanyák dajkanyelvi beszédében. Azt találták, hogy a felolvasott mondatok utolsó szótagjainak abszolút időtartama hosszabb volt a dajkanyelvi beszédben, mint a felnőttekhez szólóban. A dajkanyelvi beszédről ugyanakkor ismert, hogy lassabb a felnőttekhez szóló beszédnél. Ez azt jelenti, hogy a szótagok időtartama átlagosan hosszabb a gyermekekhez szóló beszédben, mint a felnőttekhez szólóban. Továbbá azt is tudjuk, hogy a felnőttek közötti beszédben a megnyilatkozás végén az utolsó szótagok időtartama tipikusan hosszabb a megnyilatkozás belseji szótagokéhoz képest. Így felmerül a kérdés, hogy vajon a megnyilatkozásvégi szótagok a szótag belsejéhez viszonyítva is jobban nyúlnak dajkanyelvi beszéd-módban a másik regiszterhez képest, vagy csak ugyanolyan arányban hosszabbodnak. Az eredmények azt mutatták, hogy a megnyilatkozásvégi nyúlás aránya mindkét beszéd-módban ugyanakkora volt; a megnyilatkozásvégi szótagok átlagosan több mint 40%-kal bizonyultak hosszabbnak a megnyilatkozás többi részéhez képest. Ez azt jelenti, hogy az utolsó szótagok mindkét beszéd-módban jelöltek voltak az időtartamok tekintetében, azaz kitüntetett szerepet játszottak a többi, megnyilatkozás belseji szótaghoz képest. Az utolsó szótagok pedig arányaiban megközelítőleg ugyanannyira voltak jelöltek a két regiszterben. Az eredmények azt is megmutatták, hogy a gyermekekhez szóló beszéd az utolsó szótagok figyelmen kívül hagyásával is lassabb a felnőttekhez szóló beszédhez képest. Ezek alapján tehát úgy tűnik, hogy nemcsak a megnyilatkozások utolsó szótagjai miatt lassabb a gyerekekhez szóló beszéd, hanem más részei is hosszabbak, legalábbis felolvasott mondatokban (KOHÁRI és mtsai 2019; KOHÁRI et al. 2019). Nemcsak megnyilatkozások végeit, hanem kisebb prozódiai egységek határait és az ott megjelenő frázisvégi nyújtást is megvizsgálták magyar nyelvben, és hasonló eredményekre jutottak. A különböző prozódiai egységek határain előforduló szüneteket pedig nem találták gyakoribbnak vagy hosszabbnak a dajkanyelvben a felnőttekhez szóló beszédhez képest a magyar nyelv esetében (SZALONTAI et al. 2018).

## 2.2. Beszédrítmus

A dajkanyelvi beszéd ritmusa a tempóhoz képest kevésbé kutatott terület. Ez elsősorban arra vezethető vissza, hogy a beszédrítmus fogalma nehezen megragadható, nem egyértelmű, hogy milyen akusztikai paraméterek keltik a ritmus érzetét a hallgatóban. A beszédrítmus egyik megközelítése szerint a nyelvek különböző osztályokba sorolhatók, és megkülönböztethetünk szótag-időzítésű, hangsúly-időzítésű és moraidőzítésű nyelveket (PIKE 1945; ABERCROMBIE 1967; LADEFOGED 1975). Az elmélet szerint a szótag-időzítésű nyelvekben a szótagok közel azonos időtartamban realizálódnak, míg a hangsúly-időzítésű nyelvben a hangsúlytól hangsúlyig terjedő szakaszok időtartama lenne majdnem ugyanannyi (PIKE 1945; ABERCROMBIE 1967). A moraidőzítésű nyelvekben pedig a morák (másképpen a szótag súlyát megadó fonológiai egységek) valósulnának meg hasonló időtartamban (LADEFOGED 1975). A tényleges felvételeken történő mérések nem támasztották alá a beszédrítmusosztályokról szóló hipotéziseket (ROACH 1982, 2002; HOEQUIST 1983), egy későbbi kutatásból (RAMUS–NESPOR–MEHLER 1999) ugyanakkor kiderült, hogy ez a perceptuális élményen (észlelésen) alapuló osztályozás nem teljesen alaptalan. Az újszülöttek ugyanis különbséget tudtak tenni az anyanyelvükön elhangzott mondatok és más, nem ugyanabba a beszédrítmusosztályba tartozó nyelven elhangzott mondatok között. Az újszülöttek ezen különbségtételre akkor is képesek voltak, amikor a beszédet speciális módon szűrték, mégpedig úgy, hogy csak az alacsonyabb frekvenciák voltak érzékelhetők a beszédből, azaz a beszédértést lehetővé tevő tulajdonságokat eltávolították. A francia családba született újszülöttek például meg tudták különböztetni a hangsúly-időzítésű nyelvekhez sorolt angol és a klasszikusan moraidőzítésű nyelvnek tartott japán beszédrészleteit. A babák reakcióit az alapján számszerűsítették, hogy percenként hányszor szívták a cumit a különböző ingereket hallgatva. A kísérleti csoportban a babák az angol és a japán beszédet 2-2 különböző személytől felváltva hallgatták meg, míg a kontrollcsoportban szintén négy különböző beszélőtől hallottak beszédingereket, de a nyelv végig azonos maradt. Az eredmények azt mutatták, hogy az átlagos szívásszám megnőtt a kísérleti csoportban nyelvváltáskor, míg a kontrollcsoportban a kísérlet alatt végig változatlan maradt (NAZZI–BERTONCINI–MEHLER 1998). A babák észlelési tesztjei miatt továbbra is feltételezték, hogy a nyelvek elkülöníthetők beszédrítmusuk szerint, ezért a jelenséget akusztikailag másképpen próbálták megközelíteni. A beszédet magánhangzós és mássalhangzós szakaszok váltakozásaként definiálták (RAMUS–NESPOR–MEHLER 1999; GRABE–LOW 2002). Egy egységnek tekintették azt a beszédrészletet, amelyben mássalhangzók vannak közvetlenül egymás mellett, és nem választja el őket egymástól magánhangzó. Ezt az egységet nevezzük **mássalhangzós szakasznak**. Egy **magánhangzós szakasz** pedig olyan beszédrészlet, amelyben a magánhangzó vagy magánhangzók közvetlenül egymást követik, és nem különíti el őket mássalhangzó egymástól. A meghatározás szerint az alábbi megnyilatkozást így tagolhatnánk magánhangzós és mássalhangzós szakaszokra: |A|p|e||é|ksz|e|r|e|tn|e|kf|e||é|sl|eu|gr|á|ln|i|. Különböző mérőszámokat hoztak létre a magánhangzós és a mássalhangzós szakaszok időtartamaiból azért, hogy összehasonlítsák az eltérő nyelvek beszédrítmusát. A magánhangzós és mássalhangzós szakaszok időtartamának szórásai ( $\Delta V$ ,  $\Delta C$ ), illetve a magánhangzós szakaszok összegzett időtartamának aránya a teljes időtartamhoz

képest (%V) olyan mérőszámok, amelyek alapján a nyelvek elhelyezhetők egy skála mentén beszédritmusuk szerint. Az így létrehozott skálákon pedig valamelyest elkülönültek az eredeti beszédritmusosztályok nyelvei. Az 1. ábrán látható, hogy a klasszikusan szótag-időzítésüként besorolt spanyol és francia távolabb helyezkednek el a hangsúly-időzítésűnek tekintett angol vagy holland nyelvektől (RAMUS–NESPOR–MEHLER 1999).



1. ábra

*A nyelvek beszédritmus-mérőszámok alapján történő csoportosítása (AN = angol, HO = holland, OL = olasz, FR = francia, SP = spanyol, KA = katalán, JA = japán, LE = lengyel, MA = magyar) (RAMUS–NESPOR–MEHLER 1999; KOHÁRI 2018)*

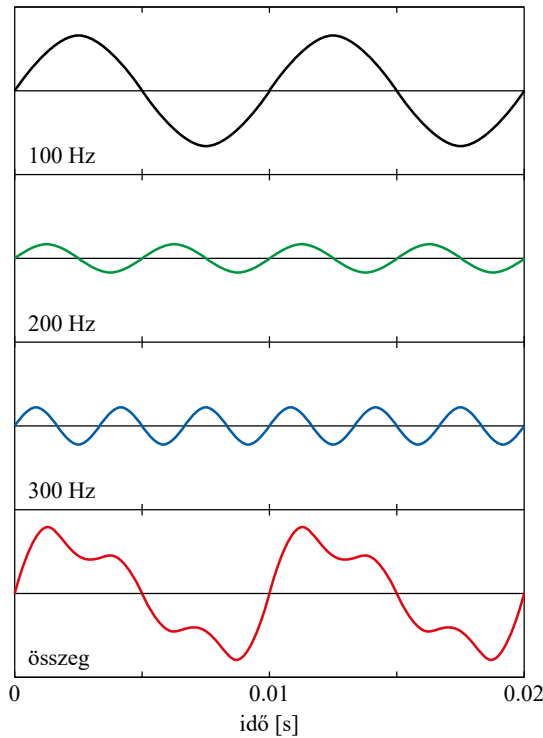
A dajkanyelv beszédritmusát is megkísérelték az imént tárgyalt keretben a beszédritmus-mérőszámok segítségével mérni. A kísérlet során a gyerekek képernyőn látható jeleneteket kellett szavaival leírnia, majd az édesanya megismételte ezeket a szavakat, utána az interjúztató is elmondta, hogy mi látható a klipben, amelyet az édesanya szintén megismételt. Az egyszerű jelenetek biztosították, hogy a gyerek és az interjúztató majdnem teljesen ugyanazokat a szavakat használja, így az édesanya szövege a két különböző regiszterben ugyanaz vagy majdnem ugyanaz volt. A spanyol, a katalán és a brit angol nyelvben azt találták, hogy mind a magánhangzós, mind a mássalhangzós szakaszok időtartamának variabilitása alacsonyabb volt a dajkanyelvi beszédben

a felnőttekhez szóló beszédhez képest. A beszédrimus-mérőszámok egy részéről ismert, hogy a tempó hatással lehet az értékekre, ezért olyan mérőszámokat is alkottak, amelyek a tempóingadozásoktól függetlennek tekinthetők. Amikor a két regiszter közötti beszédritmusbeli eltéréseket ezekkel a tempóra normalizált mérőszámokkal mérték, akkor is alacsonyabbnak bizonyultak az értékek a gyerekekhez szóló beszédben. A tempóeltérések tehát nem befolyásolhatták az eredményeket a két regiszterben, attól függetlenül is találtak eltérést a beszédritmusban (PAYNE 2009). Továbbá az angol nyelvben találtak egy olyan összefüggést is, hogy a magánhangzós szakaszok aránya a teljes időtartamhoz képest (%V) nagyobb volt a gyermekekhez szóló beszédben, mint a felnőttekhez szóló beszédben. Ez azt jelenti, hogy a magánhangzók összidőtartama arányaiban nagyobb volt a mássalhangzókéhoz képest a dajkanyelvi beszédmódban a másik regiszterhez képest. A spanyolban és a katalánban viszont nem találtak ilyen eltérést, az ausztrál angol spontán beszédben pedig semmilyen beszédritmusbeli különbséget nem mutatott a két regiszter (LEE et al. 2014). A több nyelvet is összehasonlító beszédritmusvizsgálat során arra a következtetésre jutottak, hogy a gyermekekhez szóló beszédben bizonyos nyelvek (pl. az angol és a spanyol) beszédritmusa kevésbé különül el, mint a felnőttekhez szóló beszédben. A tipikusan hangsúly-időzítésűekhez sorolt angol nyelv esetében például a dajkanyelv beszédritmusa közeledett a klasszikusan szótag-időzítésűekhez sorolt spanyol nyelv beszédritmusához (PAYNE 2009). A dajkanyelv beszédritmusa tehát meglehetősen nyelvfüggetlő.

A magyar nyelvre vonatkozóan azt találták, hogy a magánhangzós szakaszok és a mássalhangzós szakaszok időtartamának a variabilitása is nagyobb volt a dajkanyelvi, mint a felnőttekhez szóló beszédben, viszont ezek az eredmények visszavezethetők voltak a két regiszter tempóbeli eltéréseire. Tehát amikor figyelembe vették a dajkanyelv lassabb tempóját, a beszédritmusbeli különbségek eltűntek a két regiszter között. Az angol nyelvhez hasonlóan (PAYNE 2009) a magyarban is nagyobb volt a magánhangzós szakaszok időtartamaránya a teljes beszéd idejéhez képest (%V). A magyar anyanyelvi beszélők tehát a dajkanyelvben megjelenő lassítást a magánhangzók nagyobb nyújtásával érték el (KOHÁRI–DEME és mtsai 2020).

### 2.3. A dajkanyelv alaphangmagasságához kapcsolható tulajdonságok

A dajkanyelv egyik legtöbbet vizsgált tulajdonsága az alaphangfrekvencia és annak változásai. Foglalkozunk össze röviden, mit is értünk alaphangfrekvencia alatt. Ha a nyakunk elülső részére teszünk a kezünket, és például magánhangzókat ejtünk ki, érezhetjük a hangszalagjaink rezgését. A hangszalagok kváziperiodikus, azaz majdnem periodikus rezgése eredményeképpen létrejövő hang a **zöngé**. A zöngé összetett hang, amelynek van alaphangja/alaphangfrekvenciája ( $f_0$ ). A zöngé képzése során megjelennek az alaphang felharmonikusai is, azaz olyan rezgésösszetevők, amelyek frekvenciája az alaphangfrekvencia egész számú többszöröse. A 2. ábrán látható egy példa az alaphangra (folytonos vonallal jelölt, 100 Hz-es frekvenciájú összetevő) és a felharmonikusaira (pontozott, illetve szaggatott vonallal jelölt, 200 és 300 Hz-es frekvenciájú összetevő), illetve az alaphang és felharmonikusainak összegére, azaz az összetett rezgésre (legalsó vonal).



2. ábra

*Az alapprofrekvencia (folytonos vonal) és felharmonikusai (pontozott, illetve szaggatott vonal), valamint az ezekből álló összetett rezgés hullámformája (pontozott és szaggatott vonal)*

A dajkanyelv alapprofrekvenciájáról azt találták, hogy átlagosan magasabb a felnőttekhez szóló beszédhez viszonyítva. Többek között amerikai angol, brit angol, francia, japán, koreai, német, olasz, tagalog, tamil nyelvekben is kimutatták a két regiszternek ezt az alapprofrekvenciabeli eltérését (FERNALD et al. 1989; NARAYAN–MCDERMOTT 2016). Nemcsak az átlagos  $f_0$ -értékek térnek el a két regiszterben, hanem a hangterjedelem is; azaz az alapprofrekvencia minimuma és maximuma között nagyobb különbségeket találtak a dajkanyelvben, mint a felnőttekhez szóló beszédben. A hangterjedelem kiterjesztésének mértéke viszont nyelvfüggő, tehát például az amerikai angolban nagyobb hangterjedelembeli különbség volt megfigyelhető a két regiszter között, mint a japánban. Több nyelvben (pl. amerikai angol, brit angol, francia, német és olasz) azt találták, hogy az  $f_0$  variabilitása is nagyobb a dajkanyelvben, mint a felnőttekhez szóló beszédben (FERNALD et al. 1989).

A dajkanyelv sajátosságaival kapcsolatban felmerült, hogy kultúrafüggő lenne, és csak az iparosodott nyugati országokban lenne elterjedt az, hogy a felnőttek másképpen beszélnek a gyerekeikhez. BROESCH és BRYANT (2015) viszont megvizsgálták fidzsi-szigeteki és kenyai falvakban élő anyákat is, és saját gyermekükhöz szóló beszédük tulajdonságait összevetették az Amerikai

Egyesült Államokban élő édesanyák dajkanyelvének sajátoságaival. Mind az alapfrekvencia, mind az  $f_0$  variabilitása nagyobb volt a dajkanyelvben, függetlenül attól, hogy az édesanyák milyen országban éltek, sőt a tempójukon is lassítottak. A dajkanyelvi sajátosságokat tehát – úgy tűnik, hogy – kultúrától, országtól függetlenül mindenhol használják, bár hozzá kell tennünk, hogy nagy az egyéni variabilitás a különböző tulajdonságokban (NARAYAN–MCDERMOTT 2016).

Nemcsak az édesanyák, hanem más felnőttek beszédében is megfigyelhetők a dajkanyelvi sajátosságok. Az édesapák esetében hasonló mintázatokat találtak az alapfrekvenciával kapcsolatban, mint az édesanyáknál. Az apák esetében is kimutatható, hogy átlagosan magasabb alapfrekvencián beszélnek a gyerekeikhez, mint a felnőttekhez (FERNALD et al. 1989; DE PALMA–VANDAM 2017). A hangterjedelemmel és az  $f_0$ -variabilitással kapcsolatban ugyanakkor eltérő eredményeket kaptak a beszélők lakhelyétől, környezetétől függően. A Tanna-szigeteken lévő Vanuatuban élő édesapák esetében azt találták, hogy az édesanyákhoz hasonlóan ők is nagyobb hangterjedelmet használnak a dajkanyelvben, mint a felnőttekhez szóló beszédben (BROESCH–BRYANT 2018). A Nagy-Britanniában, Kanadában és az Amerikai Egyesült Államokban élő édesapák esetében ugyanakkor nem sikerült kimutatni, hogy élnének ezek az eszközökkel. Hogy vajon csak a különböző anyanyelv vagy az eltérő szociokulturális körülmények lennének hatással az eredményekre, ez a kérdés további vizsgálatokat igényel (SHUTE–WHELDALL 1999; BROESCH–BRYANT 2018).

A dajkanyelvi beszédmód prozódiai jellegzetességeit magyar anyanyelvű szülők esetében is vizsgálták már. Az alapfrekvencia nagyobb volt a dajkanyelvben, mint a felnőttekhez szóló beszédben, valamint az  $f_0$  maximuma is magasabbnak bizonyult a dajkanyelvben. A hangterjedelemmel kapcsolatban ellentmondásosak az eredmények. Az egyik vizsgálatban azt találták, hogy történetmesélés közben és tanítási situációban a szülők tágabb hangterjedelmet használnak a gyerekekhez szóló beszédben, mint a felnőttekhez szólóban (GERGELY et al. 2017). Mások viszont történetmesélésbe beleszótt felolvasott mondatokon nem találtak hasonló eltérést a két regiszter között (MÁDY et al. 2018, 2020). A magyarban vizsgálták azt a kérdéskört is, hogy vajon az, hogy az édesanyának korábban volt már gyereke vagy nem, befolyásolja-e a csecsemőjéhez szóló beszéd prozódiai sajátosságait. Az eredmények azt mutatták, hogy a többgyerekes anyukák magasabb alapfrekvenciával beszélnek újszülöttjeikhez, mint azok az édesanyák, akik első gyermekükhöz beszéltek. A **dallamhangsúly** (*pitch accent*), azaz a mondatszintű prominencia vagy kiemelkedés is eltéréseket mutatott a gyerekek száma szerint. A dallamhangsúlyos szótag alapfrekvenciájának kontúrja meredekebben valósult meg a többgyerekes édesanyák esetében, mint az először szülő anyák esetében (MÁDY et al. 2018). Ez azt jelentette, hogy a mesében új, megtanítandó elemként szereplő, ismeretlen manónevek prominensebbek, jobban hangsúlyozottak voltak a korábban már szült anyák beszédprodukciónak. Véltetően a többgyerekes édesanyák tapasztalata a babákkal, illetve a kórházi körülményekkel a dajkanyelvre jellemző tulajdonságok erőteljesebb használatát eredményezhette. Az édesanyák mellett az édesapák dajkanyelvét is vizsgálták már a magyarban (GERGELY et al. 2017), és 4 hónapos és 3,5 éves közötti gyerekekhez szóló beszédben azt találták, hogy az édesapák is magasabb alapfrekvenciával beszélnek a gyerekekhez, mint a felnőttekhez. Sőt az édesapák alkalmazkodtak a gyerek korához is, ugyanis

azt találták, hogy az édesapák 4 hónapos gyerekekhez magasabb alapfrekvenciával beszéltek szabad játék közben, mint idősebb gyerekeikhez. Az édesanyáknál nem találták a gyerek korával kapcsolatos eltérést, továbbá a felolvasott mondatokban sem mutatták ki a gyermek korának hatását a dajkanyelv különböző prozódiai tulajdonságaira. Ugyanakkor hozzá kell tennünk, hogy ez a vizsgálat nem valódi longitudinális vizsgálat volt, hanem különböző életkorú gyerekek szüleinek eltérő beszédmódjait hasonlították össze. Egy másik kutatásban magyar anyanyelvű édesanyákat arra kértek, hogy a gyermekükhöz és egy felnőttöz is beszéljenek a gyermek 0, 4, 8 és 18 hónapos korában is, így valódi longitudinális vizsgálatot tudtak elvégezni (MÁDY et al. 2020). Az eredmények azt mutatták, hogy a 4, 8 és 18 hónapos gyermekhez szóló beszéd között nem volt lényeges eltérés az alapfrekvenciával kapcsolatos paraméterek tekintetében. Az újszülöttekhez szóló beszéd viszont valamelyest eltért a többitől, ugyanis több paraméter, úgymint az alapfrekvencia átlagos értéke és a maximuma alacsonyabbak voltak ebben az életkorban. Az eredményeket magyarázhatja, hogy az édesanyák első gyermeküket szülték, és talán kevésbé voltak tapasztaltak a csecsemőkkel való kommunikációban, ahogy korábban említettük, másrészt külső tényezők (pl. kórházi körülmények, fáradtság) is szerepet játszhattak az eltérésekben (MÁDY et al. 2020). Továbbá elképzelhető, hogy azért nem mutatkozott különbség a 4, 8, illetve 18 hónapos gyerekekhez szóló beszédben, mert a szülők feltehetően nem annyira az életkorhoz, hanem a gyerek fejlődési szintjéhez, nyelvi készségeihez igazítják beszédük sajátosságait (vö. QUIGLEY–NIXON–LAWSON 2019).

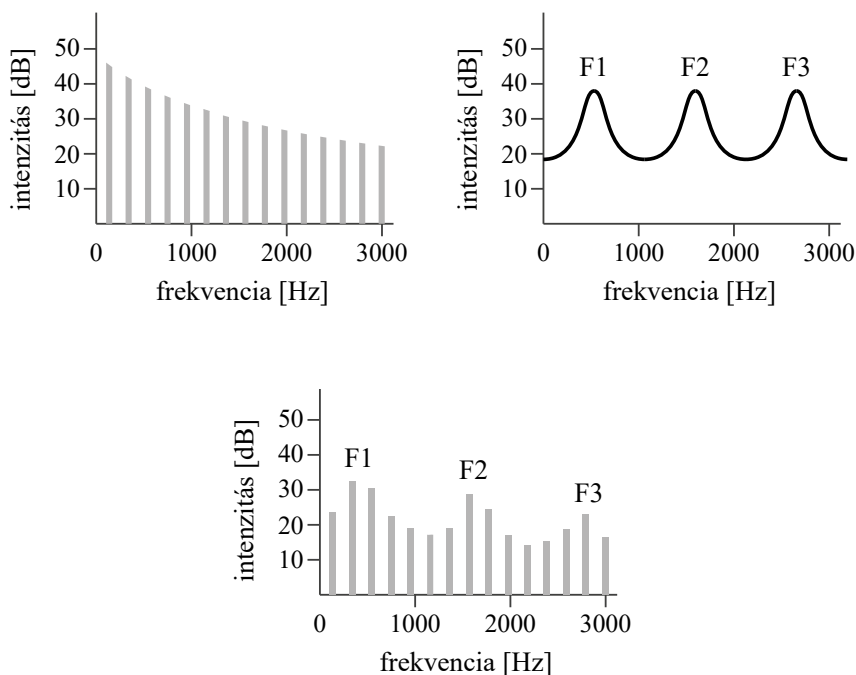
Összegezve az eddigieket elmondható, hogy a dajkanyelv alapfrekvenciája az egyik legtöbbet kutatott terület, de még továbbra is sok a tisztázatlan kérdés a témában, amely további vizsgálatokat igényel.

## 2.4. A dajkanyelv magánhangzótere

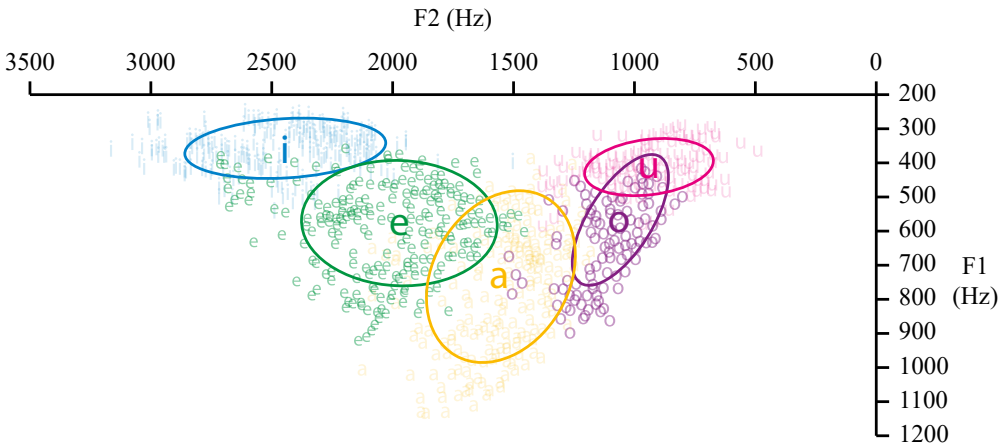
A dajkanyelvnek – a prozódia mellett – egy másik sokat vizsgált akusztikai jellemzője a **magánhangzótér**. A magánhangzók képzésekor a hangszalagok rezgésének eredményeképpen létrejön a zöngé. A kiáramló levegő a garatüregben keresztül a szájüregbe jut, ahol az artikulációs szervek, mint például a nyelv és az ajkak állása sajátos üregi rezonanciát hoz létre. Másképpen megfogalmazva, az üregek bizonyos rájuk jellemző frekvenciákat felerősítenek, másokat pedig csillapítanak. Azokat a frekvenciákat, amelyeket a toldalékcső, azaz a garatüreg, a szájüreg, az orrüreg együttesen felerősítenek, **formánsoknak** nevezzük. A 3. ábra bal oldalán látható a zöngé mint komplex hang. Az alapfrekvenciának (a 3. ábrán 200 Hz) megjelennek az egész számú többszörösei is, amelyekből csak bizonyos frekvenciák erősödnek fel a toldalékcső rezonanciájának hatására. A különböző formánsértékek  $F_1$ ,  $F_2$  stb. jelölésekkel láthatók az ábrán. Az első formáns frekvenciaértéke a nyelv függőleges, a második formánsé elsősorban a nyelv vízszintes mozgásával van összefüggésben. A felső nyelvállású magánhangzók esetében alacsonyabbak az első formánsértékek az alsó nyelvállásúakhoz képest, és az elől képzett magánhangzók esetében a második formáns értékei magasabbak a hátul képzett magánhangzókhoz képest. Ezért szokás



a magánhangzók  $F_1$  és  $F_2$  formánsait egymás függvényében is kirajzoltatni oly módon, hogy az valamelyest emlékeztessen a magánhangzónak a szájüregben található képzési helyére (4. ábra). A magánhangzók elhelyezése az első két formáns által meghatározott koordináta-rendszerben jelöli ki az akusztikai magánhangzóteret (l. részletesebben GÓSY 2004; DEME 2016). Az artikulációs szempontból a legtávolabbi helyeken képzett magánhangzók, az úgynevezett sarokmagánhangzók (a felső nyelvállással képzett elülső /i/ és hátulsó /u/, valamint a legalsó nyelvállású /a/) formánsértékei kijelölnek egy területet az első és második formáns koordináta-rendszerében (4. ábra). A sarokmagánhangzók formánsai által kijelölt háromszögnek a területe lehet kisebb vagy nagyobb, attól függően, hogy mennyire vannak távol egymástól a magánhangzók. A nagyobb magánhangzóter nem azt jelenti, hogy minden magánhangzó formánsértéke nagyobb lesz, hanem bizonyos formánsértékek nőnek, mások csökkennek, hogy nagyobb legyen a magánhangzók közti távolság (a magánhangzók jobban megkülönböztethetők legyenek egymástól), és így a háromszög alapterülete. Ha például a legalsó nyelvállású /a/ első formánsértéke nő, miközben a többi magánhangzó  $F_1$ -értéke változatlan, akkor nagyobb lesz a magánhangzóter. Ha viszont például a felső nyelvállással képzett elülső /i/ és hátulsó /u/ első formánsértéke nő, miközben a harmadik magánhangzó  $F_1$  formánsértéke változatlan, akkor kisebb lesz a magánhangzóter, ahogyan az a 4. ábráról leolvasható.



3. ábra  
A magánhangzó formánsai



4. ábra

*Különböző magánhangzó-realizációk a magánhangzótérben (az ábra a phonR [McCLOY 2014] csomaggal készült)*

A dajkanyelvben igen ellentmondásos eredményeket találtak eddig a magánhangzótérrel kapcsolatban. Több nyelvben (amerikai angol, japán, mandarin, orosz, svéd) kimutatták, hogy a magánhangzótér nagyobb a dajkanyelvben, mint a felnőttekhez szóló beszédben (KUHLE et al. 1997; ANDRUSKI–KUHLE–HAYASHI 1999; LIU–KUHLE–TSAO 2003; KALASHNIKOVA–BURNHAM 2018). Az elemzésekhez a felső nyelvállással képzett elülső /i/ és hátulsó /u/, valamint a legalsó nyelvállású /a/ sarokponti magánhangzókat használták, amelyeket az amerikai angol, az orosz és a svéd dajkanyelvben is vizsgáltak. Az első és második formáns meghatározását követően elhelyezték a mérési eredményeket a koordináta-rendszerben, és kiszámolták a magánhangzók közti távolságokat. A három magánhangzó egy háromszöget rajzolt ki, amelynek területe nagyobb volt a dajkanyelvben, mint a felnőttekhez szóló beszédben. Az eredmények azt mutatták, hogy nem egyszerűen minden formáns értéke nőtt, hanem bizonyos formánsértékek csökkentek. Például az /a/ első formánsértékei és az /i/ második formánsértékei nagyobbak voltak dajkanyelvben, mint felnőttekhez szóló beszédben, míg az /u/ második formánsértékei tendenciaszerűen alacsonyabbak voltak. Így növekedett meg a magánhangzótér, és a magánhangzó-minőségek jobban elkülönültek egymástól akusztikailag (KUHLE et al. 1997).

Más vizsgálatokban azonban nem találták a dajkanyelv magánhangzótérét nagyobbának a felnőttekhez szóló beszédéhez képest. A norvég dajkanyelvet vizsgálva éppen az ellenkezőjét mutatták az eredmények, mégpedig azt, hogy a dajkanyelv magánhangzótere kisebb (ENGLUND–BEHNE 2006). A brit angol, a francia és a japán nyelv vizsgálatában a magánhangzók első formánása nagyobb volt a zárt magánhangzók esetében, ami arra utalhat, hogy nyíltabbá vált az ejtés a dajkanyelvben. Eközben az /a/ első formánsértékei közel változatlanok vagy akár még nagyobbak is voltak. Így a magánhangzók közötti távolság közel azonos maradt, vagy akár még csökkent is. Ezek az eredmények tehát

nem támasztották alá a magánhangzótér kiterjesztését a dajkanyelvben (DODANE–AL-TAMIMI 2007). Az eltérések oka lehet például a gyerekek különböző életkora vagy más szavak, más mássalhangzó-környezet használata az anyagban, és még sorolhatnánk az anyagbeli és módszertani eltéréseket. Azt azonban fontos hozzátennünk, hogy az eddig felsorolt kutatásokban csak három magánhangzót (/i/, /a/, /u/) vizsgáltak, amelyek a magánhangzótér szempontjából a legszélsőségesebb pozíciókat vették fel, de nem elemezték a nyelvekben meglévő összes magánhangzót. Ha a magánhangzók jobb megkülönböztetése érdekében valóban távolabb valósulnának meg egymástól a magánhangzók, akkor az adott nyelv összes magánhangzójának távolabb kellene realizálódnia egymástól. Legalábbis azoknak a magánhangzóknak, amelyeknek az adott nyelvben jelentésmegkülönböztető szerepük van. A jelentésmegkülönböztető szerep azt jelenti, hogy egy nyelvben találunk olyan minimális párt, amelyben csak az a két hangzó tér el (pl. *kor* vs. *kör*), és a pár mindkét tagja értelmes szó. A mérések azonban nem erősítették meg azt a feltételezést, hogy minden jelentésmegkülönböztető magánhangzó távolabb realizálódna egymástól (CRISTIÀ–SEIDL 2013; McMURRAY et al. 2013). Több magánhangzó akusztikai vizsgálatokba történő bevonásából tehát szintén arra lehet következtetni, hogy a beszélők a dajkanyelvben nem erősítik fel a nyelvi kontrasztokat.

A magyarban azt találták, hogy a magánhangzótér nagyobb a dajkanyelvben, mint a felnőttekhez szóló beszédben, és ezt a feltevést két különböző kutatás eredményei is megerősítették (GERGELY et al. 2017; MÁDY et al. 2020). Az egyikben 4 és 25 hónap közötti gyermekekhez szóló, felolvasott mondatokat vizsgáltak, amelyben az /a: i u/ magánhangzók formánsértékei közötti távolságot mérték (GERGELY et al. 2017); míg a másikban 0 és 18 hónap közötti gyerekekhez szóló felolvasásban minden előforduló magánhangzót figyelembe vettek, és a magánhangzóteret egy másik eljárással, a magánhangzó-diszperzióval mérték (MÁDY et al. 2020). A **magánhangzó-diszperzió** egy olyan eljárás, amely során a magánhangzók formánsértékeinek kiszámolják a súlypontját (egyfajta középpont) a koordináta-rendszerben, majd megadják a formánsértékek átlagos távolságát ettől a súlyponttól. A módszertani és anyagbeli eltérések ellenére ugyanarra a következtetésre jutottak a regiszter hatása tekintetében. Az elemzések azt is feltárták, hogy nemcsak az édesanyák, hanem az édesapák gyermekükhöz szóló beszédében is kimutatható, hogy nagyobb magánhangzótér jelenik meg a dajkanyelvben, mint a felnőttekhez szóló beszédben. Az édesanyák nagyobb magánhangzóteret használtak 4, illetve 16 hónap körüli gyermekükhöz, mint az édesapák (GERGELY et al. 2017). Amikor valódi longitudinális vizsgálatot hajtottak végre, tehát az édesanyák a saját gyermeküknek 0, 4, 8 és 18 hónapos korában is elmondták ugyanazt a mesét, akkor azt találták, hogy nincs hatással az életkor az édesanyák dajkanyelvében megvalósuló magánhangzótérre, kivéve az újszülöttkort. A baba néhány napos korában ugyanis az édesanyák nem használtak nagyobb magánhangzóteret a csecsemőjükhöz a felnőttekhez szóló beszédhez képest (MÁDY et al. 2020). Hasonlóan a prozódiai paraméterekhez, itt is a kórházi körülmények, a babákkal való kommunikációban való tapasztalatlanság merülhet fel magyarázatként, ugyanakkor azt is feltételezhetjük, hogy ekkor még az édesanyák nem feltételezik, hogy babájuk megértene, mit mondanak nekik, ezért nem is erősítik fel az egyes magánhangzók közötti eltéréseket.

Egy másik vizsgálatban nem pusztán a magánhangzók közti különbségeket, hanem a fonológiai hosszúság szerinti szembenállást is elemezték a magyar dajkanyelvben (DEME és mtsai 2019).

Ehhez két magánhangzópárt elemeztek: az /u/–/u:/ és az /ɒ/–/a:/ magánhangzópárok realizációit. A rövid /u/ és párja, a hosszú /u:/ esetében a fonológiai hosszúsági szembenállás fonetikai vetülete csak az időtartamban valósul meg, hiszen mindkét hangzó zárt, kerekített, felső nyelvállású magánhangzó. A másik magánhangzópár esetében viszont a hangzók hangszínbeli és időtartambeli eltéréseket is mutatnak, mivel az /ɒ/ ajakkerekítéses, míg az /a:/ ajakréses hangzó. A kísérletben úgy kontrollálták a magánhangzók környezetét, hogy a koartikulációs hatások ne befolyásolhassák az eredményeket. Ehhez CVC (mássalhangzó–magánhangzó–mássalhangzó) szerkezetű álszavakat hoztak létre, amelyben a mássalhangzók képzéshelye azonos volt. Az édesanyáknak szövegbe ágyazva a *Szut*, *Tút*, *Zat* és *Dát* elképzelt manóneveket kellett felolvasniuk a saját gyerekeküknek és egy felnőttnek is. Mivel a kutatók azt szerették volna megvizsgálni, hogy az édesanyák a fonológiai kontrasztot felerősítik-e, ezért összevetették a magánhangzópárok időtartamát, valamint a formánsaik frekvenciaértékét. A rövid–hosszú magánhangzópárok időtartamának különbsége nagyobb volt a dajkanyelvben, mint a felnőttekhez szóló beszédben, amikor viszont a rövid–hosszú párok időtartamának arányát vették, akkor a két regiszter nem mutatott különbséget. Felmerül a kérdés, hogy vajon az időtartamok különbsége vagy az időtartamok aránya játszik jelentős (vagy jelentősebb) szerepet a babák észlelésében. Mind a csecsemők, mind a felnőttek esetében azt találták, hogy elsősorban az időtartamarányok határozzák meg az észlelést (BRANNON et al. 2008), tehát ezek alapján az időtartamok nem mutattak a babák észlelése szempontjából releváns kontraszterősítést a magánhangzópárok között. Továbbá a spektrális elemzések sem erősítették meg azt a feltételezést, hogy az édesanyák csecsemőikhez intézett közléseikben erősítenék a fonológiai magánhangzókontraszt párpai közötti minőségbeli szembenállást. Habár az /ɒ/–/a:/ magánhangzópár mutatott hangszínbeli eltérést, azaz a magánhangzótérben elkülönültek egymástól, a regiszternek nem volt hatása a távolságukra. Összefoglalva az elemzéseket, a dajkanyelv nem mutatta a fonológiai kontrasztok felerősítését a rövid–hosszú magánhangzópárokban (DEME és mtsai 2019).

## 2.5. A dajkanyelv zöngeminősége

A dajkanyelv **zöngeminőségét** eddig meglehetősen keveset vizsgálták, mégis egy olyan új, ígéretes kutatási terület, amelyet feltétlenül meg kell említenünk. A dajkanyelv zöngeminőségével kapcsolatos eredmények ismertetése előtt azonban röviden áttekintjük, mit is jelent ez a terminus, és hogyan tudunk hozzá mérőszámokat kapcsolni. Ismétlésként: a zöngé a hangszalagok mozgásának eredményeként létrejövő kváziperiodikus, azaz majdnem periodikus rezgés. A zöngé minősége alatt pedig azt a sajátosságot értjük, amelynek alapján a hallgató két azonos magasságú és intenzitású hangot eltérőnek talál (MARKÓ 2013). Többféle zöngeminőséget különböztetnek meg a szakirodalomban, amelyek elhelyezhetők egy skála mentén, a hangszalag nyitottsága alapján. A képzeletbeli skála két végpontjában a zöngétlen képzési mód és a teljes glottális zár van, közöttük pedig középen a modális képzésmódú zöngé helyezkedik el. Ha a zöngé periodikus rezgés eredményeként, a megszokott módon jön létre, modális zöngéről beszélünk. Ha a rezgés nem annyira periodikus, és a hangszalagok rezgése, valamint a zöngé

hangzása a megszokottól valamilyen módon eltér, akkor nem modális zöngéről van szó. A nem modális zöngeminőségben belül megkülönböztethetünk recsegő zöngét (*creaky voice*) és leheletes zöngét (*breathy voice*) aszerint, hogy a hangszalagok mennyire zártak a hangképzés során (GARELLEK 2019). A leheletes zöngéképzés során a hangszalagok nyitottabbak (l. a videót<sup>1</sup>), recsegő zöngéképzés során pedig a hangszalagok zártabbak (l. a videót<sup>2</sup>), mint a modális zöngé létrejöttékor. A különböző zöngéképzési módok összehasonlíthatók beszédfelvételek akusztikai elemzésével, ugyanis a magánhangzók esetében mérhetőek olyan paraméterek, amelyek az észlelhető zöngeminőségbeli különbségekhez kapcsolhatók.

Egy nemrég megjelent kutatásban azt találták, hogy az édesanyák hajlamosak leheletebb zöngét használni 20 hónapos gyerekekhez szóló beszédükben, mint a felnőttekhez szólva (MIYAZAWA et al. 2017). Sőt a dajkanyelvnek ez a zöngeminőségbeli sajátossága már a szülés előtt megjelenhet. Az édesanyák egy része a terhesség időszakában is beszél gyermekéhez, és a cseh nyelvre azt találták, hogy a leheletebb zöngé már ekkor is jelen van a dajkanyelvben (CHLÁDKOVÁ et al. 2019).

A magyarra is vizsgálták a dajkanyelv zöngeminőségét. 4 hónapos csecsemőkhöz szóló beszéd magánhangzóin (/v/, /a:/, /e/, /i/, /o/ és /u/) azt találták több paraméter alapján, hogy a dajkanyelv leheletebb zöngével képzett a felnőttekhez szóló beszédnél. Az /i/ és /u/ hangzók viszont nem mutattak eltérést a két regiszter között, aminek a hátterében az állhat, hogy mindkettő felső nyelvvállású magánhangzó, és azokról ismert, hogy hajlamosabbak a leheletes zöngéhez hasonló tulajdonságokkal realizálódni, így elképzelhető, hogy ez elmosta a két regiszter közti különbséget. Másrészről a zöngeminőség akusztikai méréseit befolyásolják a környező mássalhangzók, amely ezen két magánhangzó esetében szintén befolyásolhatta az eredményeket (KOHÁRI–REICHEL et al. 2020). Összességében viszont elmondható, hogy a dajkanyelvnek a leheletes zöngé egy olyan sajátossága, amelyet nemcsak cseh, japán, hanem magyar nyelvben is kimutattak a különböző életkorú babákhoz szóló beszédben.

### 3. A dajkanyelv hatása a gyerekekre és szerepe az interakcióban

A dajkanyelv, illetve annak használata a felnőtt-gyerek interakcióban sokféle funkciót tölthet be egyszerre. A különböző tulajdonságai révén megragadja a gyermekek figyelmét, pozitív érzelmeket közvetít feléjük, segíti a nyelvelsajátítást is. A dajkanyelvnek a különböző kommunikációban betöltött szerepeit tekintjük át röviden a következő alfejezetekben.

---

<sup>1</sup> COMPANION WEBSITES 2011. Breathy-Voice.mov. <https://www.youtube.com/watch?v=9cKnUFZjs8k> (A letöltés ideje: 2021. február 12.)

<sup>2</sup> COMPANION WEBSITES 2011. Creaky-Voice.mov. <https://www.youtube.com/watch?v=BYSZS1LaABQ> (A letöltés ideje: 2021. február 12.)

### 3.1. A figyelem felkeltése és fenntartása a dajkanyelv sajátosságaival

A dajkanyelv sajátosságai elősegítik, hogy a gyerekek jobban bevonódjanak a szociális interakcióba, mivel jobban felkelti az érdeklődésüket, jobban figyelnek rá, mint a tipikusan felnőttekhez szóló beszédre. A vizsgálatok megmutatták, hogy a gyerekek hosszabban néztek a vizuális ingerekre akkor, amikor dajkanyelvet hallottak, mint akkor, amikor felnőttekhez szóló beszédet (PEGG–WERKER–MCLEOD 1992; COOPER et al. 1997). Az újszülöttek is a dajkanyelvet preferálják a felnőttekhez szóló beszédhez képest (COOPER 1993), sőt a frontális agyi aktivitásuk akkor is megnövekszik a dajkanyelvet hallgatva, amikor alszanak (SAITO et al. 2007). Az egyik vizsgálatban 98 újszülött agyi aktivitását vizsgálták, miközben magyar nyelvű dajkanyelvet és felnőttekhez szóló beszédet hallgattak. Az agyban lévő idegsejtek elektromos aktivitásának mérését elektroenkefalográfiával (EEG) végezték, amellyel elektródák közötti potenciálkülönbségeket mértek. A különböző regiszterekben kimondott szavakra adott agyi válaszok vizsgálata feltárta, hogy az újszülöttek megkülönböztetik a két regisztert, és más agyi területek aktivizálódnak attól függően, hogy dajkanyelvi vagy felnőttekhez szóló beszédre jellemző akusztikai ingert kapnak (HÁDEN et al. 2020). A babák tehát már újszülött koruktól kezdve észlelik, ha dajkanyelven szólnak hozzájuk, sőt ez a beszédmód jobban felkelti az érdeklődésüket is.

Felmerül a kérdés, hogy vajon milyen akusztikai tulajdonságok váltják ki azt, hogy a gyerekek jobban figyelnek a dajkanyelvi beszédmódra, mint a felnőttekhez szóló beszédre. Az egyik tanulmány szerint a magasabb alaphfrekvencia, illetve a nagyobb alaphfrekvenciabeli variabilitás nagyobb figyelmet eredményez a gyerekeknél (FERNALD–KUHL 1987; SEGAL–NEWMAN 2015), míg más jellemzőknél, mint például a hangerő vagy a beszédritmus, nem találtak ilyen összefüggést. Ugyanakkor hozzá kell tennünk, hogy a dajkanyelvi tulajdonságok ilyen módon történő szétválasztása csak nem természetes beszéd létrehozásával lehetséges, holott ismert, hogy a természetes beszédre a gyerekek jobban figyelnek, mint a nem természetes beszédre (DUNST–GORMAN–HAMBY 2012). Nem kizárható tehát, hogy a dajkanyelvnél más, könnyen felismerhető sajátosságai is hozzájárulnak a figyelem felkeltéséhez és fenntartásához, csak a szintetizált beszéddel ezt nem sikerült eddig kimutatni. Habár vannak olyan kutatások, amelyek szerint a dajkanyelv preferálása csökken a gyerekek egyéves korára (SAINT-GEORGES et al. 2013), más vizsgálatok eredményei azt mutatták, hogy 12 és 16 hónapos gyerekek hosszabban figyeltek a magasabb és változatosabb alaphfrekvenciájú beszédre (SEGAL–NEWMAN 2015). Egy átfogó vizsgálat keretében, amelyben 67 különböző laboratórium vett részt a Föld minden tájáról (Ausztrália, Ázsia, Észak-Amerika, Európa), azt találták, hogy a babák 4 hónapos korától kezdve 14 hónapos korukig egyre jobban preferálták a dajkanyelvet a nem dajkanyelvi jellegzetességekkel rendelkező beszédhez képest (MANYBABIES CONSORTIUM 2020). A gyerekek tehát életük első másfél évében a dajkanyelvet részesítik előnyben a felnőttekhez szóló beszédhez képest, hosszabban figyelnek rá, és ez elsősorban feltehetően a dajkanyelv alaphfrekvenciával kapcsolatos tulajdonságainak köszönhető.

A dajkanyelv nem pusztán felkelti a gyerekek figyelmét, és jelzi, hogy az adott üzenet a gyerekeknek szól, hanem elősegíti azt is, hogy a gyerekek kövessék a felnőttek tekintetét egy tárgyra. Amikor egy felnőtt dajkanyelven köszöntötte a 6 hónapos csecsemőket, majd egy tárgyra nézett,

akkor a babák hajlamosak voltak követni a felnőtt tekintetét. Amikor viszont a felnőtt nem használta a dajkanyelvre jellemző prozódíát, nem néztek ugyanarra a tárgyra (SENJU–CSIBRA 2008). A dajkanyelv tehát annak jelzéséként (is) funkcionál az interakcióban, hogy a kommunikációban részt vevő személy valamilyen új vagy valamilyen szempontból releváns információt fog közölni. A dajkanyelv tehát ily módon elősegíti az ismeretszerzést, a tanulást, és előremozdíthatja a nyelvelsajátítást.

### 3.2. A dajkanyelv mint érzelemkifejező beszéd

A dajkanyelv egy másik lehetséges funkciója, hogy a beszélő érzelmeit közvetíti a gyerek felé. Az érzelmi állapotunk ismerten hatással van a beszédünkre. Jó néhány vizsgálat kimutatta, hogy a beszéd prozódiai sajátosságai és az érzelmi aktivációs szint között kapcsolat áll fenn. Az öröm, boldogság például magasabb alaphangfrekvenciával és nagyobb alaphangfrekvencia-variabilitással jár együtt (SCHERER 2003). Elfogadott nézet, hogy a dajkanyelvben azért jelennek meg hasonló akusztikai jellegzetességek, mert azok is a pozitív érzelmek kifejezői (SPINELLI–FASOLO–MESMAN 2017).

A nagyobb magánhangzótérrel is felvetették, hogy az érzelemkifejezés, illetve az agresszióhiány kifejeződésének következményeként jelenne meg. Egy artikulációs vizsgálat során ugyanis azt találták, hogy a gyerekekhez szóló beszéd során az ausztrál angol anyanyelvű beszélők meg rövidítik a toldalékcsovet oly módon, hogy megemelik a géjét. Ennek következtében nagyobb lesz a feszültség a hangszalagok között, ez pedig magasabb alaphangfrekvenciát eredményez. A gége megemelése által létrejött rövidebb toldalékcso pedig magasabb magánhangzóformáns-értékekhez is vezet. Az ajkak és a nyelv mozgása nem mutatott eltérést a két regiszterben, tehát azok nem kapcsolhatók a formánstér megváltozásához. Ebből a szerzők arra a következtetésre jutottak, hogy a magánhangzók formánsértékeinek alakulása csak egy mellékterméke a megemelkedett gégének. Magyarozatuk szerint mivel a gyerekek kisebb toldalékcsovel rendelkeznek a felnöttekhez képest, és magasabb alaphangfrekvenciával beszélnek, ezért a felnöttek pusztán megpróbálnak hozzájuk hasonlóbban beszélni, ezzel mutatva közeledésüket. A nagyobb formánstér dajkanyelvi megjelenése mögött szerintük tehát nincs kifejezett tanítási cél, hanem inkább a barátságosság, a nem agresszív magatartás és az interakcióban részt vevő beszédéhez való igazodás kifejeződése állhat a háttérben (KALASHNIKOVA–CARIGNAN–BURNHAM 2017). Ehhez hasonlóan holland édesanyák dajkanyelvét vizsgálva azt találták, hogy nem nagyobb a magánhangzótér a dajkanyelvben, mint a felnöttekhez szóló beszédben, hanem csak bizonyos formánsértékek magasabbak. Ezek az eltérések pedig magyarázhatók lennének a boldog érzelmi állapottal is, amelynek hasonló akusztikai lenyomatai vannak a beszédben (BENDERS 2013).

A dajkanyelv figyelemfelhívó és -fenntartó aspektusaival kapcsolatos vizsgálatokat (3.1. fejezet) illetően is felmerült, hogy az eredmények magyarázhatók lennének a különböző regiszterek érzelmi töltetével. Azt feltételezték, hogy a gyerekek azért részesítik előnyben a dajkanyelvet a felnöttekhez szóló beszéddel szemben, mert az előbbi pozitív érzelmeiket közvetíti, míg az utóbbi

inkább semleges az érzelmek szempontjából. Ezt a hipotézist látszik megerősíteni az a kísérlet, amelyben hat hónapos babák preferenciáit vizsgálták. A vizsgálat során egy felnőttet arra kértek, hogy különböző érzelmi állapotokban (boldog, semleges, szomorú) olvasson fel egy szöveget egy felnőttnek, aztán egy fotón lévő gyerekeknek. Majd ezeket a felvételeket játszották le csecsemőknek, akiknek a tekintetét követték a különböző ingerek közben. Ahogy az előző fejezetben láthattuk, a babák alapvetően a dajkanyelvre figyelnek hosszabban, jobban a felnőttekhez szóló beszédhez képest. Ebben a kísérletben viszont azt találták, hogy amikor mindkét regiszter pozitív érzelmekkel valósult meg, akkor a babák nem tettek különbséget a beszédmódok között. Az eredményekből azt a következtetést vonták le, hogy a beszédben megjelenő magasabb alapfrekvencia és nagyobb alapfrekvenciabeli variabilitás nem elegendő ahhoz, hogy a babák hosszabban, jobban figyeljenek. A két regiszter közötti fő különbség szerintük a beszédmódok érzelmi töltetében rejlik (SINGH–MORGAN–BEST 2002).

### 3.3. A dajkanyelv szerepe a nyelvelsajátításban

A dajkanyelv nemcsak felkeltheti a figyelmet, és érzelmet közvetíthet a gyerek felé, hanem segítheti a nyelvelsajátítást is. Kísérletes úton is igazolták, hogy a dajkanyelv elősegíti a szótanulást. A kisebb gyerekek (21 hónap körül) ugyanis könnyebben tudnak belőle szavakat tanulni, mint a felnőttekhez szóló beszédből (MA et al. 2011). De vajon a dajkanyelvnek milyen tulajdonságai járulnak hozzá ahhoz, hogy a gyerekeknek egyre nagyobb legyen a szókinccse?

A lassabb tempó az egyik olyan tulajdonsága a dajkanyelvnek, amelyet közvetlen kapcsolatba hoztak a szótanulással. Egyrésztől a lassabban elhangzó szavak és megnyilatkozások segítik a gyerekek szövegfelismerését (SONG–DEMUTH–MORGAN 2010; WANG–LLANOS–SEIDL 2017). Másrészt összefüggést találtak az amerikai angol anyanyelvű édesanyák beszédének alacsonyabb dajkanyelvi tempója és az **expresszív szókinccs**, azaz a gyerekek által aktívan használt szavak száma között (RANERI et al. 2020). A 7 hónapos babához való lassabb beszéd kapcsolatot mutatott a gyerek kétéves kori expresszív szókinccsével. Minél lassabban beszélt az édesanyja a gyermekéhez, annál nagyobb lett a gyerek későbbi szókinccse. Az még magyarázatra szorul, hogy miért a baba fél éves kora körüli dajkanyelvi tempóértékek mutattak korrelációt a szókinccsel, és miért nem más életkor (10, 18 vagy 24 hónapos kor). Emellett kérdés még, hogy az amerikai angolon kívül ez az összefüggés más nyelvekben is kimutatható-e. Ugyanakkor úgy tűnik, hogy a dajkanyelv lassabb tempója kísérletileg is igazolhatóan fontos szerepet tölt be a beszédfeldolgozásban és a nyelvelsajátításban.

A másik dajkanyelvi tulajdonság, amelyről azt feltételezik, hogy segíti a szavak felismerését, és ezáltal azok elsajátítását, az a nagyobb magánhangzótér. Az akusztikai adatok azonban – mint láttuk – nem támasztották alá, hogy a magánhangzótér minden nyelv esetében egyértelműen nagyobb lenne a dajkanyelvben. Továbbá a kísérletek azt sem igazolták, hogy a különböző magánhangzók ejtése a nyelv precízebb mozgásának következtében távolabb valósulna meg a szájüreg belsejében, és hogy ennek következtében a magánhangzók távolabb helyezkednének el egymástól



a formánstérben, ezáltal pedig jobban elkülöníthetők lennének (1. 1.4. fejezet). Ráadásul a magánhangzótér megnagyobbodásáról azt gondolják, hogy a gége megemeléssel létrehozott rövidebb toldalékcso eredményeként jön létre egyfajta melléktermékként (KALASHNIKOVA–CARIGNAN–BURNHAM 2017; 1. 2.2. fejezet). Ettől függetlenül viszont a létrejött akusztikai sajátosság még befolyásolhatja a befogadó percepcióját, beszédfeldolgozását és ezáltal a nyelvelsajátítást.

Egy kísérletben azt találták, hogy a gyerekek jobban, illetve gyorsabban ismerték fel azokat a szavakat, amelyekben a magánhangzók távolabb valósultak meg egymástól a magánhangzótérben, mint azokat, amelyekben közelebb helyezkedtek el (LIU–KUHL–TSAO 2003; SONG–DEMUTH–MORGAN 2010), tehát a nagyobb magánhangzótér az alacsonyabb tempó mellett szintén hozzájárulhat a könnyebb beszédfeldolgozáshoz. Továbbá a dajkanyelv magánhangzótere összefüggést mutatott a gyerekek szókincsével, hiszen egy másik kutatás eredményei szerint a 9 hónapos babákhoz szóló dajkanyelv magánhangzóterének adataiból meg lehetett jósolni a gyerekek későbbi, 15 és 19 hónapos kori expresszív szókincsének méretét (KALASHNIKOVA–BURNHAM 2018). Egy harmadik kísérletben kimutatták, hogy a másfél éves gyerekhez szóló dajkanyelv magánhangzótere nemcsak a két éves gyerekek expresszív szókincsével függött össze, hanem a **receptív szókinccsel**, azaz a megértett szavak számával is (HARTMAN–BERNSTEIN RATNER–NEWMAN 2017). Ezek az eredmények arra engednek következtetni, hogy a dajkanyelvi nagyobb magánhangzótér elősegíti a szavak felismerését, a beszéd feldolgozását és hosszú távon a szókincs növekedését. A szókinccsel kapcsolatos kísérletek azonban angol dajkanyelvet vizsgáltak, nem tudjuk, hogy vajon más nyelvekben is hasonló összefüggések találhatók-e.

Az alapfrekvenciával kapcsolatos dajkanyelvi sajátosságok esetében többször felmerült, hogy az alapfrekvenciában megjelenő nagyobb variabilitás, a mérhető különbségek a szótagokon esetleg segítenék a szavak határainak a felismerését és ezzel a nyelvelsajátítást (SPINELLI–FASOLO–MESMAN 2017). Ennek a hipotézisnek az alátámasztására meglehetősen kevés kísérletileg igazolt eredmény áll rendelkezésre. Egy vizsgálatban azt találták, hogy a három hónapos babákhoz szóló beszéd nagyobb hangterjedelme összefüggésben van a csecsemő 12 hónapos kori expresszív szókincsének méretével (PORRITT et al. 2014). Annak megértéséhez, hogy a dajkanyelv prozódiai sajátosságai pontosan milyen szerepet töltenek be a nyelvelsajátításban, még további vizsgálatok szükségesek.

### 3.4. A dajkanyelv mint beszédalkalmazkodás

A dajkanyelv a beszédalkalmazkodásnak, azaz a beszédünk hallgatóhoz való igazításának egyik tipikus példája. A felnőttek másképp beszélnek a gyerekekhez, mint a felnőttekhez, ami számtalan akusztikai jegyben megnyilvánul (SPINELLI–FASOLO–MESMAN 2017). Felmerül a kérdés, hogy miért éppen ezek az akusztikai sajátosságai a dajkanyelvnek, és hogyan lehetséges, hogy több tulajdonság is univerzálisnak tekinthető.

A gyerekek azokat a beszédrészleteket részesítik előnyben, amelyek a saját beszédükhöz hasonlóbbak. 4 és 6 hónap közötti babák hosszabban figyeltek olyankor, amikor a szintetizált magánhangzó alapfrekvenciája a saját ejtésmódjukhoz közelebb volt, azaz magasabb

frekvenciaértékekkel rendelkezett (MASAPOLLO–POLKA–MENARD 2016). Egy másik vizsgálatban 5 és 7 hónap közötti csecsemőket vizsgálva pedig azt találták, hogy a babák hosszabban figyeltek a sajátjukéhoz hasonló formánsszerkezetű magánhangzókra, mint a felnőttekéhez hasonló formánsszerkezetű magánhangzókra (POLKA–MASAPOLLO–MENARD 2020). Az egyéves kor alatti gyerekekkel kapcsolatban tehát kísérletekkel is alátámasztották, hogy a sajátjukéhoz hasonló alaphangfrekvenciájú és hasonló formánssértékekkel rendelkező magánhangzókat részesítik előnyben a felnőttekéhez hasonló beszéddel szemben.

Az interakciók során a beszélgetőpartnerek hajlamosak imitálni a másik beszédének nyelvi és akusztikai sajátosságait, vagy legalábbis közelíteni a sajátjukat a partneréhez. Ezt a jelenséget **mimikrinek** is szokás nevezni (BERNHOLD–GILES 2020). Ha egy felnőtt beszélget egy gyerekkel, akkor ott is feltételezhetjük, hogy utánozza az interakcióban lévő másik fél beszéd-sajátosságait. Ezen megközelítés szerint mivel a gyerekek zöngéjének alaphangfrekvenciája – hangszalagjaik kisebb méretei miatt – magasabb, ezért a felnőttek is magasabb alaphangfrekvenciával beszélnek a gyerekekhez. Habár a dajkanyelv nagyobb magánhangzóterével kapcsolatos akusztikai eredmények meglehetősen ellentmondásosak, láthattuk, hogy az egyik magyarázat szerint ez az eltérés artikulációs szempontból visszavezethető a kisebb toldalékcsozománya, amelynek következtében a gyerekekéhoz hasonló magánhangzóter jön létre. Ily módon ez a tulajdonság is magyarázható lenne azzal, hogy a felnőttek megpróbálják megközelíteni a gyerekek beszédének akusztikai sajátosságait (KALASHNIKOVA–CARIGNAN–BURNHAM 2017). Az is ismert tény, hogy a gyerekek tempója lassabb, mint a felnőtteké (PAYNE 2012), így ez esetben is magyarázatként szolgálhatna a mimikri. Ugyanakkor nem kell jelen lennie egy gyereknek ahhoz, hogy a felnőttek produkálják a dajkanyelvi jegyeket. Az édesanyák a baba meg nem születése előtt, a prenatális korban is produkálják a dajkanyelvre jellemző akusztikai sajátosságokat. Továbbá a gyerekekkel nem rendelkezők is képesek módosítani a beszédüket, ráadásul egy fényképhez beszélve, egy elképzelt szituációban is előidézhető ez a beszédregiszter (KO–SODERSTROM 2013). Elképzeltető lenne, hogy élettapasztalatunk alapján tudjuk, vagy legalábbis tudni véljük, hogy a gyerekek hogyan beszélnek, és ehhez alkalmazkodnánk? Ennek a kérdésnek a megválaszolásához még további kísérleti eredményekre van szükség, mindenesetre jelen tudásunk szerint a mimikri is motiválhatja a dajkanyelvi sajátosságok használatának egy részét.

## 4. A dajkanyelv tulajdonságainak és funkcióinak összefoglalása

A jelen tanulmányban a dajkanyelv akusztikai sajátosságait tekintettük át részletesebben, illetve ezeknek lehetséges szerepét a gyerek-felnőtt interakcióban. A dajkanyelv meghatározó tulajdonságai közé sorolható a lassabb tempó, a magasabb alaphangfrekvencia és a nagyobb hangterjedelem, valamint több nyelvben a nagyobb magánhangzóter. A dajkanyelvnek nemcsak

figyelemfelkeltő, figyelemfenntartó, valamint érzelmközvetítő szerepe van a kommunikációban, hanem segíti a beszédfeldolgozást és a nyelvelsajátítást is. Habár már nagyon sok mindent tudunk, illetve tudni vélünk a dajkanyelvről, számtalan nyitott kérdés még megválaszolatlan. Például az, hogy miért éppen ezek a dajkanyelv tulajdonságai; hogy longitudinálisan (a baba életkorának előrehaladtával) hogyan alakulnak a különböző tulajdonságok; hogy hogyan lassítják beszédüket az édesanyák; hogy mi magyarázhatja a nyelvek közti eltéréseket a különböző tulajdonságok esetében; és még sorolhatnánk a nyitott kérdéseket, amelyek a jövőben válaszra várnak.

## FOGALMAK

*beszédalkalmazkodás; dajkanyelv; artikulációs tempó; beszédtempó; frázisvégi nyúlás; mássalhangzós szakasz; magánhangzós szakasz; beszédritmus; prozódia; alappfrekvencia; dallamhangsúly; magánhangzótér; formáns; zöngeminőség; leheletes zöngé; a dajkanyelv funkciói; érzelmközvetítés; expresszív és receptív szókincs; mimikri*

## IRODALOM

- ANDRUSKI, Jean E. – KUHL, Patricia K. – HAYASHI, Akiko 1999. Point vowels in Japanese mothers' speech to infants and adults. *The Journal of the Acoustical Society of America* 105/2. 1095–1096.
- BENDERS, Titia 2013. Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development* 36/4. 847–862.
- BERNHOLD, Quinten S. – GILES, Howard 2020. Vocal accommodation and mimicry. *Journal of Nonverbal Behavior* 44/1. 41–62.
- BRANNON, Elizabeth M. – LIBERTUS, Melissa E. – MECK, Warren H. – WOLDORFF, Marty G. 2008. Electrophysiological measures of time processing in infant and adult brains: Weber's Law holds. *Journal of Cognitive Neuroscience* 20/2. 193–203.
- BROESCH, Tanya L. – BRYANT, Gregory A. 2015. Prosody in infant-directed speech is similar across western and traditional cultures. *Journal of Cognition and Development* 16/1. 31–43.
- BROESCH, Tanya L. – BRYANT, Gregory A. 2018. Fathers' infant-directed speech in a small-scale society. *Child Development* 89/2. e29–e41.
- CHLÁDKOVÁ, Kateřina – ČERNÁ, Martina – PAILLEREAU, Nikola – SKARNITZL, Radek – OCELÁKOVÁ, Zuzana 2019. Prenatal infant-directed speech: vowels and voice quality. In CALHOUN, Sasha – ESCUDERO, Paola – TABAIN, Marija – WARREN, Paul (eds): *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)*. Australasian Speech Science and Technology Association Inc., Canberra. 1525–1529.
- COOPER, Robyn P. 1993. The effect of prosody on young infants' speech perception. *Advances in Infancy Research* 8. 137–167.
- COOPER, Robyn P. – ABRAHAM, Jane – BERMAN, Sheryl – STASKA, Margaret 1997. The development of infants' preference for motherese. *Infant Behavior and Development* 20/4. 477–488.

- CRISTIÀ, Alejandrina 2010. Phonetic enhancement of sibilants in infant-directed speech. *The Journal of the Acoustical Society of America* 128/1. 424–434.
- CRISTIÀ, Alejandrina – SEIDL, Amanda 2014. The hyperarticulation hypothesis of infant-directed speech. *Journal of Child Language* 41/4. 913–934.
- DE PALMA, Paul – VANDAM, Mark 2017. Using automatic speech processing to analyze fundamental frequency of child-directed speech stored in a very large audio corpus. In *2017 Joint 17th World Congress of International Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems (IFSAS-SCIS)*. IEEE. 1–6.
- DEME Andrea 2016. *Magánhangzók ejtése és észlelése a szopránéneklésben*. ELTE Eötvös Kiadó, Budapest.
- DEME Andrea – KOHÁRI Anna – REICHEL, Uwe D. – SZALONTAI Ádám – MÁDY Katalin 2019. A magánhangzós hosszúsági fonológiai kontraszt a dajkanyelvben a csecsemő életkorának függvényében. *Beszéd kutatás* 27. 221–242.
- DODANE, Christelle – AL-TAMIMI, Jaleddin 2007. An acoustic comparison of vowel systems in adult-directed-speech and child-directed speech: Evidence from French, English and Japanese. In TROUVAIN, Jürgen – BARRY, William J. (eds): *16th International Congress of Phonetics Sciences*. Pirrot GmbH, Dudweiler. 6–10.
- DUNST, Carl – GORMAN, Ellen – HAMBY, Deborah 2012. Preference for infant-directed speech in preverbal young children. *Center for Early Literacy Learning* 5/1. 1–13.
- ENGLUND, Kjellrun – BEHNE, Dawn 2006. Changes in infant directed speech in the first six months. *Infant and Child Development: An International Journal of Research and Practice* 15/2. 139–160.
- FERNALD, Anne – KUHL, Patricia 1987. Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development* 10/3. 279–293.
- FERNALD, Anne – SIMON, Thomas 1984. Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology* 20/1. 104.
- FERNALD, Anne – TAESCHNER, Traute – DUNN, Judy – PAPOUSEK, Mechthild – DE BOYSSON-BARDIES, Bénédicte – FUKUI, Ikuko 1989. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language* 16/3. 477–501.
- GARELLEK, Mark 2019. The phonetics of voice. In KATZ, William F. – ASSMANN, Peter F. (eds): *The Routledge Handbook of Phonetics*. Routledge, New York. 75–106.
- GERGELY, Anna – FARAGÓ, Tamás – GALAMBOS, Ágoston – TOPÁL, József 2017. Differential effects of speech situations on mothers' and fathers' infant-directed and dog-directed speech: An acoustic analysis. *Scientific Reports* 7/1. 1–10.
- GÓSY Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- GRABE, Esther – LOW, Ee Ling 2002. Durational variability in speech and the Rhythm Class Hypothesis. *Papers in Laboratory Phonology* 7. 515–546.
- GRIESER, DiAnne L. – KUHL, Patricia K. 1988. Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology* 24/1. 14–20.

- HÁDEN, Gábor P. – MÁDY, Katalin – TÖRÖK, Miklós – WINKLER, István 2020. Newborn infants differently process adult directed and infant directed speech. *International Journal of Psychophysiology* 147. 107–112.
- HAN, Mengru – DE JONG, Nivja H. – KAGER, René 2018. Infant-directed speech is not always slower: Cross-linguistic evidence from Dutch and Mandarin Chinese. In BERTOLINI, Anne B. – KAPLAN, Maxwell J. (eds): *Proceedings of the 42nd annual Boston University Conference on Language Development*. Cascadilla Press, Somerville. 331–344.
- HARTMAN, Kelly M. – BERNSTEIN RATNER, Nan – NEWMAN, Rochelle S. 2017. Infant-directed speech (IDS) vowel clarity and child language outcomes. *Journal of Child Language* 44/5. 1140–1162.
- HOEQUIST, Charles 1983. Syllable duration in stress-, syllable and mora-timed languages. *Phonetica* 40. 203–237.
- KALASHNIKOVA, Marina – BURNHAM, Denis 2018. Infant-directed speech from seven to nineteen months has similar acoustic properties but different functions. *Journal of Child Language* 45/5. 1035–1053.
- KALASHNIKOVA, Marina – CARNIGAN, Christopher – BURNHAM, Denis 2017. The origins of babytalk: smiling, teaching or social convergence? *Royal Society Open Science* 4/8. 170306. <https://doi.org/10.1098/rsos.170306>
- KO, Eon-Suk 2012. Nonlinear development of speaking rate in child-directed speech. *Lingua* 122/8. 841–857.
- KO, Eon-Suk – SODERSTROM, Melanie 2013. Additive effects of lengthening on the utterance-final word in child-directed speech. *Journal of Speech, Language, and Hearing Research* 56/1. 364–371.
- KOHÁRI Anna 2018. *Időzítési mintázatok a magyar beszédben*. ELTE Eötvös Kiadó, Budapest.
- KOHÁRI Anna – DEME Andrea – REICHEL Uwe D. – SZALONTAI Ádám – MÁDY Katalin 2019. A dajkanyelv temporális jellemzői 4 és 8 hónapos csecsemőkhöz szóló beszédben. *Beszéd kutatás* 27. 243–258.
- KOHÁRI Anna – DEME Andrea – REICHEL Uwe D. – SZALONTAI Ádám – MÁDY Katalin 2020. A beszédritmus időbeli dimenziójának jellegzetességei a dajkanyelvben. In BÓNA Judit – KREPSZ Valéria (szerk.): *Nyelvfejlődés csecsemőkortól kamaszkorig*. ELTE Eötvös Kiadó, Budapest. 195–208.
- KOHÁRI, Anna – MÁDY, Katalin – DEME, Andrea – REICHEL, Uwe D. – SZALONTAI, Ádám 2019. Utterance-final lengthening in infant-directed speech. In FIVELA, Barbara Gili – AVESANI, Cinzia – SAVINO, Michelina – D'APOLITO, Sonia – LORUSSO, Paolo – GRIMALDI, Mirko – SIGONA, Francesco (eds): *3rd Phonetics and Phonology in Europe (PaPE)*. LabPhon – Fondazione Puglia – AISV, Lecce. 196–197.
- KOHÁRI, Anna – REICHEL, Uwe D. – MARKÓ, Alexandra – MÁDY, Katalin 2020. Breathiness in speech directed to 4-month-old infants. In TIEDE, Mark – WHALEN, Doug – GRACCO, Vince (eds): *12th International Seminar on Speech Production*. Haskins Laboratories – Yale University, New Haven, CT. 1–2.

- KUHL, Patricia K. – ANDRUSKI, Jean E. – CHISTOVICH, Inna A. – CHISTOVICH, Ludmilla A. – KOZHEVNIKOVA, Elena – RYSKINA, Viktoria – STOLYAROVA, Elvira I. – SUNDBERG, Ulla – LACERDA, Francisco 1997. Cross-language analysis of phonetic units in language addressed to infants. *Science* 277/5326. 684–686.
- LADEFOGED, Peter 1975. *A Course in Phonetics*. Harcourt Brace Jovanovich, New York.
- LEE, Christopher S. – KITAMURA, Christine – BURNHAM, Denis. – MCANGUS TODD, Neil P. 2014. On the rhythm of infant-versus adult directed speech in Australian English. *The Journal of the Acoustical Society of America* 136/1. 357–365.
- LIU, H. M. – KUHL, Patricia K. – TSAO, Feng-Ming 2003. An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science* 6/3. F1–F10.
- MA, Weiyi – GOLINKOFF, Roberta M. – HOUSTON, Derek M. – HIRSH-PASEK, Kathy 2011. Word learning in infant-and adult-directed speech. *Language Learning and Development* 7/3. 185–201.
- MÁDY, Katalin – REICHEL, Uwe D. – SZALONTAI, Ádám – KOHÁRI, Anna – DEME, Andrea 2018. Prosodic characteristics of infant-directed speech as a function of maternal parity. In KLESSA, Katarzyna – BACHAN, Jolanta – WAGNER, Agnieszka – KARPIŃSKI, Maciej – ŚLEDZIŃSKI, Daniel (eds): *Proceedings of the 9th International Conference on Speech Prosody 2018*. International Speech Communication Association (ISCA), Poznań, Poland. 294–298.
- MANYBABIES CONSORTIUM 2020. Quantifying sources of variability in infancy research using the infant-directed-speech preference. *Advances in Methods and Practices in Psychological Science* 3/1. 24–52.
- MARKÓ Alexandra 2013. *Az irreguláris zöngé funkciói a magyar beszédben*. ELTE Eötvös Kiadó, Budapest.
- MARTIN, Andrew – IGARASHI, Yosuke – JINCHO, Nobuyuki – MAZUKA, Reiko 2016. Utterances in infant-directed speech are shorter, not slower. *Cognition* 156. 52–59.
- MASAPOLLO, Matthew – POLKA, Linda – MENARD, Lucie 2016. When infants talk, infants listen: Pre-babbling infants prefer listening to speech with infant vocal properties. *Developmental Science* 19/2. 318–328.
- MCMURRAY, Bob – KOVACK-LESH, Kristine A. – GOODWIN, Dresden – MCCECHRON, William 2013. Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence? *Cognition* 129/2. 362–378.
- MCCLOY, Daniel R. 2014. Phonetic effects of morphological structure in Indonesian vowel reduction. *Proceedings of Meetings on Acoustics* 12/1. 1–14.
- MIYAZAWA, Kouki – SHINYA, Takahito – MARTIN, Andrew – KIKUCHI, Hideaki – MAZUKA, Reiko 2017. Vowels in infant-directed speech: More breathy and more variable, but not clearer. *Cognition* 166. 84–93.
- NARAYAN, Chandan R. – McDERMOTT, Lily C. 2016. Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. *The Journal of the Acoustical Society of America* 139/3. 1272–1281.

- NAZZI, Thierry – BERTONCINI, Josiane – MEHLER, Jacques 1998. Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance* 24/3. 756–766.
- PAYNE, Elinor – POST, Brechtje – ASTRUC, Lluïsa – PRIETO, Pilar – VANRELL, Maria del Mar 2012. Measuring child rhythm. *Language and Speech* 55/2. 203–229.
- PAYNE, Elinor M. – POST, Brechtje – ASTRUC, Lluïsa – PRIETO, Pilar – VANRELL, Maria del Mar 2009. Rhythmic modification in child directed speech. *Oxford University Working Papers in Linguistics. Philology & Phonetics* 12. 123–144.
- PEGG, Judith E. – WERKER, Janet F. – MCLEOD, Peter J. 1992. Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development* 15/3. 325–345.
- PICKERING, Martin J. – GARROD, Simon 2006. Alignment as the basis for successful communication. *Research on Language and Computation* 4/2–3. 203–228.
- PIKE, Kenneth 1945. *The Intonation of American English*. University of Michigan Press, Ann-Arbor.
- POLKA, Linda – MASAPOLLO, Matthew – MENARD, Lucie 2014. Who's talking now? Infants' perception of vowels with infant vocal properties. *Psychological Science* 25/7. 1448–1456.
- POLKA, Linda – MASAPOLLO, Matthew – MENARD, Lucie 2020. Setting the stage for speech production: Infants prefer listening to speech sounds with infant vocal properties. In TIEDE, Mark – WHALEN, Doug – GRACCO, Vince (eds): *12th International Seminar on Speech Production*. Haskins Laboratories – Yale University, New Haven, CT. 1–2.
- PORRITT, Laura L. – ZINSER, Michael C. – BACHOROWSKI, Jo-Anne A. – KAPLAN, Peter S. 2014. Depression diagnoses and fundamental frequency-based acoustic cues in maternal infant-directed speech. *Language Learning and Development* 10/1. 51–67.
- QUIGLEY, Jean – NIXON, Elizabeth – LAWSON, Sarah 2019. Exploring the association of infant receptive language and pitch variability in fathers' infant-directed speech. *Journal of Child Language* 46/4. 800–811.
- RAMUS, Franck – NESPOR, Marina – MEHLER, Jacques 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 75/1. 1–28.
- RANERI, Daniele – VON HOLZEN, Katie – NEWMAN, Rochelle – BERNSTEIN RATNER, Nan 2020. Change in maternal speech rate to preverbal infants over the first two years of life. *Journal of Child Language* 47/6. 1263–1275.
- RÉGER Zita 2002. *Utak a nyelvhez*. Soros Alapítvány – MTA Nyelvtudományi Intézet, Budapest.
- ROACH, Peter 1982. On the distinction between 'stress-timed' and 'syllable-timed' languages. In CRYSTAL, David (ed.): *Linguistic Controversies*. Edward Arnold, London. 73–79.
- ROACH, Peter 2002. *A Little Encyclopaedia of Phonetics*. University College Dublin, Dublin.
- SAINT-GEORGES, Catherine – CHETOUANI, Mohamed – CASSEL, Raquel – APICELLA, Fabio – MAHDHAOUI, Ammar – MURATORI, Filippo – LAZNIK, Marie-Christine – COHEN, David 2013. Motherese in interaction: At the cross-road of emotion and cognition? (A systematic review). *Plos One* 8/10. e78103. <https://doi.org/10.1371/journal.pone.0078103>

- SAITO, Yasuhiro – AOYAMA, Shinya – KONDO, Takeo – FUKUMOTO, Rie – KONISHI, Nakao – NAKAMURA, Kazuhiro – KOBAYASHI, Masao – TOSHIMA, Tamotsu 2007. Frontal cerebral blood flow change associated with infant-directed speech. *Archives of Disease in Childhood-Fetal and Neonatal Edition* 92/2. F113–F116.
- SCHERER, Klaus R. 2003. Vocal communication of emotion: A review of research paradigms. *Speech Communication* 40/1–2. 227–256.
- SEGAL, Judith – NEWMAN, Rochelle S. 2015. Infant preferences for structural and prosodic properties of infant-directed speech in the second year of life. *Infancy* 20/3. 339–351.
- SENJU, Atsushi – CSIBRA, Gergely 2008. Gaze following in human infants depends on communicative signals. *Current Biology* 18/9. 668–671.
- SHUTE, Brenda – WHELDALL, Kevin 1999. Fundamental frequency and temporal modifications in the speech of British fathers to their children. *Educational Psychology* 19/2. 221–233.
- SINGH, Leher – MORGAN, James L. – BEST, Catherine T. 2002. Infants' listening preferences: Baby talk or happy talk? *Infancy* 3/3. 365–394.
- SONG, Jae Yung – DEMUTH, Katherine – MORGAN, James 2010. Effects of the acoustic properties of infant-directed speech on infant word recognition. *The Journal of the Acoustical Society of America* 128/1. 389–400.
- SPINELLI, Maria – FASOLO, Mirco – MESMAN, Judi 2017. Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review* 44. 1–18.
- SWANSON, Lori A. – LEONARD, Laurence B. – GANDOUR, Jack 1992. Vowel duration in mothers' speech to young children. *Journal of Speech, Language, and Hearing Research* 35/3. 617–625.
- SZALONTAI, Ádám – MÁDY, Katalin – DEME, Andrea – KOHÁRI, Anna 2018. Prosodic boundaries in Hungarian infant-directed speech. In *Proceedings Experimental and Theoretical Advances in Prosody (ETAP) 4*. University of Massachusetts, Amherst, USA. 53. [http://real.mtak.hu/89350/1/ETAP4\\_paper\\_53.pdf](http://real.mtak.hu/89350/1/ETAP4_paper_53.pdf) (A letöltés ideje: 2021. február 8.)
- THANAVISUTH, Chayada – LUKSANEYANAWIN, Sudaporn 1998. Acoustic qualities of IDS and ADS in Thai. In *Fifth International Conference on Spoken Language Processing*. ISCA, Sydney, Australia. 445–448.
- WANG, Yuanyuan – LLANOS, Fernando – SEIDL, Amanda 2017. Infants adapt to speaking rate differences in word segmentation. *The Journal of the Acoustical Society of America* 141/4. 2569–2578.

## KÖSZÖNETNYILVÁNÍTÁS

A tanulmány az „A dajkanyelv longitudinális vizsgálata multimodális módszerekkel” című pályázat (NKFI-134775) keretében készült. Ez az írás nem jöhetett volna létre Vincze Benedek Gergő és Vincze Miklós inspiráló javaslatai, valamint Mády Katalin segítségével és támogatása nélkül.





# Egyéni variabilitás a beszédben

Gráczki Tekla Etelka

NYELVTUDOMÁNYI KUTATÓKÖZPONT  
MTA–ELTE „LENDÜLET” LINGVÁLIS ARTIKULÁCIÓ KUTATÓCSOPORT

Krepsz Valéria

NYELVTUDOMÁNYI KUTATÓKÖZPONT

Huszár Anna

NYELVTUDOMÁNYI KUTATÓKÖZPONT

## 1. Bevezetés

Mi változik tíz év alatt egy ember életében? Nyilvánvalóan a válasz attól is függ, milyen életciklusában tekintjük az embert. Egy újszülöttkortól induló első 10 életév vizsgálata nagyon más válaszokat ad erre a kérdésre, mint egy 80 éves kortól induló elemzés, és mindkettőtől merőben különböző eredményekre juthatunk akkor, ha fiatal felnőtteket veszünk kutatásunk tárgyául. A válaszok attól is függenek, milyen szempontokat választunk ki. Saját területünket, a beszédet figyelembe véve ugyanezen életkorok esetében az első csoportban az anyanyelv elsajátításáról és korai fejlődéséről tehetünk megállapításokat, a második életkorban az időskori beszédbeli változásokról, míg azt gondolhatnánk, hogy a fiatalabb felnőttek vizsgálatában nem találunk lényeges eltéréseket. Mégis érdemes néhány szempont mentén végiggondolni, mi minden és hogyan befolyásolja a beszédünket ebben az életszakaszban.

Az életkorok mentén talált eltérések nem egyik pillanatról a másikra lépnek fel – esetleg egyes patológiás esetektől eltekintve –, hanem fokozatosan nyomon követhetőek. Az, hogy az életkori csoportok között mely ponton milyen mértékben változnak az értékek, érdekes vizsgálati téma (l. FOUGERON 2020). Voltak olyan jellemzők, amelyeknél lineáris változást találtak a 30–39 év közötti életkorig, majd stagnálást, esetleg később további változást, de olyanok is, ahol a fiatal és középkorú felnőttek között nem volt eltérés vagy kimutatható tendencia, míg az 50-es vagy 60-as éveikben járó beszélőkhöz képest már különbséget találtak. A mi kérdéseink arra vonatkoznak, hogy fiatal vagy középkorú beszélők összevetésével és ugyanazon beszélők 10 év utáni nyomon követésével milyen mintázatokat tudunk feltérképezni az egyes beszédjellemzőkben.

## 2. Az öregedés

Ahhoz, hogy a beszéd életkori csoportok közötti eltéréseiről vagy az egyénben végbemenő változásokról beszélhessünk, figyelembe kell vennünk az öregedés vagy idősödés folyamatát.

A WHO beosztása szerint a 60–74 évesek korcsoportja alkotja az idősödők kategóriáját. Efölött vannak az idős (75–89), az agg (90–99) és a matuzsálem (100 év felett) életkori csoportok. Az idősödés, öregedés ugyanakkor nem egy adott életkorban lép fel, hanem egy folyamat, mégpedig egy permanens, visszafordíthatatlan, progresszív folyamat. Nemcsak a kronológiai életkor határozza meg, azaz a születéstől eltelt idő, hanem az egyes személyek egyéb tulajdonságai, szokásai, az őket érő hatások is. A folyamat korán, 20–30 éves korban megkezdődik. Tulajdonképpen a fogantatástól folyamatos változással írható le a szervezetünk, és az élettani, fiziológiás funkciókban 20–30 éves kor között már megjelenhet romlás, gyengülés, azaz öregedés. Természetesen a folyamat nagyon eltérő az egyének között, sőt az egyén egyes szervei vagy egy szervének funkciói között is találhatunk különbséget, ugyanakkor minden elem öregedése kihat a rendszer többi elemére is (SEMSEI 2014).

Beszélhetünk kronológiai életkorról és biológiai életkorról (BELSKY et al. 2015). A **kronológiai életkor** az, amennyi idő eltelt születésünk óta, míg a **biológiai életkor** a szervezetünk funkcióinak állapotával írható le. BELSKY és munkatársai (2015) 954 fiatal felnőtt öregedését vizsgálták. Két alapvetésből indultak ki. Egyrészt abból, hogy az idősödés kutatásának egyik paradoxona, hogy általában időseken vizsgálják a fiziológiás változásokat, miközben náluk már gyakori, hogy krónikus állapot hatása is megjelenik. Másrészt arra hívták fel a figyelmet, hogy az öregedés folyamat, nemcsak időseken tanulmányozható, hanem a fiatalabbakon is tetten érhető, hiszen korán megindul. Kutatásukban ugyanazon személyeket vizsgálták 26, 32 és 38 éves korukban; 18 élettani funkciót mértek (pl. fehérvérsejtek száma, nitrogén a vizeletben, kardio-respirációs állapot, teljes koleszterinszint, átlagos artériás nyomás). Ezek összegzésével azt állapították meg, hogy a kronológiailag 38 éves felnőttek biológiai életkora normál eloszlású, és mintegy 28 éves és 50 éves kor közé esett. Az azonos (kronológiai) életkorú, 38 éves, de magasabb biológiai életkort mutató felnőttek esetében az öregedés tempója gyorsabb volt, vagyis a vizsgált élettani funkciók nagyobb mértékben romlottak azonos/hasonló korábbi (12 és 24 évvel korábbi) értékekről, mint a fiatalabb biológiai életkort mutatóké.

Az egyéni variabilitás egyik fő oka a személyeket érő eltérő hatások: az idősödésben meg kell különböztetni háromféle folyamatot. Az **elsődleges öregedés** a kronológiai életkor következtében fellépő változás, a **másodlagos öregedés** a patológiai folyamatok következményeképp fellépő változás, míg a **harmadlagos öregedés** a környezeti tényezők okozta változások (LAMY 1991).

A korai (fiatal, középkorú) életkorban kimutatható változásokból néhány példát említve: a tüdőerek merevebbé válása már 30–35 éves kor után, a tüdőkapacitás csökkenése kb. 30 éves kortól adatolható (összefoglalás: LALLEY 2013). Az izmok és a szervek 40 éves kortól veszítenek tömegükből (AMARYA et al. 2018). Az agy térfogatának maximuma a 20-as évekre tehető, utána csökkenni kezd, a 40-es évektől az agykéreg zsugorodik, a személyek memóriacsökkenést tapasztalhatnak (AMARYA et al. 2018).

A biológiai változások hathatnak a beszédre. A beszédet befolyásoló fiziológiás változásokat részletesen ismertette BALÁZS és BÓNA (2016). Ilyen a tüdőkapacitás csökkenése és ebből kifolyólag a gyakoribb légvétel (pl. HUBER 2008), a szív munkájának romlása, a renyhébben záró és kevésbé rugalmas hangszalagok, a meszesedő porcok, a (nemi) hormonok alultermelődése, a nehezített nyelvmozgás, a rugalmatlanabb légyszájpad-mozgás, a fogvesztés (és a fogpótlások).

Természetesen nem minden mértékű biológiai változás von maga után a beszédben mérhető, tetten érhető következményeket, főleg a fiatalabb életkorokban, ugyanakkor az egészséges beszélőkre felállítható tendenciák segíthetik a folyamat megértését és az esetleges patológiás határok, tehát már betegségre utaló értéktartományok felállítását.

### 3. A beszéd életkori sajátosságainak vizsgálata

A beszédet érintő időbeli változások vizsgálatának az egyik legfontosabb kérdése, hogy a különböző változások hogyan mutatkoznak meg a beszéddel kapcsolatos területeken az életkor előrehaladtával. Az idősödés során ugyanis számos változás zajlik mind a beszédben, mind pedig a hangjellelmzőkben: a fent említett fiziológiai, anatómiai, respirációs, hormonális módosulások, a motoros kontroll változásai, a kognitív funkciók módosulása (memória, alkalmazkodás, figyelem stb.) mellett az életút és a beszédhasználat is meghatározó, amelyek többek között a hivatásban és a szocializációban mutatkoznak meg (BROWN 1996; KETCHAM et al. 2002).

Az életkor előrehaladtával bekövetkező beszédbeli változások vizsgálata nagy érdeklődésre tart számot. Az eredmények nemcsak a fonetika számára, de a beszélőazonosításban, az egészséges és patológiás változások közötti határvonal(-sáv) meghatározásában is fontos adatokat szolgáltathatnak.

Az idő mint vizsgálati paraméter elemzésére két különböző megközelítés létezik. Az első az úgynevezett **keresztmetszeti** (*cross sectional*) **vizsgálatok** köre, amelyek az egy adott időpontban (vagy egymáshoz közeli időpontokban), valamint az életút különböző fázisaiban lévő egyének vizsgálatát jelentik. Az egy csoportba tartozó adatközlők általában egykorúak, vagy kis életkori eltérés van közöttük, illetve az egyéb vizsgálati szempontok szerint (például nem, egészségi állapot stb.) homogén csoportot alkotnak. Azt, hogy mi a „kis életkori eltérés”, egyrészt a mindenkori vizsgálati kérdés, másrészt az egyéb vizsgált életkorok / életkori csoportok határozzák meg. Szükségszerű, hogy a csoportbeli variabilitás kisebb legyen, mint a csoportok közötti. Tehát úgy kell kialakítani a csoportokat, hogy a fiatalabb csoport legidősebbje és az idősebb csoport legfiatalabbja között jelentősen nagyobb eltérés legyen, mint egyazon csoport legidősebbje és legfiatalabbja között (természetesen ezen túlmenően sem célszerű az egy csoporton belüli nagy variabilitás). További lehetőség folyamatos skálával dolgozni, ahol nem előre kialakított életkori csoportok vannak, hanem a teljes vizsgált tartományon belül ordinális (sorrendi) skálaként vehető figyelembe az életkor.

Ezzel szemben a **longitudinális** vagy más elnevezés szerint a **követéses/hosszmetszeti vizsgálat** az egyéni változásokat elemzi az idő és az adatközlők életkorának előrehaladtával.

Ez utóbbi esetben tehát az adatközlő vagy az adatközlők csoportja minden mérés esetében azonos, jellemzőik vizsgálata pedig rövidebb vagy hosszabb időintervallumon keresztül, adott időközönként történik.

A két eljárás azonban nemcsak módszertanában, de célkitűzésében, az eredmények sajátosságaiban és a levonható következtetések tekintetében is különbözik egymástól. A keresztmetszeti vizsgálatokban a különböző életkorokat különböző beszélők képviselik, ezért ezekben az eredmények a beszélők közötti különbségeket mutatják be. Ezzel szemben a követéses kutatásokban az egyénen belüli különbségek vizsgálata és leírása történik, azaz a vizsgált életkorokban ugyanazon beszélőktől kapott adatokat vetnek össze. Így tehát fontos különbség, hogy míg a hosszmetzeti eljárásnál magát a változást tudjuk leírni, a keresztmetszeti eljárással végzett mérések során az életkorral bekövetkező változásokra közvetve tudunk következtetni (tehát nem magának a változásnak a közvetlen elemzése történik) (vö. pl. BODZSÁR–ZSÁKAI 2013).

Mindkét vizsgálati eljárás során figyelembe kell venni, hogy milyen hatás érheti a beszélőket. A longitudinális vizsgálatok egyes mérési pontjai között (ezek lehetnek napok, hetek, hónapok, évek, de akár évtizedek is) olyan egyéni hatások érhetnek az adatközlőket, amelyek befolyásolják a vizsgálatban mért paramétereket. Ezáltal nehezíthetik az egyes beszélők összevetését egymással (különböző betegségek, káros szenvedélyekre való rá- vagy épp leszokás, költözés vagy a közvetlen környezet összetételének megváltozása stb.). Emellett szintén nehézséget jelent, hogy az esetlegesen detektált változások kulturális, nemi, pszichoszociális tényezőkkel, vagy pedig valóban az idősödő hang sajátosságainak változásával magyarázhatóak-e. A keresztmetszeti elemzéseknél ugyanakkor előfordulhatnak bizonyos szisztematikus vagy véletlenszerűen előforduló hibák. Ide sorolható például a **kohorszeffektus** megjelenése. Ez olyan hatást jelent, amely a keresztmetszeti vizsgálatok esetében csupán egyetlen adatközlői csoportot (életkori kategóriát) érint. Ilyen lehet például, ha egy korosztály tapasztalatait és ismereteit meghatározza az átélte háború vagy esetleg a nyelvhasználatra vonatkozó törvényi szabályozás.

Az individuális eltérések ugyanakkor egy-egy homogénnek tekintett, azonos nemi és életkori csoporton belül is meghatározók lehetnek, akár keresztmetszeti, akár hosszmetzeti kutatásról van szó. Ezeknek az esetleges okaira pedig figyelni kell az elemzések során.

A longitudinális felvételek rögzítésével kapcsolatos nehézség, hogy a követéses vizsgálatok elvégzése hosszú távú elköteleződést jelent mind a vizsgálatban részt vevők, mind a kutatók részéről a témakör és a vizsgálati paradigma iránt (az egyes vizsgálatoknál a felvételek elkészítése és elemzése akár hónapokon vagy éveken keresztül is történhet). A másik lehetőség az utólagosan kialakított longitudinális módszertan, amikor egy korábbi – eredetileg nem követéses vizsgálati paradigmával megtervezett – vizsgálatot bizonyos idő elteltével (közel) azonos körülmények között újra elvégzünk. A beszédhez kapcsolódó elemzésekben mind a követéses, mind a keresztmetszeti jellegű vizsgálat mint fogalom elsőként a szociolingvisztikában jelent meg, ahol látszólagosidő-vizsgálatként és valóságosidő-vizsgálatként nevezték meg őket (l. LABOV 1994, 2001; TRUDGILL 1978). Később a nyelvészet számos más területén is megjelent az eljárás. A fonetika területén elsősorban az anyanyelv-elsajátítást és az annak során bekövetkező változásokat, a különböző patológiás eseteket (például demencia, Alzheimer-kór, Parkinson-kór, vö. pl. ASH et al. 2017; LE et al. 2011; POWELL et al. 1989; MISONO et al. 2016; Bitchener 2010), a dohányzás

hatását, a nyelvi és beszédzavarok gyógyítására használt terápiákat, valamint fiatal életkorban az idő előrehaladtával bekövetkező változásokat elemezték longitudinális módszerrel.

Önmagában az idővel bekövetkező változások számos kutatási kérdést vetnek fel, amihez hozzáadódik a növekvő átlagéletkor miatti igény az idősödés folyamatával bekövetkező változások elemzésére. Noha az igény megvan egészséges, (fiatal vagy idős) felnőtt beszélők esetében is a követéses vizsgálatok megvalósítására, korábban csupán néhány ilyen típusú elemzés látott napvilágot, és ezek is nagyon heterogén eredményekkel szolgáltak. Ennek oka, hogy számos módszertani nehézség jelentkezhet a longitudinális elemzések végrehajtása során. Ezek a kutatások legtöbbször kis elemszámú adatközlői csoporttal valósíthatók meg, gyakran akár egyetlen beszélő utánkövetése történik. Ennek számos oka lehet. Például kevesen vállalják a hosszú követéses felmérésben való részvételt, illetve nagyobb csoportból is sokszor kevesebb adatközlő válik elérhetővé egy idő után, vagy az utólagos módszertan kialakításánál bizonyos idő után már nem minden felvétel minősége felel meg az összehasonlítás elvárt paramétereinek (pl. akusztikai sajátosságok célzott elemzése során a rossz minőségű, zajos vagy nem megfelelő körülmények között rögzített felvételek nem teszik lehetővé bizonyos elemzések elvégzését). De akár a beszélő életkörülményei, egészségi állapota is változhat, ami miatt a csoportba kerülés feltételei alapján ki kell zárni. Emellett az utánkövetett adatközlők gyakran professzionális beszélők, hiszen sokszor elsősorban tőlük áll rendelkezésre megfelelő minőségű, szabályos időközönként rögzített felvétel. Ez azonban nehezíti az összevetést, pl. amennyiben eltérnek a műfajok (sporttudósítás, hírolvasás, filmszerepek, politikai beszédek stb.), a rögzítések időpontjai, a beszédpartnerek életkora, neme. Emellett a professzionális beszélők gyakorta képzett beszélők, tehát nem a „hétköznapi” beszédmintát produkálják, legalábbis hivatásos körülmények között. Továbbá természetesen kis számú adatközlő esetén erősen dominálhatnak az egyéni különbségek, így a levonható következtetések köre is korlátozott.

Tehát az ilyen jellegű vizsgálatok elsősorban a professzionális beszélőkről és a náluk tapasztalható nyelvi változásokról nyújtanak információkat, és nem szolgálnak ismeretekkel arról, hogy milyen változások valószínűsíthetők az amatőr beszélőknél, azaz azoknál, akik nem vettek részt korábban a hangképzéssel és a beszéddel kapcsolatos képzéseken, valamint esetleg életvitelükből és munkájukból adódóan nem beszélnek olyan mennyiségben, mint a professzionális beszélők. Tehát egyrészt nem képzett beszélők, másrészt azonban a beszédterhelésük lehet hasonlóan magas, vagy épp sokkal alacsonyabb.

## 4. A beszéd variabilitása

A személyek beszédét a fiziológiás jellemzőkön túl számtalan más tényező is befolyásolja. Ilyenek lehetnek a szociális hatások, amelyek szintén változnak életünk folyamán, de a beszédet szorosabban tekintve is rengeteg meghatározó faktor lehetséges, pl. a beszédhelyzet vagy a létrehozott beszéd műfaja, de nyelvi vagy pragmatikai sajátosságok is.

Ismert például, hogy a nők és férfiak beszédjellemezői között vannak eltérések. A legszembe-tűnőbb talán az, hogy az **alapfrekvencia** ( $f_0$ ), azaz a hangszalagok rezgésének szaporasága eltér a nemek között (vö. pl. FITCH 1990; FITCH–HOLBROOK 1970; KROOK 1988; KÜNZEL 1989). Ennek oka részben az eltérő hangszalaghossz, részben a nemi hormonális hatások. Különbséget találtak az egyes nyelvek között is: az amerikaiakhoz képest a lengyelek, a japánok és a kínaiak is magasabb alapfrekvencián beszélnek (MAJEWSKI et al. 1972; YAMAZAWA–HOLLIEN 1992; KEATING–KUO 2010). ALTENBERG és FERRAND (2006) angol–oroszló két nyelvűeket vizsgálva azt találta, hogy oroszul magasabb alapfrekvencián beszélnek, mint angolul. A beszéd típus is meghatározó lehet az alapfrekvenciára, ami így az adott beszélőn belül is varianciához vezet. Számos kutatás vizsgálta az olvasás és a spontán beszéd alaphangjellemzőinek alakulását. Bár az eredmények nem voltak egységesek, a vizsgálatok többsége alapján a felolvasást magasabb alapfrekvencia jellemzi (vö. pl. SNIDECOR 1944; HANLEY–RINGEL 1966; HOLLIEN–JACKSON 1973; HUDSON–HOLBROOK 1982; SORENSEN 1982; ABU-AL-MAKAREM–PETROSINO 2007; SKARNITZL–VANKOVÁ 2017; GRÁ CZI et al. 2019). Ezzel ellentétben DALY és ZUE (1992) 45 női és 44 férfi adatközlő anyagán a spontán beszédben mért magasabb alapfrekvenciát; habár a férfiak esetében a különbségek csak 5 Hz alattiak voltak, a nők esetében 5–10 Hz közöttiek, szignifikánsnak bizonyultak. Vannak olyan kutatások is, melyek, ha találtak is különbséget a felolvasás és a spontán beszéd átlagos alapfrekvenciájában, az nem volt jelentős, illetve személyenként vagy csoportonként különbözött az eltérés iránya (CALGODNETTO et al. 1997; BÓNA 2011; BÓNA–MARKÓ 2012). A **hangterjedelem**, azaz a beszélő által használt legmagasabb és legalacsonyabb alapfrekvencia aránya (félhangban, arányszámként vagy zenei hangközmértékkel) esetében az eredmények szintén nem egységesek: több kutatásban azt találták, hogy a felolvasásban nagyobb a hangterjedelem (HUDSON–HOLBROOK 1982; BEKE 2008; SKARNITZL–VANKOVÁ 2017), ezzel szemben vannak olyan kutatások is, melyek a spontán beszédben mérték nagyobb hangterjedelmet (CALGODNETTO et al. 1997; BÓNA 2011). GRÁ CZI és munkatársai (2019) eredményei szerint a hangköz (egy adott beszédegység – mondat/megnyilatkozás/beszédszakasz – legalacsonyabb és legmagasabb alapfrekvencia-értékének aránya [ugyancsak félhangban, arányszámként vagy zenei hangközmértékkel megadható]) értéke a felolvasásban nagyobb, míg a hangterjedelem a spontán beszédben. Az egymásnak ellentmondó eredményeket okozhatják módszertani különbségek is: akár az adatközlők összetétele (nem, életkor, adatközlők száma), a vizsgálati anyag (szöveg, mondat vagy szólista felolvasása), a szöveg témája, hangulata stb. is befolyással lehet. Különbséget okozhat az is, hogy a beszélőknek írott szöveget kell-e felolvasniuk (ABU-AL-MAKAREM–PETROSINO 2007; HUDSON–HOLBROOK 1982; SORENSEN 1982, 1990), vagy a saját spontán beszédük leiratozott változatát (SWERTS et al. 1996; CALGODNETTO et al. 1997; DE SILVA 2003).

A legtöbb esetben a spontán beszédet (azaz az előre meg nem tervezett közléseket, melyek során az üzenet kitalálása, megformálása és meghangosítása kvázipárhuzamosan történik) és a felolvasást szokták elemezni a beszédstílusok vizsgálata esetén, de a beszédstílus egy skálán is felfogható, melynek két végpontja az említett két beszédstílus. WACHA (1974) négy kategóriát különböztet meg: a spontán beszédet, a félreproduktív beszédet, a reproduktív beszédet és a felolvasást. Ezek a kategóriák nem válnak el élesen egymástól. A félreproduktív jellegzetessége, hogy a beszélő előzetesen felkészül tartalmilag arra, hogy mit fog mondani, a nyelvi formát azonban

nagyrészt csak az adott beszédhelyzetben rendeli hozzá a mondanivalójához. Ebbe a kategóriába tartozhat a tanári megnyilatkozás, történetmondás, szórakoztató előadás (pl. stand-up) vagy az iskolai felelet. Reprodukzív beszéd alatt értjük azt, amikor a beszélő egy előre elkészített szöveget memorizál, és azt mondja el. Ilyen kategóriába tartozhat például a szónoklat, versmondás.

A beszédstílusok közötti eltérés magyarázatára számos elméleti keret létezik. Vannak olyanok, melyek a **beszédfigyelmet** állítják a középpontba, azaz azt, hogy a beszédprodukciónak mekkora figyelmet szán a beszélő a produktumra. Ilyen elmélet pl. JOOS (1968) vagy ZWICKY (1972) elmélete. Más elméleti keretek azonban továbblépnek, és a beszélő körülményei, attitűdje és további hatások mentén értelmezik a beszédstílust (pl. GILES–SMITH 1972; LADEGARD 1995). Ezen tényezők között szerepel például, hogy milyen személyek vesznek részt a megszólalás környezetében. Ki a címzett, ki a hallgatóság, és kire számít a beszélő, hogy meghallhatja a beszédet (BELL 1984). A címzethez közelíteni vagy távolítani törekszik a beszélő a szimpátia-antipátia következtében, mégis azok a személyek, akik még jelen vannak, vagy meghallhatják a beszédet (pl. a felvételt ki hallgathatja vissza) ugyancsak befolyásolják akár a magasabb nyelvi szinteket, pl. a szóválasztást, akár a szintaktikai szerkezetet, grammatikai struktúrát és prozódiai sajátosságokat (GILES–SMITH 1979). A közelítés vagy távolítás nem azt jelenti, hogy mindenképp az adott címzethez hasonlót vagy eltérőt akar a beszélő létrehozni, jelentheti azt is, hogy csak vélt beszédjellemzőkhöz alkalmazkodik, vagy ezektől távolodik, de azt is, hogy nem a beszéde alapján, hanem a társalgást befolyásoló egyéb tényezők, például a becsült hallásállapot / hallásviszonyok határozzák meg a beszédproduktumot. Ugyanakkor a részt vevő partnerek szociális viszonya is meghatározza a beszédprodukciónak, mint például az alá-főlérendeltség vagy azonos státusz (pl. BLUM–KULKA et al. 1989; SZILI 2004). Mindezek és további jellemzők további változatossághoz vezetnek egy beszélő adott időbeni beszédprodukciónak.

Arra, hogy milyen jellemzők változnak, milyen jellemzőket változtatnak a beszélők, a hétköznapi életünkből is szembeutó példa lehet, hogy idősekhez hangosabban beszélnek a feltételezett rosszabb hallás miatt, vagy hogy a gyermekekhez szóló beszédben megemelkedik és nagyobb tartományban realizálódik az alapfrekvencia (l. pl. NARAYAN–MCDERMOT 2016; FERGUSON 1964; magyarra MÁDY et al. 2018 – habár a tágabb  $f_0$ -tartományt a magyar tanulmány nem igazolta). Eltérő lehet a beszédhangok ejtése is. LABOV (1972) az angol beszédben az interdentális réshangok alveoláris felpattanó zárhangként való ejtésének arányát a beszédfigyelmelem, illetve az /r/ ejtését vagy törlését a szociális helyzet függvényében (LABOV 1986) elemezte.

Számos elmélet létezik annak leírására, hogy az ejtés során hogyan valósulnak meg a beszédhangok egy-egy adott helyzetben. Az egyik legelfogadottabb LINDBLOM **H&H-elmélete** (1990), amelyben a két H a *hypospeech* és a *hyperspeech* szó rövidítése, és az alul-, illetve a túlartikulációra utalnak. Az elmélet abból indul ki, hogy a beszélő és a hallgató közötti kommunikációban szükséges invariancia a hallgató feldolgozását jellemzi, nem pedig a beszélő produktumát. Vagyis az, hogy egy szót észleljünk a beszédben, a hallgató feldolgozási folyamataiban, lenyomataiban található mintázat invarianciáján múlik, az adott szó, az azt alkotó beszédhangok pedig variabilisan jelennek meg. Egyszerűbb talán onnan értelmezni ezt a gondolatot, hogy a hallgató, miközben elsajátítja a beszédfeldolgozási folyamatokat, folyamatosan eltérő realizációknak van kitéve, nagyon szemléletesen a koartikuláció (a beszédhangok képzési jellemzőinek egymásra hatása miatt) is egy-egy



beszédhang más-más hangkörnyezetben eltérően realizálódik. OHALA (1983) viszont arra hívja fel a figyelmet, hogy nem egyes beszédhangokat hallunk, tanulunk meg feldolgozni, hanem a hangsorokat. Ennélfogva az adott hangsorbeli realizációból vonjuk el a mintát, így az elsajátított észlelésben úgy van nyoma a beszédhangoknak, ahogyan a kapott mintákban szerepeltek. Visszatérve a H&H-elméletre, a túl- és alulartikulálás nyilván egy célhoz képest írható le. A beszédhang célkonfigurációjától, tehát ahogyan „ideális” esetben mozognak vagy állnak a beszédképző szervek annak ejtésekor, térhet el a beszélő a beszédprodukción során. Egyrészt törekszik arra, hogy a hallgató számára a (vélt) legérthetőbb módon hozza létre közlését, ugyanakkor a kisebb erőfeszítés irányába is törekszik, az ergonómia, energiatakarékosság jegyében. Az, hogy a skála két végpontja között hogyan hozza létre a beszélő a produktumot, folyamatos kontrollon alapul, amelyben a hallgató és a szituáció „igényeit” monitorozza. Ha visszatérünk a fentebb említett elméletekhez a beszédfigyelemről és a kommunikációs modellekről, amelyben a beszédpartnerek és a szituációk a központi meghatározó elemek, és ezeket együttesen gondoljuk végig, akkor észrevehető, hogy a beszédstílus variabilitásának ugyan más-más központi okot adnak meg, de mindegyik esetében a beszédprodukción a szituációjának tényezői mentén szerveződik: vagy pszichológiai természetű meghatározó tényezőt, vagy az ezek miatt változó beszédfigyelmet, vagy az artikulációs sajátosságok motoros kontrollját téve az elmélet központjává. Fontos eltérés a modellek között, hogy mennyiben enged egy beszédproduktum során stílusváltást, de ennek részletes tárgyalása nem célja a jelen tanulmánynak.

## 5. Beszédbeli változások az életkor mentén

Az életkorral bekövetkező változásokra visszatérve fontos kiemelni, hogy nemcsak az idősödés, hanem más változások hatása is megjelenhet, pl. a tanulóból dolgozó lesz, így változik a szociális státusza, változik a szociális körülménye, esetleg hosszabb külföldi tartózkodás is előfordulhat, ami akár akcentushoz is vezethet. Mindezek együttesen hatnak, ezért egy-egy vizsgálatban a lehető leghomogénebb csoportokat kell képezni. Így sem zárható ki az életkorhoz köthető szociális változások hatása. Azaz a fiziológiás változások mellett ezekről sem feledkezhetünk meg az adatok értelmezésekor. Egy-egy kilógó érték (a csoportra jellemző adatoktól statisztikai értelemben eltérő érték) vagy heterogén eredmények (pl. nagy szóródás, ellentétes tendenciák) esetében vissza kell térni a beszélők egyéb jellemzőihez, hogy található-e olyan, amely okozhatja esetében az eltérő adatot, illetve további csoportokra bontja a vizsgálat résztvevőit.

### 5.1. Az alapfrekvencia és a zöngé jellemzői

A **zöngé** kváziperiodikus összetett rezgés, amely egy alaphangból és annak felharmonikusaitól áll. Az alaphang frekvenciájának a felharmonikusok frekvenciái az egész számú többszörösei. A zöngé egyik legjellegzetesebb tulajdonsága ezért az alapfrekvenciája ( $f_0$ ), amely több

tényezőtől függ, egyebek között a hangszalagok hosszától, vastagságától, feszségétől, illetve (mint láttuk) a nemi hormonális hatásoktól.

### 5.1.1. Keresztmetszeti kutatások

A gyerekek alapfrekvenciája magasabb a felnőttekénél, és a nemek közti különbségek a pubertáskortól válnak jelentőssé, amikor a fiúk alapfrekvenciája lényegesen nagyobb mértékben mélyül, mint a lányoké (BENNETT 1983; HASEK et al. 1980; DUFFY 1970; HOLLIEN–MALCIK 1967; HOLLIEN et al. 1994). Az időskor szintén jelentősebb változást idéz elő az alapfrekvenciában, ez a nemek szerint eltérő irányú, a nők alapfrekvenciája mélyül, a férfiaké magasabb lesz, a két nem alapfrekvenciája tehát közelít egymáshoz (BROWN et al. 1991; DE PINTO–HOLLIEN 1982; MYSAK 1959; HOLLIEN–SHIPP 1972). Az idős és fiatal beszélők összevetésében rengeteg keresztmetszeti tanulmány született.

EICHHORN és munkatársai (2018) angol nyelvű vizsgálata is a nők alapfrekvenciájának kifejezettebb változását mutatta ki. 53 nő és 43 férfi adatközlő alapfrekvenciáját vizsgálták felolvasott, egy szótagos szavakban. Az adatközlőket három életkori csoportra osztották: 20–30, 40–60 és 70–92 évesekre. A férfiak csoportjában monoton növekedést találtak az életkor előrehaladtával az alapfrekvenciában, míg a nők esetében csökkenést találtak a második életkori csoportban az elsőhöz képest, majd növekedést a harmadik életkori csoportban a másodikhoz képest, ez a növekedés azonban kisebb mértékű volt, mint a csökkenés. HOLLIEN és SHIPP (1972) százhetvenöt – 20 és 90 év közötti – férfi felolvasásában vizsgálta az alapfrekvencia alakulását, és arra az eredményre jutott, hogy a férfiaknál 20 és 40 éves kor között csökkenés, 60 és 80 éves kor között pedig növekedés jellemzi az  $f_0$ -átlagértékeket. STATHOPOULOS és munkatársai 2011-es vizsgálatukban közel kétszáz – 4 és 93 év közötti – amerikai beszélő kitartott [a] magánhangzójának  $f_0$ -értékeit elemezték. Azt találták, hogy nemtől függetlenül az életkor szignifikáns hatással van az alapfrekvenciára, amely a férfi beszélőknél jelentősebbnek bizonyult. A legjelentősebb változás az alapfrekvenciában mindkét nem esetében kb. 20 éves korig történt, a csökkenés a férfiak csoportjában volt nagyobb mértékű. A férfiaknál további, enyhe csökkenés volt tapasztalható kb. 50 éves korig, majd ezután valamivel nagyobb mértékű egyenletes növekedés volt megfigyelhető. A női beszélőknél az átlagos alapfrekvencia értéke 60 éves korig enyhén csökkent, aztán az egyéni különbségek nagyobbakká váltak, a beszélők közti variancia nőtt, azaz nem mutatkozott egyértelmű trend a homogénnek tekintett, nemi és életkori csoportok esetében sem. A felnőttkorra kevesebb változás jellemző HOLLIEN és munkatársai szerint (1997), a férfiak alapfrekvenciája több változást mutat ebben az életszakaszban is, majd az időződés hoz mindkét nem esetén jelentősebb változásokat: a nők esetében az alapfrekvencia mélyül, ez azonban nem hirtelen, inkább fokozatos változás. Azt, hogy ez a mélyülés milyen életkortól indul, nagyrészt egyéni tényezők határozzák meg (kb. 55–60 év). A férfiak esetében az alapfrekvencia időskorban emelkedik, és a változás mértéke nagyobb, feltehetően később indul, mint a nőknél (kb. 65–70 év), de az egyéni különbségek itt is jelentősek. 75–80 éves kor felett egyre nagyobb az átfedő tartomány a két nem frekvenciaértékei között (HOLLIEN 1997).

MARKÓ (2015) magyar adatközlőkkel végzett kutatása szerint ugyanezen érték a felnőtt beszélőknél 60 éves korig a nők esetében 192 Hz, a férfiak esetében 114 Hz volt. BÓNA vizsgálatában

(2013) 90 és 100 éves nőknél átlagosan 147 Hz-et adatolt, míg a 70 és 85 év közötti férfiaknál nem igazolt számottevő különbséget az átlagértékben (119 Hz).

Azonban nem csupán az átlagértékek, de a teljes hangterjedelem, azaz a létrehozható legalacsonyabb és legmagasabb alaphfrekvencia-érték közötti távolság is változik időskorban (pl. BALÁZS 1993; NISHIO–NIMI 2008). GRÁ CZI és munkatársai (2020c) 77 nő és 73 férfi mondatfelolvasásában vizsgálták az alaphangjellemzők alakulását 20 és 79 év között. A nők esetében csökkenést találtak az életkor előrehaladtával az átlagos alaphfrekvenciában, a férfiak esetében nem találtak statisztikailag alátámasztható hatást. Az alaphfrekvencia szórása életkortól függetlenül a nők esetében magasabb volt, az életkor előrehaladtával hasonló mértékű növekedést találtak mindkét nem esetében. A hangterjedelem szintén a nők esetében volt magasabb, mindkét nem esetében nőtt az életkor előrehaladtával, azonban a nők esetében ez a növekedés nagyobb mértékű volt.

### 5.1.2. Longitudinális vizsgálatok

Az egyik első, felnőtteken végzett longitudinális vizsgálat egy II. Erzsébet angol királynő karácsonyi beszédeiből álló korpuszon készült (*Christmas Broadcast*), amelyben 50 éves időintervallumban (de nem minden évben) rögzítették a karácsonyi köszöntőt. Ebben minden alkalommal elhangzott az *I wish you a peaceful and very happy Christmas* mondat. A beszédek átlagosan körülbelül 5 perc hosszúságúak, ám vannak köztük 1 perc és majdnem 10 perc időtartamúak is. A felvételek alkalmasak az összevetésre, hiszen jó minőségű hanganyagokat tartalmaznak, minden esetben azonos a beszédhelyzet, ugyanis a királynő évértékelő és évzáró beszédének utolsó mondataként jelenik meg a közlés. Emellett a felvételeket azonos időközönként rögzítették (évente egy alkalommal), és a kontextuális heterogenitást csökkenti, hogy azonos mondatot vizsgálnak azonos szövegbeli pozícióban, így azonos az intonációs helyzet is. Emellett rendkívül hosszú időszakot, 50 évet ölelnek fel a felvételek: 1950 és 2000 között (HARRINGTON et al. 2007; REUBOLD et al. 2010). A királynő beszédét HARRINGTON és munkatársai más korpuszok beszélőivel is összehasonlították: a Machine Readable Spoken English Corpus-szal (ROACH et al. 2009), melyet a BBC rögzített az 1980-as években. Erzsébet királynő beszédében az  $f_0$  átlagértéke 267 Hz-ről 208 Hz-re csökkent 50 éves távlatban.

Egy másik kutatás 20 holland férfi felolvasásának hanganyagát vetette össze 30 éves távlatból az alaphfrekvencia, az  $f_0$ -szórásértékek és az egyéni különbségek tekintetében (DECOSTER–DEBRUYNE 2000). Eredményeik szerint az  $f_0$  átlagértéke és szóródása is fokozatosan csökkent a későbbi felvételek értékeiben a korábbi hanganyagokhoz viszonyítva, noha igen jelentős individuális különbségeket is találtak. FOUQUET (2016) 10 férfi alaphfrekvencia-értékeit vizsgálta, és eredményei szerint egyrészt a férfiaknál igen jelentős változások voltak kimutathatóak az életkor előrehaladtával, másrészt ez a tendencia nem volt egységes. A szerző 7 éves korban adatolta a legmagasabb  $f_0$ -értékeket, majd mélyülést tapasztalt 21 éves korig, és egészen 28 éves korig stagnáltak a vizsgált értékek. Ezt követően, 42 éves korig újabb csökkenést figyelt meg, később pedig 56 éves korig újabb növekedés volt kimutatható.

KASUYA és YOSHIDA (2010) elemzésében 20 férfi és 38 női beszélő jellemzőit vizsgálta, ám a felvételeket nem azonos időközönként rögzítették, a legrövidebb utánkövetés időtartama 11,

a leghosszabbé 19 év volt. Nem csupán az  $f_0$ , de a *jitter* és a *shimmer* értékeit is mérték a 20 férfi és 38 nő beszédét 11–19 éves távlatból tartalmazó korpuszukon. A *jitter* /dʒit:ɛr/ és a *shimmer* /ʃim:ɛr/ a hangszalagok rezgésének szabályosságát mérő számok. A *jitter* a szaporaság, azaz a frekvencia ingadozását számszerűsíti, míg a *shimmer* a kitérését, azaz amplitúdóját. Mindkettő számos módon mérhető az alapján, hogy hány periódus értékeit hány szomszédos periódus értékeire vetítjük. A nőknél az életkor előrehaladtával jelentősebb volt az alapprofrendencia csökkenése, mint a férfiak esetében, valamint a *shimmer* értéke nagyobb változáson ment keresztül az életkor előrehaladtával, mint a *jitter*é.

REUBOLD és HARRINGTON közös tanulmányukban (2015) az életkorhoz kötődő fonetikai változásokat Alistair Cooke brit–amerikai rádiókommentátor hanganyagai alapján vizsgálták, különös tekintettel az alapprofrendencia, valamint az  $F_1$  és  $F_2$  változásaira. Eredményeik szerint a nem alsó nyelvállású magánhangzóknál az első formáns ( $F_1$ ) és az alapprofrendencia ( $f_0$ ) egyaránt csökkenő–emelkedő mintát mutatott az életkor növekedésével.

### 5.1.3. Összegzés

Érdemes megfigyelni, hogy a fiatalabb életkorokban, 20–50 éves kor között több esetben is alapprofrendencia-csökkenés volt kimutatható mindkét nem esetében mind a longitudinális, mind a keresztmetszeti vizsgálatok alapján, de olyan eredmény is volt, amely szerint ebben az életkorban stagnáltak az értékek a férfiak esetében. Bár ebben az életkorban a hangszalagok nem nőnek, a hormonháztartás, a szociális státusz és számos további tényező változik folyamatosan, amely hatással lehet az értékekre. A nők esetében érdemes kiemelni, hogy ezen időtartomány vége felé a menopauza következtében fellépő hormonális és egyéb változások okozhatnak nagyobb mértékű változást.

## 5.2. A magánhangzók formánsértékei

Az alapprofrendencián túl a magánhangzók formánsértékeit is vizsgálták az életkor előrehaladtával. A **formánsok** a szonoráns beszédhangokban a **toldalékcső** (garatüreg, szájüreg, orrüreg, pitvar) által felerősített felhangnyalábok (a zöngé felharmonikusainak egy-egy egymáshoz közel eső tartománya). Mivel a toldalékcső működése határozza meg őket, ennek bármilyen változása módosulást okozhat a formánsértékekben. Gondoljunk például arra, hogy ha renyhül a szájpad izomzatának működése, az uvula nem záródik úgy, mint korábban, a beszéd orrhangzósabbá válik. Ennek hátterében az áll, hogy azon beszédhangok ejtésében, amelyek szájüregi képzésűek, megjelenik az orrüreg rezonanciás hatása is.

### 5.2.1. Keresztmetszeti eredmények

A kutatások többsége a formánsértékek tekintetében csökkenő tendenciát talált az életkor mentén, azonban nem minden vizsgált magánhangzó minden formánsértékére mutat ki összefüggést az életkorral. LINVILLE és RENS (2001) 19–24 évesek és 68–74 évesek

formánsértékeit elemezte, és eredményei szerint az első három formáns ( $F_1, F_2, F_3$ ) az idősebbek korcsoportjában alacsonyabb. DECOSTER és WIVINE (1999) az első két formánsra találta ezt a tendenciát, hasonló két életkori csoportot vizsgálva. EICHHORN és munkatársai (2018) az első négy formáns ( $F_1, F_2, F_3, F_4$ ) vizsgálták, de csak az  $F_1$  esetében találtak csökkenő tendenciát az életkorral, a többi formáns nem különbözött statisztikailag az eltérő életkori csoportokban. SCUKANEC és munkatársai (1991) négy magánhangzó (/i/, /æ/, /u/ és /a/) ejtését vizsgálták, és eredményeik szerint az  $F_1$  mind a négy magánhangzónál csökkent az életkorral, míg az  $F_2$  csak az /u/ és /a/ esetében mutatott csökkenést. XUE és HAO (2003) szintén több magánhangzót vizsgált, és azt találta, hogy az  $F_1$  mindegyik magánhangzó esetében csökkent az életkorral, az  $F_2$  és  $F_3$  esetén ez a tendencia már kevesebb magánhangzónál volt kimutatható. Vannak olyan kutatások is, melyek a magánhangzó-minőségtől függően eltérő tendenciákat találtak: RASTATTER és JAUQUES (1990) eredményei szerint a férfiak esetében az elől és középen képzett magánhangzók (/i/, /æ/)  $F_1$ -értékei növekedtek, míg a hátul képzett magánhangzók (/ʌ/, /u/ és /a/)  $F_1$ -értékei csökkentek az életkor előrehaladtával. A nők esetében az elől képzett magánhangzó (/i/)  $F_1$ -értékei hasonló tendenciát mutattak, míg a középen képzett magánhangzó (/æ/) nem mutatott változást az életkorral. Az  $F_2$ -re mindkét nem esetében az  $F_1$ -hez képest fordított trend volt megfigyelhető. TORRE és BARLOW (2009) eredményei szerint négy magánhangzó ([ɪ ɛ æ ʌ])  $F_1$ -értéke csökkent az életkorral mind a nők, mind a férfiak esetében, másik két magánhangzó ([i ʊ])  $F_1$ -értéke a nőknél emelkedett, a férfiaknál pedig enyhén csökkent az életkorral. Az  $F_2$ - és  $F_3$ -értékekre nem volt szignifikáns hatása az életkornak sem a nők, sem a férfiak csoportjában. A 60 év feletti beszélők (60 és 95 év közöttiek) formánsértékeiben nem találtak szignifikáns különbséget az életkor mentén (SEBASTIAN et al. 2012; FLETCHER et al. 2015).

### 5.2.2. Longitudinális eredmények

ENDRES és munkatársai (1971) 3 nő és 2 férfi adatközlő formánsértékeinek ( $F_1, F_2, F_3, F_4$ ) alakulását vizsgálta az /a/, /e/, /i/ és /o/ hangok ejtésében, és mindkét nem esetén csökkenést talált az első és második formáns értékében az életkor előrehaladtával, ami zártabb és hátrább történő képzésre utal. HARRINGTON és munkatársai (2007) 2 nő és 2 férfi formánsértékeinek alakulását vizsgálták longitudinálisan, 24–50 év távlatából. Sajnos a kutatásból levonható következtetések körét csökkenti, hogy a beszélőket más körülmények között rögzítették, más életkorúak voltak, más típusú professzionális beszélőknek tekinthetők, más időközönként rögzítették a hanganyagaikat (50, 24, 39 és 37 év), és más adatbázisból származtak a felvételek. Az első formáns esetén csökkenő tendenciát találtak az életkor előrehaladtával, a második formánsnál is csökkenést találtak, viszont sokkal kisebb mértékben. A harmadik formáns esetén a változások kevésbé voltak konzisztensek, de 3 beszélő esetén enyhe növekedést találtak az életkor előrehaladtával. MWANGI és munkatársai (2009) II. Erzsébet királynő fentebb említett korpuszában vizsgálta a formánsokat, a királynő 26 éves és 76 éves kora között. Az első formáns esetén negatív korrelációt talált a formánsértékek és az életkor között a magánhangzók többségében, tehát az  $F_1$  értéke csökkent az életkor

előrehaladtával. A második és harmadik formáns esetén a magánhangzók többségénél nem talált erős korrelációt a formánsértékek és az életkor között, azaz ezek kevésbé változtak. REUBOLD és munkatársai (2010) a II. Erzsébet királynő és az Alistair Cooke által ejtett /ə/ magánhangzók első formánsainak alakulását követték végig. A királynő esetében 29, Cooke esetében pedig 47 felvétel anyagát elemezték, mintegy 30 év távlatában. II. Erzsébet királynő esetén folyamatos csökkenést találtak, míg Cooke  $F_1$ -értékei is nagyjából hasonló mértékben csökkentek az adatközlő 80 éves koráig, utána azonban nagyobb meredekségű növekedést mutattak.

### 5.2.3. Összegzés

A legtöbb vizsgálat azt mutatta, hogy az első formáns értéke csökken az életkor előrehaladtával. A második formáns esetében nem volt ennyire egységes az irodalom, hol csökkenést, hol stagnálást, hol pedig a mért beszédhanggal való összefüggést találtak. A további formánsok változása nem volt egységes. A formánsok eltérését indokolhatják a toldalékcsoben végbemenő artikulációs beállítások (pl. renyhébb légyszájpad-mozgás, pontatlanabb motoros kivitelezés).

II. Erzsébet királynő formánsainak elemzése során egy további fontos szempont is felmerült. Mégpedig az, hogy melyik társadalmi réteghez célozza beszédét. A várt tendenciának ellentmondó eredményeket azzal magyarázták HARRINGTON és munkatársai (2000), hogy beszédében fiatalabb alsóbb osztálybeliekre jellemző dél-brit köznyelvi magánhangzók irányába tolódtak a formánsértékek. Azaz a szerzők véleménye szerint nem a beszélő életkori sajátosságai határozták meg az eredményt, hanem a célközönségé.

## 5.3. A mássalhangzók akusztikai szerkezete

A mássalhangzók akusztikai szerkezete a képzési mód függvényében eltérő. A felpattanó zárhangok esetében teljes zárat képzünk, majd a zár feloldása felpattanással történik, ezért az akusztikai lenyomatban zárszakaszra és felpattanásra oszthatjuk a tipikus realizációt. A zöngétlen felpattanók esetén a zár feloldása után bizonyos idővel indul újra a zöngékepzés a követő magánhangzóhoz, míg a zöngések esetében általában a teljes zárszakasz zöngés, így a felpattanáshoz képest korábban indul el a zöngé. Ez a jellemző a **zöngékezdési idő** (*voice onset time, VOT*), pontos definícióval: a zár feloldásának kezdetétől a zöngé megindulásáig eltelt idő (részletes bemutatásért l. pl. GRÁ CZI 2016), ami tehát a zöngés mássalhangzók esetében negatív, a zöngétlenek esetében pozitív zöngékezdési idővel jellemezhető. A réshangok esetében folyamatosan áramlik a levegő a szájüregtől a külső tér felé, miközben egy résen, két beszédképző szerv közötti nagyon kis keresztmetszetű területen átáramolva a levegő sűrűlódik. Amennyiben zöngés, a lenyomaton a zöngesség általában végig jellemző, ha zöngétlen, akkor az akusztikai lenyomat a rés okozta sűrűlódási zörejjel jellemezhető csak. Minden mássalhangzó esetében meghatározza az akusztikai lenyomatot a képzési hely, azaz az akadály helye. Ennek egyik mérőszáma a **CoG**, amely a mássalhangzók esetében a **spektrális súlypontot** jelenti. Ez a beszédhang

frekvencia-összetevőinek az intenzitásukkal súlyozott középpontját jelenti. (A mássalhangzók képzésének teljes ismertetéséhez lásd MARKÓ [2017] munkáját, a zöngesség kérdésköréhez pedig BÁRKÁNYI és KISS munkáit [pl. 2018], illetve GRÁCZI és munkatársai munkáit [pl. 2012, 2020a, 2020b]).

Mivel a felpattanó zárhangok zöngेкеzdési ideje nagymértékben nyelvspecifikus, erről csak magyar eredményeket közlünk a jelen összevetésben.

### 5.3.1. Keresztmetszeti vizsgálatok

BÓNA (2013), illetve BÓNA és BEKE (2013) a réshangok közül az /s/ és /ʃ/ ejtését 20–32 és 70–85 éves nők beszédében, illetve a zöngétlen felpattanó zárhangok közül a /p/, /t/, /k/ realizációit 21–32 és 70–90 éves nők ejtésében elemezte. Eredményeik szerint a réshangok esetében a spektrális súlypont az idősök beszédében kevésbé szór, de magasabb átlagértéket vesz fel, mint a fiatalok ejtésében. Ezt azzal magyarázták, hogy az idősök gyakorlottabb motoros kontrollal, de előrébb képzik ezeket a beszédhangokat. A felpattanó zárhangok zöngेкеzdési ideje mindhárom beszédhang esetében az idősök ejtésében mutatott nagyobb szórást, a /p/ és /t/ esetében magasabb, a /k/ esetében alacsonyabb átlaggal, mint a fiatalok beszédében. A nagyobb szórás hátterében a pontatlanabb motoros kivitelezés, a /p/ és /t/ hosszabb zöngेкеzdési ideje esetében az artikulációs mozgások lassulása, a /k/ magasabb VOT-jének esetében a pontatlanabb zárképzés állhat az idősök beszédében.

SZÁRAZ és GRÁCZI (2020) a /t d k g/ mássalhangzókat vetette össze fiatal és idősödő beszélők ejtésében (23–35, 65–77 év, 15-15 nő és férfi mindkét életkori csoportban). A zöngेкеzdési idő a két zöngés mássalhangzóban és a /t/-ben hosszabb volt az idősödők ejtésében. Az előbbi a teljesen zöngésen megvalósult mássalhangzók értékeit tartalmazta csak, így a mérőszám gyakorlatilag megegyezett a zárszakasz időtartamával, ami a lassabb artikulációs tempó következményeként hosszabb az idősebbek beszédében. A /t/ esetében az eredmény megegyezik BÓNA (2013) eredményével, feltehetően a lassabb artikuláció miatt elért magasabb szájüregi nyomás és lassabb zárfelnyílás áll a hátterében. A /k/-ban mért érték nem tért el az életkori csoportok között, míg BÓNA (2013) kutatásában igen, az idősebbek beszédében rövidebb volt. Feltehetően ezen a képzési helyen a zárképzés nagyobb pontatlansággal valósult meg náluk, ami miatt az akadály mögött felgyülemelő levegő „szökött”, így nem emelkedett meg a nyomás lassabb artikuláció esetében sem. Az alternatív realizációk, amely során a felpattanó zárhang képzése helyett réshang vagy approximáns jelent meg, ritkább volt az idősök ejtésében. Ez az elemzés azonban kategoriális volt, azaz nem a két képzési mód között skálázta, hanem az egyik vagy a másik csoportba sorolta a megvalósulást. A fiatalok esetében esetleg gyakrabban megjelenő lezser beszédmód, alulartikulált beszédstílus lehetett jellemző, ami eltérő képzési móddal való megjelenéshez vezetett, míg az idősödők feltehetően törekedtek a feszebb beszédre, de a beszédképző szervek mozgása pontatlanabb volt, és egy, képzési mód váltásával nem járó, de levegőszivárgást okozó zárképzést hozhattak létre. Ennek bizonyítására azonban artikulációs, aerodinamikai vizsgálatok lennének szükségesek.

Mássalhangzókon végzett longitudinális kutatásokról nem tudnak a jelen tanulmány szerzői.

### 5.3.2. Összefoglalás

A mássalhangzók életkori eltéréseiről csak fiatal és idős(ödő) beszélők összevetése alapján tudunk beszámolni. Az, hogy a fiatal-középkorú felnőttek beszédében milyen tendenciák jelenhetnek meg, ha megjelennek, további vizsgálatok tárgyát képezi. A vizsgált életkori csoportok közötti eltérések hátterében többnyire a motoros kontrollt, a nyomásviszonyokat és az esetleges fogazati változásokat említették a szerzők.

## 5.4. Az időzítés

A beszéd időzítése vagy temporális szerkezete számos jellemzőt foglal magában. Ilyen az **artikulációs tempó** (amelyet az artikuláció tiszta idejére eső beszédjelek – pl. beszédhang, fonéma, szótag – számával számszerűsítünk) és a **beszédtempó** (ez a beszéd teljes idejére eső beszédjelek száma – azaz a szüneteket is tartalmazza az időtartam), illetve ezekkel összefüggésben a szünetezés. Ezek mellett idetartozik a beszédhangok időtartama, vagy akár belső időszerkezete (amennyiben egy beszédhang képzésének artikulációs konfigurációja több mozzanatot tartalmaz, mint pl. a fent elemzett felpattanó zárhangoké). Mindezekből és további tényezőkből eredően a ritmus is a temporális szerkezet része.

### 5.4.1. Keresztmetszeti vizsgálatok

Az artikulációs tempó és az életkor kapcsolata is ellentmondó volt az egyes kutatások esetében. Számos tanulmány igazolta, hogy az idősek lassabban beszélnek, mint a fiatalabbak (OYER–DEAL 1985; SMITH et al. 1987; BROWN et al. 1989; ZELLNER–KELLER 2006; YUAN et al. 2006). Más kutatások azonban mindennek ellentmondó eredményekre jutottak: néhány vizsgálat nem talált összefüggést az életkor növekedése és az artikulációs tempó csökkenése között. Ilyen volt például HOIT és munkatársai (1994) vizsgálata 80 fő részvételével, akik 20–30, 40–50, 60–70 évesek és 80 év felettiak voltak; illetve LINVILLE (2000) kutatása ugyancsak 80 fővel, akik közül negyvenen 19 és 24 év közötti beszélők és negyvenen 62 és 79 év közötti beszélők voltak.

SMITH és munkatársai (1987) 10 fiatal és 10 idős beszélő normál és gyorsított tempóban történő szó- és mondatvisszmondása alapján azt találta, hogy az idősek 20–25%-kal lassabb produkcióra voltak képesek, mint a fiatalok. A lassulást számos tényező okozhatja a szerzők szerint: az idősebb beszélők hosszabb szótagidőtartamokat és hosszabb szüneteket hoztak létre, mint a fiatalok. A különbségeket az időskort jellemző fiziológiás változásokkal magyarázták: a tüdőszövet rugalmassága csökken (LINVILLE 1996), de meghatározó a változó gégefunkció, a szájizmok és az ajkak csökkenő aktivitása is (LINVILLE–RENS 2001).



#### 5.4.2. Longitudinális eredmények

GESTERNBERG és munkatársai (2018) idős (70–80 éves) felnőtt német és francia beszélők beszédében (LangAge korpusz) az artikulációs tempó különböző prozodikus paramétereinek, a beszédszakaszok időtartamának és a szünetek közötti összetett kölcsönhatásnak az összefüggéseit vizsgálták mintegy 10 éves különbséggel. Nagyon jelentős egyéni különbségek rajzolódtak ki az életkor előrehaladtával minden vizsgált tényezőben: mind abban, hogy növekedés vagy csökkenés volt mérhető, mind pedig ezek mértékében. Összességében azonban a nyelvek (feltehetően inkább a kultúrák) közötti eltéréseket is tapasztaltak: míg a franciák esetében gyorsulás, a németek ejtésében lassulás volt jellemzőbb. Az eredmények mögött általában a kardiovaszkuláris rendszer változása, az ezzel együtt járó rövidebb légzési idő és a kisebb tüdőkapacitás állhat. A nagy egyéni különbségek az életviteli eltérésre is visszavezethetők, amit a szerzők nem kontrolláltak, azaz nem képeztek ezek alapján homogén csoportokat (pl. dohányzás, hangterhelés).

Míg GESTERNBERG és munkatársai (2018) eredményei szerint a beszédszakaszok időtartama negatívan korrelált a beszélők korával, miközben a szünetek és az életkor között nem volt szignifikáns összefüggés, más kutatások szerint az idősebbek artikulációs tempójára nagyobb fokú variabilitás volt jellemző, mint a fiatalokéra, így nem vonható le egységes következtetés a beszélők produkciójára vonatkoztatva csupán azok életkora alapján (LINVILLE–RENS 2001).

Egy másik vizsgálatban SHUM (2008) kitért II. Erzsébet királyné artikulációs tempója változásának vizsgálatára a *Christmas Broadcast* alapján. Noha a kutatás nem szolgáltat részletes leírással a vizsgálat módszertanáról és eredményeiről, a bemutatott diagram alapján leolvasható, hogy a királynő artikulációs tempója fokozatosan gyorsult 20 és kb. 55 éves kora között, majd ezt követően tendenciális csökkenés volt megfigyelhető a 80-as éveinek végéig.

QUENÉ (2013) egy másik kutatásban Beatrix királynő beszédének temporális sajátosságait elemezte longitudinális módszertannal az 1980 és 2012 között rögzített *Troonrede*-i („trónbeszéd”: minden év szeptember 3. keddjén az uralkodó által megtartott beszéd, amely felvázolja a következő év kormánypolitikáját) alapján (a királynő életkora ebben az időintervallumban 42 és 74 év közé esett, összesen 9 felvételt elemeztek). Az eredmények ellentmondtak a korábbi egyértelmű tendenciákat bemutató tanulmányoknak: az artikulációs tempó kismértékben csökkent az első néhány évtizedben, az utóbbi időszakban ezzel ellentétesen gyorsulás volt megfigyelhető. Tempóváltozás a beszédekben belül is mérhető volt: míg a korábbi felvételeket (1980 és 1996 között) a lassabb–gyorsabb–lassabb mintázat, a későbbi felvételeket (1996 és 2012 között) fokozatos gyorsulás jellemezte.

HUSZÁR és KREPSZ (2021) 13 fiatal és középkorú férfi felolvasott és spontán beszédét vetette össze 10 év távlatában (az első felvétel időpontjában a beszélők 19–40 év közöttiek voltak, 10 évvel később 29–50 évesek). Az első és a második felvétel között mutatkozott gyorsulás, lassulás és stagnálás is. Fiatal felnőttkorban kevésbé jelentős változások várhatók, még ilyen hosszabb időtartamot, akár egy évtizedet követően is, egészséges beszélőknél. A spontán beszéd esetében a felvételek közti különbségek teljesen elhanyagolhatóak voltak, az olvasott szöveg esetén azonban a második felvételek esetében nagyobb volt a szünetek aránya az elsőhöz viszonyítva. A szünetek

gyakoriságát tekintve szignifikáns különbséget találtak a két felvétel között: az olvasott szöveg esetében gyakoribbak voltak a szünetek az első felvételen, mint a másodikon. A 10 év alatt a vizsgált temporális jellemzőkben nem vagy csak kismértékben jelent meg változás, és GERSTERNBERG és munkatársainak (2011) eredményeihez hasonlóan jelentős egyéni eltéréseket találtak.

### 5.4.3. Összefoglalás

Az időzítés, a temporális sajátosságok jelentős összefüggést mutattak tehát az életkorral általában, idősebb korban lassabb tempót, nagyobb szórásértékekkel, több, hosszabb szünetet vonva maga után. A rövid távú longitudinális vizsgálatok eredményeiben mind fiatalabb, mind idősebb korban az egyéni különbségek meghatározóbbnak bizonyultak, mint az életkor hatása.

## 6. Magyar fiatal felnőttek longitudinális beszédkorpusza

Láthatjuk, hogy a felnőttek beszédjellemezőinek változását általában nagy időtávlatban vizsgálják, jellemzően képzett, a beszédet mint mindennapi munkaeszközt használó beszélők (pl. hírbemondók, politikusok) ejtésében. Az alapfrekvencia és a formánsok elemzése gyakori, míg a többi fonetikai jellemzőt nem vizsgálták utánkövetéses módszerrel. Ugyancsak érdemes figyelembe venni, hogy az utánkövetéses kutatásokban többnyire alacsony beszélőszámmal tudnak dolgozni a kutatók.

Mindezen tényezőket és problémákat figyelembe véve a Nyelvtudományi Kutatóközpontban egy olyan longitudinális korpusz építése kezdődött meg (GRÁCZI et al. 2020d), amely több szempontból is újdonságot jelent. Az első vizsgálat időpontjában a Longitudinális korpuszban szereplő adatközlők közül a legfiatalabb 19 éves volt, a kiválasztott csoport legidősebb tagjai a 40-es éveikben jártak, azaz a célcsoport fiatal és középkorú felnőttekből áll. Őket hívjuk vissza mintegy 10 (10-11) év elteltével utánkövetéses vizsgálatra. A korpusz egyrészt amatőr, azaz nem professzionális beszélők adatait tartalmazza, másrészt – noha a protokoll kialakítása retrospektív, azaz utólagosan kialakított szempont szerint történt – az eredeti felvételek is hang- és beszédelemzés céljából készültek, így jó minőségűek, és alkalmasak az összevetésre. Ahhoz, hogy az idő előrehaladtával bekövetkező változásokról képet kapjunk, a beszédfelvételek után egy 22 kérdésből álló kérdőív segítségével igyekszünk felmérni az egyes, a hanggal kapcsolatos befolyásoló tényezőket. Ez alapján az esetleges csoportmintázattól eltérő egyéni tendencia megjelenése esetén ellenőrizhető, hogy nincs-e valamilyen oka annak. A kérdések kitérnek a beszélő fizikai sajátosságaira (súly, magasság), nyelvismeretekre, káros szenvedélyekre, a napi szintű beszédterhelésre (pl. beszéd- és énekmenyiségre, hangképzési sajátosságokra), a hangot érintő esetleges betegségekre (légzőszervi problémák, fogászati problémák, géjét érintő betegségek stb.), további, a beszédet befolyásoló tényezőkre (pl. hogy a beszélő napi szinten érintkezik-e vegyszerekkel;

hogy mennyi folyadékot fogyaszt; illetve arra, hogy eltérnek-e a rögzítéskor mért hangjellelzők a beszélő által átlagosnak ítélttől).

Az első felvételek minden beszélőtől egy másik, az ún. BEA adatbázisból származnak. A BEA (= BEszélt nyelvi Adatbázis vagy néha Spontánbeszéd-adatbázis) 2008-ban indult ugyancsak a Nyelvtudományi Intézet Fonetikai Osztályán (pl. NEUBERGER et al. 2014). Mára 461 beszélő felvételeit tartalmazza többféle beszédstílusban. Azokat az adatközlőket, akik utolérhetőek voltak, kerestük fel a 10 évvel későbbi, utánkötéses felvételre. Az új felvételt az eredeti, BEA adatbázis protokollja alapján állítjuk össze, személyre szabottan. Egyrészt a feladatok többsége azonos, másrészt ugyanazokat a témákat vetjük fel az adatközlőnek, mint a korábbi felvételeken. Egy további és egy változtatott feladattal módosultak az új felvételek. A teljes felépítést a következőkben ismertetjük.

Az eredeti felvételen 25, az interjúkészítő által felolvasott mondatot ismételnék meg, illetve a felvétel végén papírlapról olvasnak fel az adatközlők. Az utánkötéses felvételeken ezek száma 15-re csökkent, mivel maga a felvételi idő jelentősen nyúlt a plusz feladat (azonos mondatlista többszöri felolvasása) miatt. Mindkét felvételen ugyanazon szöveget olvassák fel. Az azonos mondatok és a szöveg így az azonos beszédmódok között összevethető, és az olvasás, mondatismétlés, tehát a nem spontán beszédprodukcióban mérhető változások elemzését teszi lehetővé.

Az eredeti és a követéses felvételek is tartalmaznak az adatközlő életéről egy kvázimonologikus interjút (ahol az interjúkészítő elsődlegesen olyan kérdéseket tesz fel a beszélő számára, amelyre ez utóbbi hosszán, monológyszerűen válaszol), egy témában véleménykifejtést, illetve két meghallgatott szöveg tartalmának visszamondását. Ezekben az esetekben tehát spontán vagy félspontán beszédprodukciókat vethetünk össze. Míg a felolvasott és utánmondott feladatok esetében teljesen azonos az egyes beszédhangok/szavak kontextusa, ebben az esetben a variabilitás nagyobb.

További feladat volt az eredeti felvételek során egy háromfős társalgás. Ebben egy témát vet fel az interjúkészítő, de attól a társalgás előrehaladtával eltérhetnek a beszélgetőpartnerek. Ezt az új felvételek során egy harmadik személy részvétele nélküli témakifejtés váltotta fel. Így a beszédmód maga megváltozott, aminek a hatását elemezhetjük.

Az új, követéses felvételeken egy további beszédfeladattal egészítettük ki a felvételi protokollt, mégpedig egy olyan mondatlista felolvasásával, amelyben egyes hangkapcsolatok kontrollált környezetben szerepelnek. Ennek oka, hogy a korábbi mondatokban nem voltak kontrolláltak a fonetikai pozíciók, míg ezekben igen, így a célzott beszédhangok variabilitása pontosabban leírhatóvá válik. Ezeket az adatközlők négyszer olvassák fel, random sorrendben, két-két másik beszédfeladat között.

A készülő korpusz két szempontból jelent újdonságot és kutatási lehetőséget az érdeklődők számára. Egyrészt az újonnan rögzített felvételek által lehetőség nyílik az adott beszélőn belüli rövid (felvételen belüli) és hosszú távú (10 év) variabilitás nyomon követésére, valamint a beszélők közötti változások rendszerszerű fonetikai összevetésére. Emellett a Longitudinális korpusz nyílt hozzáféréssé, és elsődlegesen longitudinális és más összevető szempontú kutatásokhoz

biztosít hanganyagokat, valamint folyamatban van a beszédszakaszszintű annotálás (a közlések lejegyzése szünettől szünetig tartó szakaszokra tagolva) és az automatikusan létrehozott beszédhangszintű címkézés (azaz a közlések felbontása beszédhangokra egy algoritmus segítségével), amelyet ugyancsak elérhetővé teszünk a beszéd változásai iránt érdeklődő kutatóknak.

Az így épülő adatbázison végzett elemzések, az alaphangfrekvencia jellemzőinek, a magánhangzók formánsértékeinek, a mássalhangzók spektrális szerkezetének és a temporális jellemzők változásának vizsgálatai számos területen hasznosíthatók. A tipikus fejlődésű, egészséges, amatőr (nem professzionális) beszélők beszédjellemzőinek változása hozzájárulhat az életkorral együtt járó és a beszédpatológiás esetek szétválasztásához, ehhez azonban nagyszámú adatközlő bevonásával történt standard életkori referenciaértékek meghatározása szükséges (FOUGERON 2020). A korpusz és a rajta végzett vizsgálatok eredményei a beszélőazonosítás egyes kérdéseihez is hozzájárulhatnak. Többször évek telnek el a bűnesethez tartozó kérdéses felvétel és egy gyanúsított beszédének felvétele között, amikor is az összehasonlítás során fontos tudni, milyen eltérés alapján feltételezhető még azonosnak vagy eltérőnek két beszélő. Mindemellett az egyes életkori jellemzők megismerése segítséget nyújthat a vizsgálatok megtervezése során a megfelelő adatközlők megválasztásához. A kutatás segítséget nyújthat olyan kérdések megválaszolásához is, mint hogy hogyan változnak a nyelvi jellemzők a gyakran homogénnek vélt felnőttkorban. Azonos módon változik-e a különböző beszélők hangja? A férfiak és a nők hangja azonos módon változik-e? Mennyiben tekinthető egyéninek az életkorral bekövetkező változás? Hogyan adaptálódnak a beszélők az életkorral bekövetkező változásokhoz? Hogyan választható el az öregedés folyamata más szocializációs változásoktól, mint a hivatásnak a hangra gyakorolt hatása vagy a fizikai aktivitás?

## 7. Kitekintés

A korábbi szakirodalom áttekintéséből láthatjuk, hogy gyakran nagyobb távlatban vizsgálják a beszéd változását az életkor mentén. A hosszabb és a rövidebb távú longitudinális vagy épp keresztmetszeti vizsgálatok között is előfordultak eltérések (gyakorlatilag minden vizsgálati tényezőben). Ezeknek az oka lehet az esetleges egyedi és kohorszhatások, vagy épp az eltérő metodológia is. Egy eredetileg is beszédelemzésre készült keresztmetszeti korpusz kibővítése hosszmetzeti felvételekkel olyan lehetőséget ad, amelyben a nagyrészt azonos, másrészt egymást kiegészítő protokoll következtében az eredmények összevethetővé válnak a két vizsgálati módszer között, és az életkorok mentén bekövetkező eltérések jobban függetleníthetővé válnak az ilyen jellegű hatásoktól. Az egészséges beszélőkön végzett longitudinális és keresztmetszeti vizsgálatok segíthetnek akár a patológiás és egészséges változások/eltérések közötti határok meghatározásában (pl. a jelen kötetben KISS 2021 és TULICS 2021), de akár a beszélőazonosítás néhány kérdésében is (l. pl. TATÁR – VARGA – FŐZŐ 2021 a jelen kötetben).

## FOGALMAK

*kronológiai életkor; biológiai életkor; keresztmetszeti vizsgálatok; hosszmetzeti (longitudinális) vizsgálatok; kohorszeffektus; H&H-elmélet; Longitudinális korpusz; elsődleges, másodlagos, harmadlagos öregedés; alaphfrekvencia; hangterjedelem; hangköz; spontán beszéd; beszédfigyelem; zöngé; formáns; toldalékcső; zöngékezdési idő; spektrális súlypont, CoG; artikulációs tempó; beszédtempó*

## IRODALOM

- ABU-AL-MAKAREM, Ali – PETROSINO, Linda 2007. Reading and spontaneous speaking fundamental frequency of young Arabic men for Arabic and English languages: A comparative study. *Perceptual and Motor Skills* 105/2. 572–580.
- ALTENBERG, Evelyn P. – FERRAND, Carole T. 2006. Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. *Journal of Voice* 20/1. 89–96.
- AMARYA, Shilpa – SINGH, Kalyani – SABHARWAL, Manisha 2018. Ageing process and physiological changes. In D’ONOFRIO, Grazia – SANCARLO, Daniele – GRECO, Antonio (eds): *Gerontology*. IntechOpen, London. 1–24.
- ASH, Sharon – JESTER, Charles – YORK, Collin – KOFMAN, Olga L. – LANGEY, Rachel – HALPIN, Amy – FIRN, Kim – PEREZ, Sophia Dominguez – CHAHINE, Lama – SPINDLER, Meredith – DAHODWALA, Nabila – IRWIN, David J. – McMILLAN, Corey – WEINTRAUB, Daniel – GROSSMAN, Murray 2017. Longitudinal decline in speech production in Parkinson’s disease spectrum disorders. *Brain and Language* 171. 42–51.
- BALÁZS Boglárka – BÓNA Judit 2016. Életkori sajátosságok a beszédképzésben és a beszédfeldolgozásban. In BÓNA Judit (szerk.): *Fonetikai olvasókönyv*. ELTE Fonetikai Tanszék, Budapest. 7–18. <http://fonetikaitanszek.elte.hu/wp-content/uploads/2017/11/OlvasokonyvTeljes.pdf> (A letöltés ideje: 2020. december 12.)
- BÁRKÁNYI Zsuzsanna – G. KISS Zoltán 2019. A fonetikai korrelátumok szerepe a zöngékонтрасzt fenntartásában. Beszédproduktions és észleléses eredmények. *Általános Nyelvészeti Tanulmányok* 31. Akadémiai Kiadó, Budapest. 57–102.
- BEKE András 2008. A felolvasás és a spontán beszéd alaphangszerkezeteinek vizsgálata. *Beszédkutatás* 2008. 93–107.
- BELL, Alan 1984. Language style as audience design. *Language in Society* 13/2. 145–204.
- BELSKY, Daniel W. – CASPI, Avshalom – HOUTS, Renate – COHEN, Harvey J. – CORCORAN, David L. – DANESE, Andrea – HARRINGTON, Hona Lee – ISRAEL, Salomon – LEVINE, Morgan E. – SCHAEFER, Jonathan D. – SUGDEN, Karen – WILLIAMS, Ben – YASHIN, Anatoli I. – POULTON, Richie – MOFFITT, Terrie E. 2015. Quantification of biological aging in young adults. *Proceedings of the National Academy of Sciences of the United States of America*. 112/30. E4104–4110.
- BENNETT, Suzanne 1983. A 3-year longitudinal study of school-aged children’s fundamental frequencies. *Journal of Speech, Language, and Hearing Research* 26/1. 137–142.

- BITCHENER, John 2010. The relationship between the negotiation of meaning and language learning: A longitudinal study. *Language Awareness* 13/2. 81–95.
- BLUM-KULKA, Shoshana – HOUSE, Juliana – KASPER, Gabriele 1989. *Cross-Cultural Pragmatics: Request and Apologies*. Ablex Publishing Corporation, Norwood.
- BODZSÁR Éva – ZSÁKAI Annamária 2013. *Antropológiai/Humánbiológiai gyakorlatok*. ELTE, Digitális Tankönyvtár. [https://regi.tankonyvtar.hu/hu/tartalom/tamop412A/2011-0073\\_antropologiai\\_humanbiologiai\\_gyakorlatok/ch11.html](https://regi.tankonyvtar.hu/hu/tartalom/tamop412A/2011-0073_antropologiai_humanbiologiai_gyakorlatok/ch11.html) (A letöltés ideje: 2021. január 10.)
- BÓNA Judit 2011. A különböző beszédstílusok az akusztikai-fonetikai és percepció vizsgálatok tükrében. *Alkalmazott Nyelvtudomány* 11/1–2. 39–48.
- BÓNA Judit 2013. *A spontán beszéd sajátosságai az időskorban*. ELTE Eötvös Kiadó, Budapest. [http://www.eltereadler.hu/media/2014/10/Bona\\_Judit\\_Idoskori\\_READER.pdf](http://www.eltereadler.hu/media/2014/10/Bona_Judit_Idoskori_READER.pdf) (A letöltés ideje: 2020. április 10.)
- BÓNA Judit 2016. Különböző életkorú idősök beszédének temporális sajátosságai In BALÁZS Géza – VESZELSZKI Ágnes (szerk.): *Generációk nyelve*. Tanulmánykötet. Magyar Szemiotikai Társaság – ELTE BTK Mai Magyar Nyelvi Tanszék, Budapest. 221–231.
- BÓNA Judit – BEKE András 2013. A zöngétlen réshangok akusztikai szerkezete fiatal és idős korban. In GECSŐ Tamás – SÁRDI Csilla (szerk.): *Az interkulturális kommunikáció elmélete és gyakorlata*. Tinta Könyvkiadó, Budapest. 38–43.
- BÓNA, Judit – MARKÓ, Alexandra 2012. Fundamental frequency patterns: the factors of age and discourse type. In *Proceedings of 'Sociophonetics, at the crossroads of speech variation, processing and communication'*. Pisa. 45–48.
- BROWN, Steven R. 1996. Q methodology and qualitative research. *Qualitative Health Research* 6/4. 561–567.
- BROWN, William S. – MORRIS, Richard J. – HOLLIEN, Harry – HOWELL, Elizabeth 1991. Speaking fundamental frequency characteristics as a function of age and professional singing. *Journal of Voice* 5/4. 310–315.
- CALDOGNETTO, Magno E. – ZMARICH, Claudio – FERRERO, Franco 1997. A comparative acoustic study of spontaneous and read Italian speech. In *Proceedings of EUROSPEECH 1997*. 770–782.
- DALY, Nancy A. – ZUE, Victor W. 1992. Statistical and linguistic analyses of F<sub>0</sub> in read and spontaneous speech. In *Proceedings of the International Conference on Spoken Language Processing* 1. 763–766.
- DEBRUYNE, Frans – DECOSTER, Wivine 1999. Acoustic differences between sustained vowels perceived as young or old. *Logopedics Phoniatrics Vocology* 24/1. 1–5.
- DECOSTER, Wivine – DEBRUYNE, Frans 2000. Longitudinal voice changes: Facts and interpretation. *Journal of Voice* 14/2. 184–193.
- DUFFY, Robert 1970. Fundamental frequency characteristics of adolescent females. *Language and Speech* 13/1. 14–24.

- EICHHORN, Julie T. – KENT, Raynold D. – AUSTIN, Diane – VORPERIAN, Houri K. 2018. Effects of aging on vocal fundamental frequency and vowel formants in men and women. *Journal of Voice* 32/5. 644.e1–644.e9. doi: 10.1016/j.jvoice.2017.08.003.
- FERGUSON, Charles A. 1964. Baby talk in six languages. *American Anthropologist* 66/6. 103–114.
- FITCH, James L. 1990. Consistency of fundamental frequency and perturbation in repeated phonations of sustained vowels, reading and connected speech. *Journal of Speech Hearing Disorders* 55/2. 360–363.
- FITCH, James L. – HOLBROOK, Anthony 1970. Modal vocal fundamental frequency of young adults. *Archives of Otolaryngology* 92/4. 379–382.
- FLETCHER, Annalise R. – MCAULIFFE, Megan J. – LANSFORD, Kaitlin L. – LISS, Julie M. 2015. The relationship between speech segment duration and vowel centralization in a group of older speakers. *The Journal of the Acoustical Society of America* 138/4. 2132–2139.
- FOUGERON, Cécile 2020. Multidimensional account of changes in speech throughout adulthood. Tudományos előadás az International Seminar on Speech Production 2020 konferencián. 2020. december 15. <https://www.youtube.com/watch?v=NOejo5uN66A&feature=youtu.be> (A letöltés ideje: 2021. január 5.)
- FOUQUET, Roger 2016. Historical energy transitions: speed, prices and system transformation. *Energy Research & Social Science* 22. 7–12.
- GERSTENBERG, Anette – FUCHS, Susanne – KARIET, Julie – FRANKENBERG, Claudia 2018. A cross-linguistic, longitudinal case study of pauses and interpausal units in spontaneous speech corpora of older speakers of German and French. In KLESSA, Katarzyna – BACHAN, Jolanta – WAGNER, Agnieszka – KARPIŃSKI, Maciej – ŚLEDZIŃSKI, Daniel (eds): *Proceedings of 9th International Conference on Speech Prosody* 2018, Poznan, Poland. 211–215.
- GILES, Howard – SMITH, Philip 1979. Accommodation theory: Optimal levels of convergence. In GILES, Howard – ST. CLAIR, R. (eds): *Language and Social Psychology*. Blackwell, Oxford. 45–65.
- GRÁCZI Tekla Etelka 2013. Zörejangok akusztikai fonetikai vizsgálata a zöngésségi oppozíció függvényében. Doktori disszertáció. ELTE BTK, Budapest. <http://doktori.btk.elte.hu/lingv/gracztekla/diss.pdf>. (A letöltés ideje: 2020. november 1.)
- GRÁCZI Tekla Etelka 2016. A zöngékezdési időről. In BÓNA Judit (szerk.): *Fonetikai olvasókönyv*. Egyetemi E-jegyzet. ELTE Fonetikai Tanszék, Budapest. 61–75. <http://fonetikaitanszek.elte.hu/wp-content/uploads/2017/11/OlvasokonyvTeljes.pdf>. (A letöltés ideje: 2020. május 1.)
- GRÁCZI, Tekla Etelka – CSAPÓ, Tamás Gábor – BARTÓK, Márton – DEME, Andrea – MARKÓ, Alexandra 2020a. The realization of voicing opposition in alveolar fricatives in Hungarian. Preliminary study on articulation and acoustics. *Beszédtudomány – Speech Science* 1. 22–56. DOI: 10.15775/Besztud.2020.22-56
- GRÁCZI Tekla Etelka – CSAPÓ Tamás Gábor – DEME Andrea – JUHÁSZ Kornélia – MARKÓ Alexandra 2020b. A réshangok zöngésségével összefüggő nyelvpozíciós jellemzők a megelőző magánhangzóban. *Nyelvtudományi Közlemények* 116. 155–190. 2060-7644. [http://real.mtak.hu/119160/1/teljes\\_kotet\\_v5.pdf](http://real.mtak.hu/119160/1/teljes_kotet_v5.pdf) (A letöltés ideje: 2021. május 2.)

- GRÁCZI Tekla Etelka – GÓSY Mária – KREPSZ Valéria – MARKÓ Alexandra – HUSZÁR Anna – DAMÁSDI Nóra – GOCSÁL Ákos 2020c. Az alapfrekvencia jellemzői az életkor és nem függvényében. In FÓRIS Ágota – BÖLCSKEI Andrea (főszerk.)– BÓNA Judit – GRÁCZI Tekla Etelka – MARKÓ Alexandra (szerk.): *Nyelv, kultúra, identitás. Alkalmazott nyelvészeti kutatások a 21. századi információs térben: III. Fonetika*. Akadémiai Kiadó, Budapest. [https://mersz.hu/dokumentum/m675nyki3f\\_\\_25](https://mersz.hu/dokumentum/m675nyki3f__25) (A letöltés ideje: 2020. november 2.)
- GRÁCZI Tekla Etelka – HUSZÁR Anna – KREPSZ Valéria – SZÁRAZ Bettina – DAMÁSDI Nóra – MARKÓ Alexandra 2020d. Longitudinális korpusz magyar felnőtt adatközlőkről. In BEREND Gábor – GOSZTOLYA Gábor – VINCZE Veronika (szerk.): *XVI. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem, Informatikai Intézet, Szeged. 103–114. [https://rgai.inf.u-szeged.hu/sites/rgai.inf.u-szeged.hu/files/mszny2020\\_0.pdf](https://rgai.inf.u-szeged.hu/sites/rgai.inf.u-szeged.hu/files/mszny2020_0.pdf) (A letöltés ideje: 2020. augusztus 20.)
- GRÁCZI Tekla Etelka – KREPSZ Valéria – MARKÓ Alexandra – HUSZÁR Anna – SZÁRAZ Bettina 2019. Az  $f_0$ -jellemzők felolvasásban és spontán beszédben. *Alkalmazott Nyelvtudomány* 19/2. doi:<http://dx.doi.org/10.18460/ANY.2019.2.003>
- HARRINGTON, Jonathan – PALETHORPE, Sallyanne – WATSON, Catherine I. 2007. Age-related changes in fundamental frequency and formants: A longitudinal study of four speakers. In *Proceedings of Interspeech*. 2753–2756.
- HASEK, Carol S. – SINGH, Sadanand – MURRY, Thomas 1980. Acoustic attributes of preadolescent voices. *The Journal of the Acoustical Society of America* 68/5. 1262–1265.
- HOLLIEN, Harry – GREEN, Rachel – MASSEY, Karen 1994. Longitudinal research on adolescent voice change in males. *Journal of Speech and Hearing Research* 96/5. 2646–2654.
- HOLLIEN, Harry – HOLLIEN, Patricia A. – DE JONG, Gea 1997. Effects of three parameters on speaking fundamental frequency. *Journal of the Acoustical Society of America* 102/5. 2984–2992.
- HOLLIEN, Harry – JACKSON, Bernard 1973. Normative data on the speaking fundamental frequency characteristics of young adult males. *Journal of Phonetics* 1/2. 117–120.
- HOLLIEN, Harry – MALCIK, Ellen 1967. Evaluation of cross-sectional studies of adolescent voice change in males. *Speech Monograph* 34/1. 80–84.
- HOLLIEN, Harry – SHIP, Thomas 1972. Speaking fundamental frequency and chronological age in males. *Journal of Speech and Hearing Research* 15/1. 155–159.
- HUBER, Jessica E. 2008. Effects of utterance length and vocal loudness on speech breathing in older adults. *Respiratory Physiology & Neurobiology* 164/3. 323–330.
- HUDSON, Amelia I. – HOLBROOK, Anthony 1982. Fundamental frequency characteristics of young black adults: Spontaneous speaking and oral reading. *Journal of Speech and Hearing Research* 25/1. 25–28.
- HUSZÁR Anna – KREPSZ Valéria 2021. A beszélőn belüli variabilitás alakulása a beszéd típus, a szövegbeli helyzet és az életkor mentén. *Magyar Nyelvőr* 145/1. 59–80.
- JOOS, Martin 1968. The isolation of styles. In FISHMANN, Joshua A. (eds): *Readings In the Sociology of Language*. Mouton, The Hague. 185–191.



- KEATING, Patricia – KUO, Grace 2010. Comparison of speaking fundamental frequency in English and Mandarin. *UCLA Work Papers in Phonetics* 108. 164–187.
- KETCHAM, Caroline J. – SEIDLER, Rachael D. – VAN GEMMERT, Arend W. – STELMACH, George E. 2002. Age-related kinematic differences as influenced by task difficulty, target size, and movement amplitude. *Journals of Gerontology: Psychological Sciences and Social Sciences* 57/1. 54–64.
- KISS Gábor 2021. A depresszió automatikus becslése a beszéd akusztikai-fonetikai jellemzői alapján. In MARKÓ Alexandra (szerk.): *Tanulmányok a beszédtudomány alkalmazásainak köréből*. ELTE Eötvös Kiadó, Budapest. 65–86.
- KROOK, M. I. Pegoraro 1988. Speaking fundamental frequency characteristics of normal Swedish subjects obtained by glottal frequency Analysis. *Folia Phoniatria* 40/2. 82–90.
- KÜNZEL, Hermann J. 1989. How well does average fundamental frequency correlate with speaker height and weight? *Phonetica* 46/1–3. 117–125.
- LABOV, William 1972. *Sociolinguistic Patterns*. University of Pennsylvania Press, Philadelphia.
- LABOV, William 1986. Social stratification of (r) in New York department stores. *Dialect and Language Variation* 1986. 304–329. <https://doi.org/10.1016/B978-0-12-051130-3.50029-X>
- LABOV, William 1994. *Principles of Linguistic Change, Internal Factors*. Wiley-Blackwell, Oxford.
- LALLEY, Peter M. 2013. The aging respiratory system – Pulmonary structure, function and neural control. *Respiratory Physiology & Neurobiology* 187/3. 199–210.
- LAMY, Peter P. 1991. Physiological changes due to age. *Drugs & Aging* 1. 385–404. <https://doi.org/10.2165/00002512-199101050-00006>
- LE, Xuan – LANCASHIRE, Ian – HIRST, Graeme – JOKEL, Regina 2011. Longitudinal detection of dementia through lexical and syntactic changes in writing: A case study of three British novelists. *Literary and Linguistic Computing* 26/4. 435–461.
- LINDBLOM, Björn 1990. Explaining phonetic variation: A sketch of the H&H theory. In HARDCASTLE, William – MARCHAL, Alain (eds): *Speech Production and Speech Modeling*. Kluwer, Dordrecht. 403–439.
- LINVILLE, Sue Ellen 1996. The sound of senescence. *Journal of Voice* 10/2. 190–200.
- LINVILLE, Sue Ellen 2000. The aging voice. Voice Quality Measurement. In KENT, Raymond D. – BALL Martin J. – DIEGO, San (eds): *Thomson Learning*. 359–376.
- LINVILLE, Sue E. – RENS, Jennifer 2001. Vocal tract resonance analysis of aging voice using Long-Term Average spectra. *Journal of Voice* 15/3. 323–330.
- MÁDY, Katalin – REICHEL, Uwe D – SZALONTAI, Ádám – KOHÁRI, Anna – DEME, Andrea 2018. Prosodic characteristics of infant-directed speech as a function of maternal parity. In KLESSA, Katarzyna – BACHAN, Jolanta – WAGNER, Agnieszka – KARPIŃSKI, Maciej – ŚLEDZIŃSKI, Daniel (eds): *9th International Conference on Speech Prosody 2018*. International Speech Communication Association (ISCA), Poznan. 294–298.
- MARKÓ Alexandra 2015. *A spontán beszéd prozódiai szerkezete. Időzítés és beszéddallam*. Nyelvtudományi Értekezések 166. sz. Akadémiai Kiadó, Budapest.

- MARKÓ Alexandra 2017. Hangtan. In IMRÉNYI András – KUGLER Nóra – LADÁNYI Mária – MARKÓ Alexandra – TÁTRAI Szilárd – TOLCSVAI NAGY Gábor: *Nyelvtan*. Osiris Kiadó, Budapest. 77–198.
- MAJEWSKI, Wojciech – HOLLIEN, Harry – ZALEWSKI, Janusz 1972. Speaking fundamental frequency characteristics of Polish adult males. *Phonetica* 25/2. 119–125.
- MISONO, Stephanie – MARMOR, Schelomo – ROY, Nelson – MAU, Ted – COHEN, Seth M. 2016. Multi-institutional study of voice disorders and voice therapy referral: Report from the CHEER Network. *Journal of Otolaryngology-Head & Neck Surgery* 155/1. 33–41.
- MYSAK, Edward 1959. Pitch and duration characteristics of older males. *Journal of Speech, Language, and Hearing Research* 2/1. 46–54.
- MWANGI, Samuel – SPIEGL, Werner – HÖNIG, Florian – HADERLEIN, Tino – MAIER, Andreas – NÖTH, Elmar 2009. Effects of vocal aging on fundamental frequency and formants. In Acoustical Society of the Netherlands (NAG) – German Acoustical Society (DEGA) (eds): *Proceedings of the International Conference on Acoustics NAG/DAGA*. 1761–1764.
- NARAYAN, Chandan R. – C. McDERMOTT, Lilly 2016. Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. *The Journal of the Acoustic Society of America* 139/3. 1272–1281.
- NEUBERGER, Tilda – GYARMATHY, Dorottya – GRÁCZI, Tekla Etelka – HORVÁTH, Viktória – GÓSY, Mária – BEKE, András 2014. Development of a large spontaneous speech database of agglutinative Hungarian language. In SOJKA, Petr – HORÁK, Ales – KOPEČEK, Ivan – PALA, Karel (eds): *Proceedings of TSD 2014*. Heidelberg–Springer, New York – Berlin. 424–431.
- NISHIO, Masaki – NIIMI, Seiji 2008. Changes in speaking fundamental frequency characteristics with aging. *Folia Phoniatrica et Logopedica* 60/3. 120–127. DOI: 10.1159/000118510
- OHALA, John J. 1983. The origin of the sound patterns in vocal tract constraints. In MACNEILAGE, Peter F. (ed.): *The Production of Speech*. Springer Verlag, New York. 189–216.
- DE PINTO, Olive – HOLLIEN, Harry 1982. Speaking fundamental frequency characteristics of Australian women: Then and now. *Journal of Phonetics* 10/4. 367–375.
- POWELL, Marsha – FILTER, Maynard D. – WILLIAMS, Barbara 1989. A longitudinal study of the prevalence of voice disorders in children from a rural school division. *Journal of Communication Disorders* 22/5. 375–382.
- QUENÉ, Hugo 2013. Longitudinal trends in speech tempo: The case of Queen Beatrix. *The Journal of the Acoustical Society of America* 133/6. 452–457.
- RASTATTER, Michael P. – JACQUES, Richard D. 1990. Formant frequency structure of the aging male and female vocal tract. *Folia Phoniatrica et Logopaedica* 42/6. 312–319.
- REUBOLD, Ulrich – HARRINGTON, Jonathan 2015. Disassociating the effects of age from phonetic change. In GERSTENBERG, Annette – VOESTE, Anja (eds): *Language development*. John Benjamins Publishing Company, Amsterdam. 9–38.
- REUBOLD, Ulrich – HARRINGTON, Jonathan – KLEBER, Felicitas 2010. Vocal aging effects on  $F_0$  and the first formant: A longitudinal analysis in adult speakers. *Speech Communication* 52/7–8. 638–651.

- ROACH, Peter – KNOWLES, Gerry – VARADI, Tamás – ARNFIELD, Simon 2009. MARSEC: A Machine-Readable Spoken English Corpus. *Journal of the International Phonetic Association* 23/2. 47–53.
- SCUKANEC, Gail P. – PETROSINO, Linda – SQUIBB, Kevin 1991. Formant frequency characteristics of children, young adult, and aged female speakers. *Perceptual and Motor Skills* 73/1. 203–208.
- SEBASTIAN, Swapna – BABU, Sonia – OOMMEN, Neethu – BALLRAJ, Achamma 2012. Acoustic measurements of geriatric voice. *Journal of Laryngology and Voice* 2/2. 81–84.
- SEMSEI Imre 2014. *Gerontológia*. Debreceni Egyetem Egészségügyi Kar, Debrecen.
- DE SILVA, Viola – ANTTI, Iivonen – BONDARKO, Liya V. – POLS, Louis C. W. 2003. Common and language dependent phonetic differences between read and spontaneous speech in Russian, Finnish and Dutch. In *Proceedings 15th International of Phonetic Sciences ICPHS*. 2977–2980.
- SKARNITZL, Radek – VAŇKOVÁ, Jitka 2017. Fundamental frequency statistics for male speakers of common Czech. *Acta Universitatis Carolinae Philologica* 3. 7–17.
- SNIDECOR, John C. 1944. A comparative study of the pitch and duration characteristics of impromptu speaking and oral reading. *Speech Monographs* 10/1. 50–56.
- SORENSEN, David – HORII, Yashiyuki 1982. Cigarette smoking and voice fundamental frequency. *Journal of Communication Disorders* 15/2. 135–144.
- STATHOPOULOS, Elaine T. – HUBER, Jessica E. – SUSSMAN, Joan E. 2011. Changes in acoustic characteristics of the voice across the life span: Measures from individuals 4–93 years of age. *Journal of Speech, Language, and Hearing Research* 54/4. 1011–1021.
- SWERTS, Marc – STRANGERT, Eva – HELDNER, Mattias 1996. F<sub>0</sub> declination in read-aloud and spontaneous speech. In *Proceedings of Fourth International Conference on Spoken Language* 96. 1501–1504.
- SZÁRAZ Bettina – GRÁCZI Tekla Etelka 2020. Explozívák realizációja fiatal és öregedő felnőttek beszédében. *Beszédtudomány – Speech Science* 1. 73–98. DOI: 10.15775/Besztud.2020.73-98
- SZILI Katalin 2004. *Tetté vált szavak*. Tinta Könyvkiadó, Budapest.
- TATÁR Zoltán – VARGA Zoltán – FŐZŐ Eszter 2021. Beszélőazonosítás a kriminalisztikában. In MARKÓ Alexandra (szerk.): *Tanulmányok a beszédtudomány alkalmazásainak köréből*. ELTE Eötvös Kiadó, Budapest. 229–250.
- TORRE, Pete – BARLOW, Jessica A. 2009. Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders* 42/5. 324–333.
- TRUDGILL, Peter J. 1978. Introduction: Sociolinguistics and sociolinguistics. In TRUDGILL, Peter J. (ed.): *Sociolinguistic Patterns in British English*. Edward Arnold, London. 1–18.
- TULICS Miklós Gábor 2021. A diszfónia és automatikus felimerése. In MARKÓ Alexandra (szerk.): *Tanulmányok a beszédtudomány alkalmazásainak köréből*. ELTE Eötvös Kiadó, Budapest. 33–63.
- WACHA Imre 1974. Az elhangzó beszéd főbb akusztikus stíluskategóriáiról. *Általános Nyelvészeti Tanulmányok* 10. 203–216.

- XUE, Steve A. – HAO, Grace J. 2003. Changes in the human vocal tract due to aging and the acoustic correlates of speech production. *Journal of Speech Language and Hearing Research* 46/3. 689–701.
- YAMAZAWA, Hiromi – HOLLIEN, Harry 1992. Speaking fundamental frequency patterns of Japanese women. *Phonetica* 49/2. 128–140.
- ZWICKY, Arnold 1972. On causal speech. In *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, Chicago. 607–615.

### **KÖSZÖNETNYILVÁNÍTÁS**

A kutatás az NKFIH FK-128814-es pályázatának keretében valósul meg.



# Beszélőazonosítás a kriminalisztikában

Tatár Zoltán

NEMZETBIZTONSÁGI SZAKSZOLGÁLAT SZAKÉRTŐI INTÉZET

Varga Zoltán

NEMZETBIZTONSÁGI SZAKSZOLGÁLAT SZAKÉRTŐI INTÉZET

Főző Eszter

NEMZETBIZTONSÁGI SZAKSZOLGÁLAT SZAKÉRTŐI INTÉZET

## 1. A beszélőazonosítás fogalma

Az emberek a hangjuk alapján általában felismerhetők, hiszen „az emberi beszédben előforduló hangjelenségek mindig nagymértékben egyéniek és egyszerűek” (LAZICZIUS 1979: 5).

Idővel azonban olyan változások mehetnek végbe a beszédünkben, amely következtében akár az is előfordulhat, hogy egy rég nem látott ismerősünk például telefonban nem ismeri fel a hangunkat. A beszédsajátosságok időbeli változása a nyomozó hatóságok és a bíróságok számára is nehézségeket okozhat. Előfordulhat ugyanis olyan eset, amikor egy több, akár tíz évvel korábbi kérdéses hanganyag alapján a beszélő személy kilétét kell megállapítani. A hatóságok igazságügyi szakértőkhöz fordulhatnak, akik szakértői véleménnyel segítik a munkájukat. Két eset állhat fenn: amennyiben van az ügyben gyanúsított, akitől feltételezhetően származhat a kérdéses szöveg, úgy a szakértő hangmintát vesz tőle, és elvégzi a **beszélőazonosítást**, azaz megvizsgálja a gyanúsított személy és a kérdéses hangfelvételen szereplő beszélő azonosságának feltételezhető mértékét. Az így készülő szakvélemény a bíróságon bizonyítékként használható. Akkor azonban, ha nincs feltételezett elkövető, a szakértők **beszélőprofil** készíthetnek a kérdéses beszélőről, azaz az elhangzott szöveg alapján megállapítható szociodemográfiai és egyéb jellemzőket gyűjtik össze (pl. nem, életkor, iskolázottság, nyelvjárás, beszédhiba). Ez segítheti a nyomozást, hiszen az elkövető(k) feltételezett tulajdonságait körvonalazza (TATÁR 2013a).

Ezeket a szakvéleményeket két szakterület képviselői készíthetik: a hangtechnikai és a nyelvész szakértők, akik sok esetben összedolgoznak. A **hangtechnikai szakértő** a kérdéses hangfelvételeken percepció és akusztikai-fonetikai vizsgálatokat, valamint a felvétel eredetiségének elemzését végzi. A **nyelvész szakértő** szövegnyelvészeti, megfelelő szakértelemmel pszicho-, továbbá szociolingvisztikai elemzéseket végezhet. Ezen túlmenően az adott munkához szükséges egyéb szakterület képviselője – például pszichológus vagy egy adott idegen nyelvet ismerő személy – is bevonható az elemzésekbe. Abban az esetben, amikor rendelkezésre áll gyanúsítottól származó hangminta, és minden vizsgálandó hanganyag megfelelő minőségű és hosszúságú, akkor a hangtechnikai szakértő által önállóan is elvégezhető a beszélőazonosítás, és erős valószínűségi szintű vélemény alakítható. Ha viszont bármilyen oknál fogva hangtechnikai szakértői

módszerekkel kevesebb bizonyítékkal támasztható alá a vélemény, úgy a nyelvész szakértő segíthet, sőt speciális esetben akár önállóan is elvégezheti a beszélő személy azonosítását.

A következőkben bemutatjuk, hogy a Nemzetbiztonsági Szakszolgálat (NBSZ) Szakértői Intézetében milyen módszerekkel végezzük a beszélőazonosítást, illetve milyen kutatásokat és módszertani fejlesztéseket végzünk, különös figyelmet fordítva a beszélők beszédsajátosságainak feltérképezésére, összevethetőségére, valamint a beszédsajátosságok változásainak detektálására.

Beszélőazonosítás során a hangtechnikai szakértő hallásalapú (percepciós), fonetikai alapú (számítógépes mérésekkel támogatott) és biometrikus (automatikus) összehasonlító vizsgálatokat végez, és egy valószínűségi skála mentén alkotja meg a véleményét. Ez a skála szakértőnként, illetve szakértői intézetenként kismértékben eltérő lehet, az NBSZ Szakértői Intézetében az alábbi 9 fokozatú skálát használjuk (a lenti felsorolásban a *beszédhang* terminus nem nyelvészeti értelemben, azaz a beszéd legkisebb egységeként értendő, hanem az igazságügyi hangszakértői módszertani leírásokban szereplő módon alkalmazzuk, 'beszédminta' értelemben):

- +4 A két beszédhang nagyon nagy valószínűséggel azonos személytől származik.
- +3 A két beszédhang nagy valószínűséggel azonos személytől származik.
- +2 A két beszédhang valószínűleg azonos személytől származik.
- +1 A két beszédhang származhat azonos személytől.
- 0 A beszélők azonossága nem zárható ki, de nem is bizonyítható.
- 1 A két beszédhang származhat különböző személytől.
- 2 A két beszédhang valószínűleg különböző személytől származik.
- 3 A két beszédhang nagy valószínűséggel különböző személytől származik.
- 4 A két beszédhang nagyon nagy valószínűséggel különböző személytől származik.

Ezek a kategóriák a módszertani fejlesztések következtében változhatnak. A jelenlegi módszertan egyik említést érdemlő jellegzetessége, hogy nem adunk kategorikus véleményt, vagyis nem állítjuk azt, hogy a két beszédhang kétséget kizáróan azonos vagy különböző személytől származik. Ezt az automatikus beszélőazonosítás nemzetközi térhódításának következtében kezdtük el alkalmazni, ugyanis a hangtechnikai szakértői területen is nagy jelentőséget kapott a Bayes-analízisen alapuló statisztikai szemlélet (ROSE 2002), amely szerint gyakorlatilag soha nem lehet kijelenteni 100%-os biztossággal, hogy a két beszélő azonos/különböző, mert az ellenkezőjének is mindig van statisztikai esélye (TATÁR–OSZTERTÁG–MORBER 2019). Ezen álláspont bírálói azzal érvelnek, hogy az összehasonlítás eredményei alapján a szakértő vonja le a következtetést, és ha a beszélők azonosságának esélyéhez képest elhanyagolhatóan kicsi a különbségük esélye, akkor igenis van létjogosultsága a kategorikus véleménynek, mert a szakértő ezzel jobban tudja segíteni a bíróságot a döntéshozatalban.

## 2. Kérdésfelvetés

Egy hangfelvételen beszélő kérdéses személy és a gyanúsított hangjának összehasonlításakor sok nehézség adódhat, ilyen például, ha rossz minőségű vagy rövid a rendelkezésre álló hanganyag (zajos, frekvenciaszűrt, pl. telefon), ha torzított a beszéd (pl. suttogott), vagy ha

az inkriminált és az összehasonlítandó hangfelvétel rögzítése között hosszabb idő, akár évek teltek el. Ez utóbbi esetben annál nehezebb a feladat, minél nagyobb az összehasonlítandó hanganyagok felvételi időpontja közti távolság. A beszédsajátosságok változásai miatt máshogy kell kezelni egy beszélőnek egy 10 évvel ezelőtti és egy jelenlegi hanganyagát. Feltételezhető, hogy egy beszélő hangját az egy évtizeddel korábbi saját hangjával összehasonlítva nagyobb különbség mutatkozik, mint a beszélő két, egymáshoz közeli időpontban rögzített beszédmintájának összevetésekor. Ezen egyszerű oknál fogva a hangtechnikai szakértőnek az összehasonlítás eredményeinek értékelésekor és a véleményalkotáskor is figyelembe kell vennie az időfaktort. A kérdés az, vajon milyen mértékben. Ahhoz, hogy a szakértői beszélőazonosítás a bíróságon minden körülmények között megalapozott legyen, szükséges tudni a választ erre a kérdésre. A helyes válaszhoz pedig – a szakértői tapasztalaton túlmenően – a beszédsajátosságok változásainak részletes megismerése szükséges, amely az utánkövetéses adatbázison végzett kutatások alapján lehetséges.

A Nemzetbiztonsági Szakszolgálat Szakértői Intézete és a Nyelvtudományi Kutatóközpont (korábban Intézet) közötti szakmai-tudományos együttműködés keretében megkezdődött a beszédsajátosságok változásainak vizsgálata, kifejezetten gyakorlati megközelítésből, amelynek céljai között szerepel az igazságügyi hangtechnikai szakértői és az igazságügyi nyelvész szakértői módszertan fejlesztése.

Az egyéni beszédstílust a fonetikai, a nyelvhasználati, a fiziológiai, a pszichikai és az egyéb sajátosságok összessége adja. Ebből kiindulva elfogadható az az alapvetés, miszerint minél több paramétert hasonlítunk össze, annál pontosabb és jobban alátámasztott azonosítást és véleményt alkothatunk. Azonban a szakértő nem végezhet mindenre kiterjedő, időigényes vizsgálatokat, ezért meg kellett találni azt az egy-két jellemzőt, amelyek nagymértékben meghatározzák az egyén beszédét. Ilyen jellemzők – a fonetikai alapú beszélőazonosításban széles körben elfogadottan (MCMENAMIN 2002; JESSEN 2010) – a formánsok és az alapfrekvencia.

Az ember hangképző szerveire jellemző biológiai egyediségnek a jeleit többek között a formánsok – „a zöngének a rezonátorüregben felerősödött felharmonikusai” (GÓSY 2004: 99) – is magukon viselik, ezért két beszélő összehasonlításával a magánhangzók formánsai alapján következtethetünk a beszélők artikulációjának egyes sajátosságaira, illetve a hangképző szervek hasonlóságára. Az első két formáns ( $F_1$  és  $F_2$ ) „a magyarban rendszerint egyértelműen meghatározza a magánhangzó minőségét” (GÓSY 2004: 116), tehát a magánhangzó nyelvi azonosítását, a különféle magánhangzók megkülönböztetését teszi lehetővé. Ez akaratunktól és egyéb körülményektől függően különböző lehet, hiszen más-más stílusban, lelkiállapotban stb. szoktunk beszélni, illetve idővel, évek alatt is változik. Emellett az  $F_1$  és  $F_2$  különböző beszélők közti különbsége (éppen amiatt, hogy ezek szerepe a nyelvi azonosíthatóság biztosítása) sok esetben nem mutatható ki. Az egyedi azonosításhoz van segítségünk a 3. és a 4. formáns ( $F_3$ ,  $F_4$ ), ugyanis ezek az ejtett beszédhangtól kevésbé függenek, az egyénre jellemzőbbek, vagyis ha ezeket összehasonlítjuk, és hasonlóak, akkor az a két beszélő azonosságát erősíti. Az  $F_3$ ,  $F_4$  ugyanakkor nem önmagában jellemző az adott beszélőre, hanem az első két formánsal együtt. Ha a két beszélő ugyanazt a beszédhangot egyformán ejti (pl. az *e* hangot hasonló  $F_1$ -,  $F_2$ -értékekkel), valamint az  $F_3$  és az  $F_4$  is hasonló értékeket mutat, akkor tekinthető az adott



eredmény bizonyítéknak a beszélők azonossága mellett. (A formánsok beszélők közötti variabilitásának és a beszélőazonosításban betöltött szerepének összefoglalását és elemzését lásd pl. CAO, DELLWO 2019.)

A formánsok összehasonlítási módszere annyira nagy jelentőségű, hogy ha a hangfelvétel gyenge minősége miatt nem jelenik meg a spektrogramon a 3., 4. formáns, akkor fonetikai alapú beszélőazonosításra alkalmatlan besorolást kap a hanganyag. Ennek következtében az egyéb vizsgálatokkal, beleértve az automatikus beszélőazonosítást és a nyelvész szakértői összehasonlítást is, a jelenlegi módszertanunk szerint legfeljebb csak „valószínűleg azonos” vagy „valószínűleg különböző” megállapítás tehető. Speciális esetben, amikor a kérdéses hangfelvételek rossz minőségűek, de az elhangzottak érthetőek, és nagy mennyiségű hangzó szöveg áll rendelkezésre, akkor szövegnyelvészeti összehasonlítás és szakvélemény önállóan is készülhet a nyelvész szakértői módszerek alkalmazásával.

A hangtechnikai szakértői módszertan fejlesztéséhez elengedhetetlen ismerni, hogy egy személy több év különbséggel rögzített beszédének összehasonlításakor a formánsok és az alfrekvencia-értékek milyen mértékben változnak. Az adatok milyen mértékű eltérése engedi meg az azonosságot, és mekkora különbség utal már különböző beszélőkre? Ezekre a kérdésekre célzott kutatások adhatnak választ. Jelenleg a Nyelvtudományi Kutatóközponttal zajló együttműködés keretében a Longitudinális korpuszon 10 év elteltével rögzített hanganyagokon végzünk elemzéseket az egy évtized elteltével bekövetkező változások mintázatainak kimutatása céljából (vö. GRÁCZI–KREPSZ–HUSZÁR 2021 a jelen kötetben).

A fonetikai elemzés mellett automatikus beszélőazonosítási kutatásokat is tervezünk, hiszen a biometrikus azonosítás a hangtechnikai szakértői véleményalkotás fontos részét képezi. A 10 évvel korábbi felvételek felhasználásával végzett elemzések is szükségesek ahhoz, hogy egy teljes beszélőazonosítási módszertant kidolgozzunk arra a különleges esetre, amikor nagy időkülönbséggel rögzített hanganyagok összehasonlításával alkotunk véleményt.

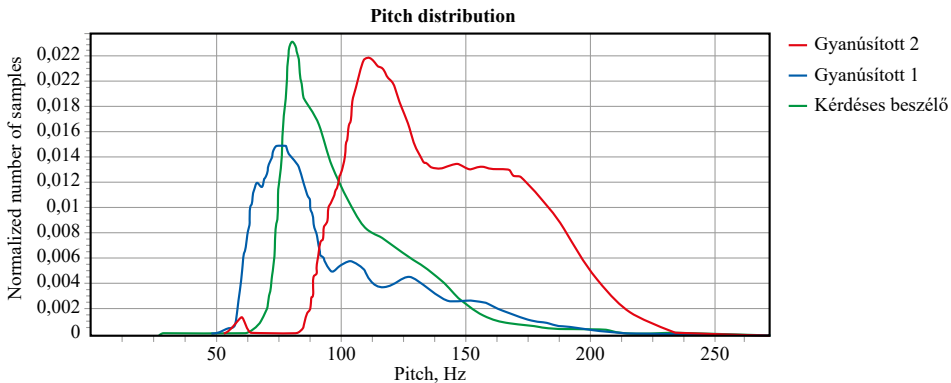
### 3. A beszélőazonosítás folyamatának bemutatása: a hangtechnikai szakértői elemzés

Ebben a részben a Longitudinális korpusz felhasználásával végzett előzetes kutatásokat mutatjuk be. A gyakorlatban is használatos kifejezéseket alkalmazva vezetjük végig az esettanulmány-jellegű vizsgálatot, mintha egy igazságügyi beszélőazonosítási feladat lenne. Az összehasonlításához három felvételt használunk fel. A kérdéses felvétel egy 2009-ben rögzített hanganyag, amely egy 25 éves egyetemi hallgató férfi beszédfelvétele (kérdéses beszélő). Az összehasonlító (a gyanúsítottaktól származó) minták pedig ugyanezen beszélő 11 évvel későbbi (2020-ban felvett) beszédanyaga („Gyanúsított 1”) és egy öt évvel idősebb férfi beszédfelvétele („Gyanúsított 2”) lesznek.

### 3.1. Percepció és akusztikai-fonetikai elemzések

A percepció (hallásalapú) elemzéseket az ügy jellegéhez igazítva elvégezhetjük kevésbé részletesen vagy nagyon részletesen. Amikor a szakértő meghallgatja a hanganyagokat, akkor a beszélők hangmagassága, beszédtempója, egyedi beszédjellemzői (pl. beszédhibák) alapján kialakul benne egy hipotézis a beszélők azonosságára kérdésében. Abban az esetben, amikor jól mérhetőek a formánsok, és az automatikus azonosító is erősen alátámasztja a beszélők azonosságát vagy különbözőségét – ahogy jelen esetben is –, akkor a percepció összehasonlítás alapján csak egy rövid megállapítást teszünk, például azt, hogy a kérdéses beszélő és a „Gyanúsított 1” férfi beszéde között nagy fokú hasonlóság, míg a kérdéses személy és a „Gyanúsított 2” viszonyában különbség tapasztalható.

A következő lépés a fonetikai paraméterek összehasonlítása, amelyet az ACU-EXPERT (<https://acustek.com/en/>) nevű szakértői szoftverrel végzünk el. Elsőként az átlagos alaphangfrekvencia-értékeket vizsgáljuk, és az 1. ábrán látható módon szemléltetjük az eredményt. A vízszintes tengelyen az alaphangfrekvencia értéke, a függőleges tengelyen pedig az adott frekvencia relatív gyakorisága látható, azaz ahol a legmagasabb az adott görbe csúcsa, azt az alaphangfrekvenciát (vagy frekvenciatartományt) használja a leggyakrabban a beszélő. Az ábrán mindhárom beszélőnél megfigyelhető, hogy a leggyakoribb alaphangfrekvencián kívül egy rendszeresen használt magasabb tartomány is megjelenik, és egyértelműen leolvashatóak a görbék közötti hasonlóságok és különbségek. Megállapítható, hogy a kérdéses beszélő és a „Gyanúsított 1” alaphangfrekvencia-értékei hasonlóbbak, míg a „Gyanúsított 2” tőlük nagyobb mértékben tér el. Ez az eredmény összhangban áll a percepció vizsgálat során megfogalmazott hipotézissel.



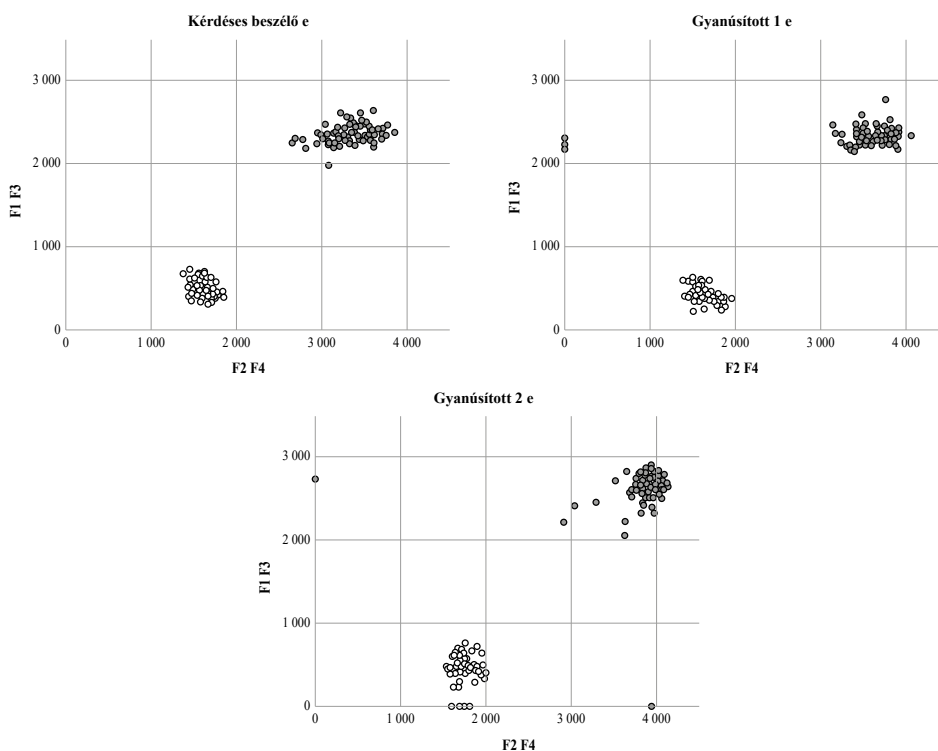
1. ábra

*Az alaphangfrekvencia-értékek histogramjai (folytonos vonallal jelölt: a kérdéses beszélő; pontozott vonallal jelölt: „Gyanúsított 1”; szaggatott vonallal jelölt: „Gyanúsított 2”)*

A következő lépés a formánsok összehasonlítása. Ezen belül vizsgálhatjuk a formánsmenetet a spektrogramon (hangszínekép, amely megmutatja a hangspektrum időbeli változását  $x = \text{idő}$ ,  $y = \text{frekvencia}$  grafikonban, és a hangerősséget színskála szerint szemlélteti; erre látható példa

a 3. ábrán), amely egyedi lehet, továbbá FFT- (*Fast Fourier Transformation*, amely kirajzolja az egy rövid időintervallumban mért spektrális összetevőket az  $x$  = frekvencia,  $y$  = hangerősség grafikonon; erre szolgáltat példát a 4. ábra) és LPC- (*Linear Predictive Coding*, amely megbecsüli az egy rövid időintervallumban mért spektrális összetevőket az  $x$  = frekvencia,  $y$  = hangerősség grafikonban) görbék segítségével hasonlíthatjuk össze a hangok spektrális összetevőit, illetve a formánsértékek összességét is elemezhetjük, amelyek szintén egyénre jellemzőek.

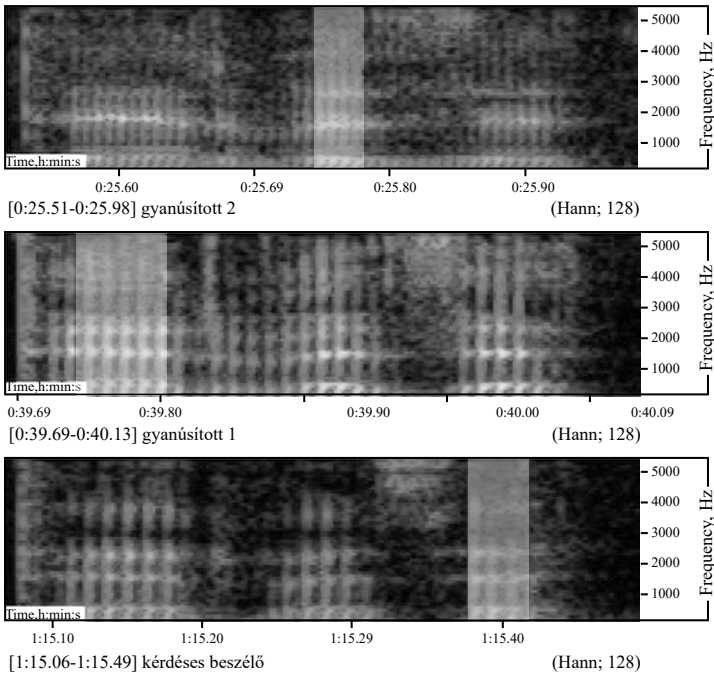
A 2. ábra paneljei a hanganyagokban előforduló összes  $e$  hang formánsértékeit mutatják ( $x$  = frekvencia,  $y$  = frekvencia). A 2. ábra grafikonjait összehasonlítva megállapítható, hogy az  $F_1$  és az  $F_2$  mindhárom beszélőnél hasonló értékeket mutat, az  $F_3$ -,  $F_4$ -értékek tekintetében viszont csak a „Gyanúsított 1” hasonlít a kérdéses beszélőre. A „Gyanúsított 2” 3., 4. formánsai ugyanis szinte minden esetben magasabb értékeket vesznek fel. Ez azt jelenti, hogy az  $e$  hangot hasonló üregi rezonanciával ( $F_1$ ,  $F_2$ ) ejti mindhárom beszélő, és a 3., 4. formáns alapján tudunk következtetni a beszélők azonosságára/különbözőségére. Ez a vizsgálat is megerősíti a korábbi hipotézist, és újabb bizonyítékot szolgáltat a kérdéses beszélő és a „Gyanúsított 1” azonosságához, illetve a kérdéses beszélő és a „Gyanúsított 2” különbözőségéhez.



2. ábra

A vizsgált beszélők  $e$  hangjainak formánsértékei (Hz): a bal felső panelen a kérdéses beszélő; a jobb felső panelen a „Gyanúsított 1”-é; az alsó panelen a „Gyanúsított 2”-é (szürke [felső]:  $F_1 \times F_2$ , fehér [alsó]:  $F_3 \times F_4$ )

A következő vizsgálat a beszédhangok egyes előfordulásain belüli formánsmenetek összehasonlítása. A 3. ábrán látható spektrogramokon megfigyelhetők a kijelölt beszédhangok (a kérdéses beszélőnél a *tervezett* szó harmadik *e* hangja, a „Gyanúsított 1”-nél a *tervezett* szó első *e* hangja, a „Gyanúsított 2”-nél a *tervezett* szó második *e* hangja), amelyekből elkészítettük a 4. ábrán látható LTA-görbéket (*Long Term Average*, amely kirajzolja a kijelölt időintervallumban mért átlagolt spektrális összetevőket az  $x$  = frekvencia,  $y$  = intenzitás grafikonban). Az összehasonlított hangok, amelyeket a 3. ábrán kiemeléssel jelöltünk, nem azonos hangkörnyezetből származnak, de a módszertanunk szerint nem ez a követelmény, hanem az, hogy stabil formánsok és hasonló  $F_1$ -,  $F_2$ -értékek legyenek mérhetőek. Az összehasonlítandó hangok akár három különböző szóból is származhatnak, a legbiztosabban akkor lehet megállapítani a beszélők azonosságát vagy különbözőségét, ha az első két formáns minden beszélőnél hasonló helyen van, ezért választottuk ezeket a hangokat, a harmadik és a negyedik formáns pedig megmutatja a beszélők azonosságát vagy különbözőségét.

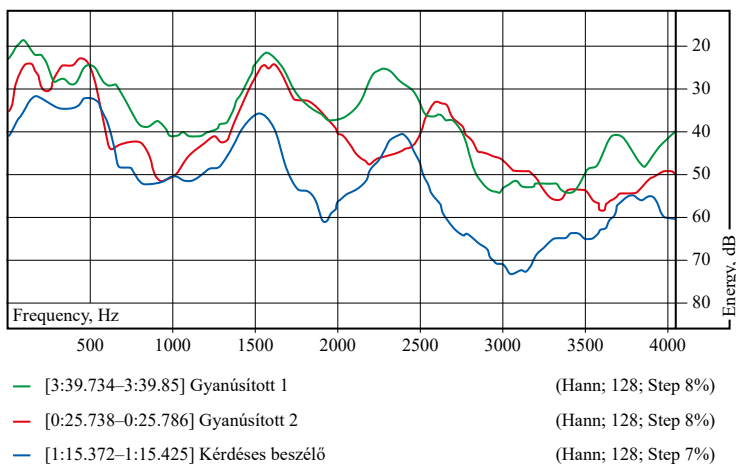


### 3. ábra

*A kérdéses beszélő (alsó panel); a „Gyanúsított 1” (középső panel) és a „Gyanúsított 2” (felső panel) által ejtett tervezett szó spektrogramjai*

A 4. ábrán bemutatott görbék – amelyek a 3. ábra kijelölt hangjainak felhasználásával készültek – elemzése alapján megállapítható, hogy az  $F_1$  500 Hz és az  $F_2$  1500 Hz környékén mindhárom beszélőnél hasonló értéket mutat. Az  $F_3$ -csúcspont a kérdéses beszélő és a „Gyanúsított 1” esetében hasonló helyen, kb. 2250-2300 Hz-nél található, míg a „Gyanúsított 2”-nél 2600 Hz körül van.

Az  $F_4$  a jelen esetben nem mérhető, viszont a példa enélkül is felhasználható bizonyítékként, és erősíti a korábbi hipotézisünket.



#### 4. ábra

*Az összehasonlított tervezett szóban lévő e hangok LTA-görbéi (pontozott vonallal jelölt görbe: a kérdéses beszélő; folytonos vonallal jelölt görbe: „Gyanúsított 1”; szaggatott vonallal jelölt görbe: „Gyanúsított 2”)*

A gyakorlatban a negyedik formáns gyakran nem mérhető, de ahogy a példa mutatja, ez nem gátolja minden esetben a véleményalkotást. Amikor azonban a negyedik formáns hiányzik, és az első három formáns megegyezik, az még lehet azonos vagy különböző beszélő is. Előfordulhat ugyanis olyan eset, amikor két beszélő között csak a negyedik formánsban van különbség. Ilyenkor a nagy mennyiségű hang-összehasonlítás és egyéb vizsgálatok segítségével lehet véleményt alkotni.

Ezekből a formáns-összehasonlító vizsgálatokból a lehető legtöbbet végzünk el. Lehetőség szerint annyi magánhangzót vizsgálunk, amennyi alapján egyértelműen megállapítható az azonosság vagy a különbözőség, illetve rövid felvételeknél az összes vokálist bevonhatjuk a vizsgálatba. A darabszám függ a percepciós elemzésektől és az automatikus azonosítás eredményétől is, ugyanis ha a biometrikus azonosító eredményei alapján nagyon nagy valószínűséggel azonos beszélőkről van szó, akkor nem feltétlenül szükséges 30-40 magánhangzó-realizációt formánsvizsgálattal összehasonlítani, hanem elég lehet 10-20 előfordulás is. Ebben a kérdésben a szakértőnek van némi szabadsága.

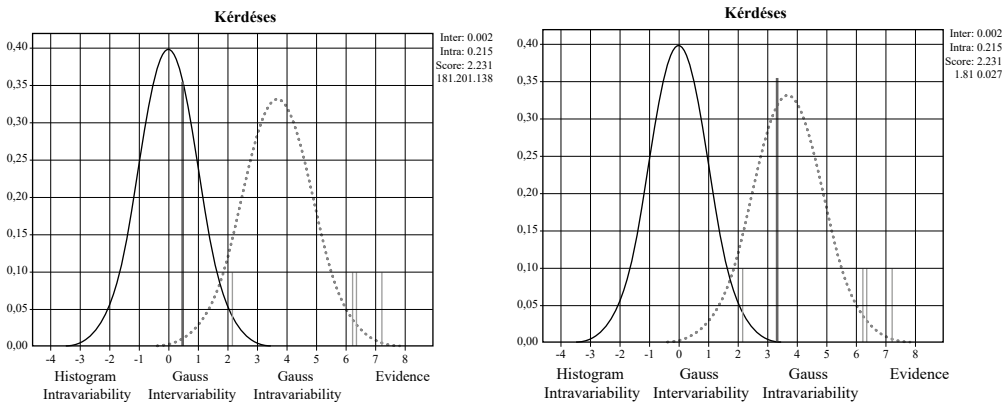
A percepciós és az akusztikai-fonetikai összehasonlítások alapján kialakítjuk a szakértői vélemény első részmegállapításait a valószínűségi skála szerint. Ezek a részmegállapítások – mivel minden mérési eredmény megerősítette azt a hipotézisünket, miszerint a kérdéses beszélő és a „Gyanúsított 1” azonos, illetve a „Gyanúsított 2” tőlük különböző – a +4-es és a -4-es, vagyis a) a kérdéses beszélő és a „Gyanúsított 1” nagyon nagy valószínűséggel azonos személy; b) a kérdéses beszélő és a „Gyanúsított 2” nagyon nagy valószínűséggel különböző személy. Ezt

követően elvégezzük az automatikus azonosítási vizsgálatokat, amelyek eredményei alapján szintén megteesszük majd a részmegállapításokat.

### 3.2. Biometrikus, azaz automatikus összehasonlító elemzések

A biometrikus beszélőazonosítás a térhódítását annak köszönheti, hogy már a hangtechnikai szakértői vélemények esetében is követelmény, hogy azok kváziobjektívek legyenek (TATÁR 2013b; DRYGAJLO et al. 2015). A módszer alapelve az, hogy minden ember biológiailag meghatározott, csak rá jellemző beszédszervekkel – tüdővel, hangajakkal, toldaléksóvel (garatüreg, szájüreg, orrüreg, pitvar) – rendelkezik. Ez minden körülmények között meghatározza az egyén beszédét, olyannyira, hogy akkor is felismerhető, ha különböző nyelven, más-más kommunikációs helyzetekben beszél, illetve különféle felvevőeszközzel rögzítik a hangját. Ezen alapelvek mentén fejlesztették ki az általunk használt Batvox automatikus beszélőazonosító rendszert (<https://www.nuance.com/omni-channel-customer-engagement/security/public-safety.html>) is.

A Batvox rendszer **valószínűségi arányszámmal** (*Likelihood Ratio, LR*) jellemzi az összehasonlított beszélők egymáshoz való hasonlóságát. Ennek kiszámítását úgy végzi el, hogy a kérdéses személyt összeveti a gyanúsítottal, illetve a populációs adatbázisában lévő más beszélőkkel. Ennek eredményeképpen megállapítja, hogy hányszor nagyobb a valószínűsége annak, hogy a kérdéses személy megegyezik a gyanúsítottal, mint annak, hogy nem. Ha ez a szám 1, akkor ugyanannyi a valószínűsége annak, hogy a két beszélő azonos, és annak, hogy különböző. Ha 1-nél nagyobb az érték, akkor nagyobb a valószínűsége annak, hogy azonos beszélőtől származnak a hanganyagok, ha 1-nél kisebb, akkor különböző a kérdéses beszélő és a gyanúsított. Az LR-eredmények mellett – a fent bemutatott 9 fokozatú skála alapján – számokkal jelöltük a hasonlóságot. Az 5. ábrán látható az automatikus azonosítás eredményeinek grafikonos ábrázolása.



5. ábra

A Batvox-szal készült hangminta-összehasonlítás eredménye: a bal oldali panelen a kérdéses beszélő és a „Gyanúsított 2”, a jobb oldali panelen a kérdéses beszélő és a „Gyanúsított 1”

A magasabb/meredekebb hisztogramgörbe szemlélteti a populációs adatbázis beszélőit, a laposabb hisztogram a kérdéses beszélőt, a görbét a csúcs közelében metsző vastag egyenes mutatja a gyanúsított eredményét. Ha ez az egyenes túlnyomórészt a magasabb/meredekebb hisztogramban helyezkedik el, akkor a populációs adatbázis beszélőihez hasonlít, ha inkább a laposabb görbén belül helyezkedik el, akkor a kérdéses beszélőhöz hasonlít a gyanúsított. A bal oldali panelről a kérdéses beszélő és a gyanúsított különbsége, míg a jobb oldali panelről azok hasonlósága olvasható le. Az összehasonlítás LR-értékeit az 1. táblázat tartalmazza. A két összetartozó minta esetében 204,43766-et kaptunk, ami azt jelenti, hogy ennyiszor nagyobb az esélye annak, hogy a két hanganyag beszélője azonos, mint annak, hogy nem. A kérdéses beszélő és a másik gyanúsított hangmintájának összevetésére 0,0272-es értéket kaptunk, azaz a két beszélő eltérése valószínű.

1. táblázat

*A kérdéses beszélő és a két gyanúsított biometrikus összehasonlítására kapott LR-értékek*

Beszélőpár	LR	Az LR érték szintje a valószínűségi skálán
<b>Kérdéses beszélő – „Gyanúsított 1”:</b>	204,43766	+1
<b>Kérdéses beszélő – „Gyanúsított 2”:</b>	0,0272	-2

Az automatikus azonosítás eredménye támogatja az eredeti hipotézisünket, azonban mindkét LR igen alacsony érték, amely gyengíti a korábbi vizsgálatok eredményeit. A 2. táblázat szemlélteti, hogy a részeredmények alapján milyen következtetések vonhatók le a beszélők azonosságára vonatkozóan. Nyelvész szakértői vizsgálatok nem minden esetben készülnek. Amennyiben tartalmaz nyelvészeti összehasonlítást a hangtechnikai szakértői vélemény, vagyis nem különálló nyelvész szakértői véleményről van szó, úgy annak eredménye a percepció és az akusztikai-fonetikai összehasonlítás részeredményeihez tartozik. Jelen példában nem mutattunk be nyelvészeti összehasonlítást, ezért a táblázatban „a nyelvészeti” zárójelben szerepel.

2. táblázat

*A beszélőazonosítási vizsgálatok összesítése*

	A percepció, (a nyelvészeti) és az akusztikai-fonetikai alapú azonosítás eredménye	Az automatikus azonosítás eredménye	Vélemény
<b>Kérdéses beszélő – „Gyanúsított 1”</b>	+4	+1	+2
<b>Kérdéses beszélő – „Gyanúsított 2”</b>	-4	-2	-3

A két részeredmény alapján tehát a következő vélemény fogalmazható meg:

1. A kérdéses beszélő és a „Gyanúsított 1” valószínűleg azonos személy (+2).
2. A kérdéses beszélő és a „Gyanúsított 2” nagy valószínűséggel különböző személy (-3).

### 3.3. A hangtechnikai szakértői azonosítás eredményének összegzése

A fenti beszélőazonosítási módszer összhangban van a nemzetközi módszertanokkal és a Magyar Igazságügyi Szakértői Kamara (2020) módszertani ajánlásával is. A szakértőnek ezekhez igazodva van némi mozgásteret abban a tekintetben, hogy milyen ábrákkal szemlélteti az eredményeket, milyen bizonyítékokkal támasztja alá, és hogyan fogalmazza meg a véleményét. Az itt bemutatott grafikonok és magyarázatok a szakértői bizonyítási eljárás egy lehetséges módját példázzák.

A bemutatott elemzés egy előzetes kutatásnak tekinthető, amely alapján hipotézisek is megfogalmazhatók a további vizsgálatokhoz. A fonetikai alapú összehasonlítás alapján a leg-erősebb megállapítást tehetjük, míg az automatikus azonosítás jóval alacsonyabb valószínűségi szintű vélemény megalkotását tette lehetővé. Ezt okozhatják a beszédsajátosságok változásai, a hangfelvételek minőségbeli különbségei is, de legnagyobb mértékben a módszertanunk sajátosságai a felelősek érte. Tudniillik a Szakértői Intézetben korábban elvégzett vizsgálataink alapján dolgoztuk ki azt a metódust, amely szerint a jelen vizsgálatban kapott 204-es LR-értéket – ami azt jelenti, hogy 204-szer nagyobb az esélye annak, hogy a két beszélő azonos, mint annak, hogy különböző – a valószínűségi skálán a +1-es, „a két beszédhang származhat azonos személytől” szintnek feleltettük meg, ami a „valószínűleg azonos” szintnél is alacsonyabb. Ebben a vizsgálatban a biometrikus összehasonlítás gyengítette az akusztikai-fonetikai részeredményeket, de olyan eset is előfordulhat, amikor az automatikus azonosítás eredménye jóval erősebb bizonyítékot nyújt, mint a többi vizsgálat.

Az eredmények azt mutatják, hogy a vizsgált kérdéses beszélő alaphangfrekvencia-értékeinek és formánsértékeinek 11 év után is megmaradtak az egyedi jellemzői, sikeresen beazonosítottuk az egyik gyanúsítottat, és egyértelműen kizártuk a másikat. Ez az eredmény a beszélőazonosítási gyakorlatunk szempontjából is fontos, mert alátámasztja a jelenlegi módszertanunk helyességét, és feltelezhetjük, hogy a nagy időkülönbséggel rögzített hanganyagok összehasonlítása és a beszélők azonosítása egy éles beszélőazonosítási feladat során is sikeresen elvégezhető lesz. A kutatást több beszélőre kiterjedően folytatjuk, hogy általános képet kapjunk a beszédsajátosságok változása és a beszélők azonosíthatósága kérdéskörében, különös tekintettel a biometrikus beszélőazonosítási metódusra, annak érdekében, hogy erősítse a vélemény valószínűségi szintjét.

A továbbiakban egy másik szakértői területről lesz szó, amely segítheti a beszélőazonosítást, hiszen egy ismeretlen beszélő egyedi stílusjegyei nemcsak a fonetikai, hanem más, például a nyelvhasználati sajátosságokban is megmutatkoznak. A hangzó szöveg nyelvészeti szempontú összehasonlításával további bizonyítékokkal támasztható alá a szakértői vélemény. Ezt az elemzést nyelvész szakértő végzi, aki hagyományosan írott szövegek összehasonlításával foglalkozik, azonban a hangzó beszéd nyelvi jellemzőinek elemzése is beleillik a nyelvész szakértői módszertanba.



## 4. Az igazságügyi nyelvészet

Igazságügyi nyelvészet kb. az 1960-as évek óta létezik; szerteágazó feladatkörét az *International Association of Forensic Linguists* (<http://www.iafl.org>) rendszerezte. A Nemzetbiztonsági Szakszolgálat Szakértői Intézet Nyelvész Szakértői Laboratóriumában alapvetően két módszerrel dolgozunk, melyekkel az angol nyelvű szakirodalomban a *forensic stylistic* és a *forensic authorship identification* (PERLMAN 2018) néven találkozhatunk. Ez a két tevékenység a **csoporthatározás** (nyelviprofil-alkotás) és a **szerzőazonosítás**; mindkettő a szerzőséget vizsgálja a kriminalisztikai szövegnyelvészet eszközeivel és módszereivel.

A tanulmány további részében ennek a két eljárásnak a rövid ismertetésére térünk rá, valamint egy, a szerzőazonosítás metodikájára épülő, de új szemléletű vizsgálatot mutatunk be.

### 4.1. A kriminalisztikai szövegnyelvészet

A kirendelő hatóság (nyomozó hatóság, ügyészség, bíróság) általában fenyegető, zsaroló, zaklató, rágalmozó, becsületsértő tartalmú szövegek vizsgálatára kér fel nyelvész szakértőt, kérdései pedig a következők:

- Ha egyetlen kérdéses írásműről van szó: *Mi állapítható meg a fogalmazójáról a nyelvhasználat alapján, pl. nem, életkor, iskolai végzettség, anyanyelv, terület (regionális nyelvhasználat) stb.?* A feladat az elkövetői kör szűkítése.
- Ha több kérdéses írásműről van szó: *Ugyanaz a személy fogalmazta-e a kérdéses írásműveket?* A feladat annak megállapítása, hogy egy elkövetőt kell-e keresni, vagy többet.
- Ha egy vagy több kérdéses írásmű és konkrét személy(ek)től származó összehasonlító szövegmintáról van szó: *A szövegminta alapján a gyanúsított/terhelt személy fogalmazta-e a kérdéses írásműve(ke)t?* A feladat a gyanúsított/terhelt személy beazonosítása.

**A kriminalisztikai nyelvészet vizsgálatának tárgya** lehet bármilyen magyar nyelvű szöveg, mely minimális terjedelemmel, gondolati tartalommal és struktúrával rendelkező írás- vagy beszédmű. Klasszikusan a nyelvész szakértő írott szövegekkel dolgozik, a hangzó szövegek elemzése a hangtechnikai szakértő feladata. Ugyanakkor a hangszakértők sok esetben kéri a beszélőprofil-alkotás vagy a beszélőazonosítás műveletéhez a nyelvész szakértő segítségét, ilyenkor a nyelvész szakértő a beszédleírásokat mint írott szövegeket elemzi. Azokban a speciális esetekben, amikor a hangszakértő nem tudja objektív okokból elvégezni a beszélőazonosítás műveletét (pl. a felvétel nagyon rossz minőségű, de még érthető szöveget tartalmaz, ezért akusztikai-fonetikai méréseket nem lehet rajta végezni, de nyelvészeti elemzésre alkalmas), a hatóság kirendelhet nyelvész szakértőt is, hogy ő végezze el a beszélőazonosítást (ilyen esetekben a nyelvész szakértő a fent írt módon jár el, és valószínűsítő véleményt ad).

A nyelvész szakértői vizsgálat szempontjából irreleváns, hogy az adott szöveg eredetileg szóbeli vagy írott, illetve az utóbbi esetben annak sincs jelentősége, hogy milyen kivitelezési eljárással készült (kézírás, írógép, számítógép, ragasztás, betűsablon, online környezetben létrehozott stb.) vagy milyen műfajban íródott (levél, e-mail, szerződés, végrendelet, poszt, komment, fórum- és chatszöveg, röpcédula, azonnali üzenetküldő szolgáltatáson keresztül továbbított üzenet, SMS stb.). A szövegvizsgálatokat minőségi (kvalitatív) és mennyiségi (kvantitatív) elemzéssel végezzük. A csoportbehatárolás és a szerzőazonosítás metodikájának kidolgozása Nagy Ferenc nevéhez fűződik, akinek 1980-ban jelent meg *Kriminalisztikai szövegnyelvészet* című könyve, mely a nyelvész szakértés eddigi egyetlen összefoglaló műve.

A csoportbehatárolás során és egyetlen kérdéses vagy egyetlen szövegminta rendelkezésre állása esetén a véleményalkotáskor a valószínűség szintjén maradunk, míg a szerzőazonosítás eredményeként akár kategorikus véleményt is megfogalmazhatunk (bár nagyon ritkán tesszük) egy 7 fokozatú skála alkalmazásával:

A fogalmazó–szerző azonos.

A fogalmazó–szerző nagy valószínűséggel azonos.

A fogalmazó–szerző valószínűleg azonos.

A fogalmazó–szerző azonossága nem zárható ki, de nem is bizonyítható.

A fogalmazó–szerző valószínűleg nem azonos.

A fogalmazó–szerző nagy valószínűséggel nem azonos.

## 4.2. A kriminalisztikai szövegnyelvészet háttere

Minden embernek van egy sajátos stílusa, idegen szóval **idiolektusa** (WARDHAUGH 1995: 112; ÜRMÖSNÉ SIMON 2019: 66), mely valamennyi nyelvi szinten vizsgálható, úgymint morfológia, szintaxis, szemantika, pragmatika, valamint helyesírás, szövegmegformálás, stilisztika (általános fogalmazási készség) stb. Ez az egyediség úgy működik, mint egy **nyelvi ujjlenyomat** (TOLNAINÉ KABÓK 2015: 135), vagyis a fogalmazásmód árulkodik a szerzőről. Az egyéni nyelvhasználat az a képesség, ahogyan az adott nyelv (jelen esetben anyanyelv) szabályrendszerét alkalmazzuk; tulajdonképpen egy adott nyelv eszköztárából való válogatás és ezek kombinálásának eredménye. Ez a válogatás kialakítja az idiolektust abban az értelemben, hogy a szövegalkotó nyilván csak az általa ismert nyelvi formák és szabályok közül képes válogatni a nyelv használata során, a válogatás és kombinálás korlátait a szerző társadalmi pozíciója, tanulmányai, olvasmányélményei, nyelvi tudatossága, életkora, neme, szakmája stb. szabja meg (SZILÁK 1980: 67). A **nyelvprofil-alkotás** során a fogalmazó ezen szociokulturális/szociodemográfiai tényezőit gyűjtjük össze, és próbálunk következtetéseket levonni a fogalmazó–szerző elsődleges csoportjegyére vonatkozóan, úgymint nem, életkor, iskolai végzettség.

A **szerzőazonosítás** módszerének alapja a kérdéses szerzőségű írásművek és a konkrét személyhez tartozó szövegminták összehasonlító vizsgálata során feltárt **általános és különös stílusjegyek** azonosságának, hasonlóságának és különbségének értékelése; az általános és különös

szövegsajátosságok feltárását, vagyis a (statisztikai) stíluselméletet perzepciós úton és konkordancia-programok használatával – pl. Laurence Anthony-szoftverek, <http://www.antlab.sci.waseda.ac.jp> – végezzük. A nyelvi sajátosságok különböző azonosítási értékkel rendelkeznek, és ez az érték fordított arányban áll a standard gyakoriság fokával. Például a köznyelvben ritkán használt szavak egy adott szerző írásaiban lehetnek egyéni szövegezési sajátosságok, feltéve, ha abban nagyobb gyakorisággal jelennek meg (kulcsszavak). Ha egy nyelvi tulajdonság egy adott személy írásaiban dominánsan jelentkezik, akkor az a tulajdonság általában jellemző a fogalmazóra, tehát általános sajátosság. Az általánoshoz viszonyítva kivételnek tekinthető tulajdonságok a különös sajátosságok, melyeknek szintén következetesen meg kell jelenniük az adott fogalmazó írásaiban (NAGY 1980: 60–61). A szerzőazonosítás művelete során ezen szövegsajátosságok hasonlóságának, illetve különbségének a mértéke alapján a nyelvész szakértő következtetéseket von le a fogalmazók azonosságára vagy különbözőségére vonatkozóan. CRYSTAL (1997: 92) szerint

Augustus de Morgan (1806–1871) brit matematikus és logikus, aki először ismerte fel a szerzőazonosítás lehetőségét a statisztikai stíluselmélet módszerével, 1851-ben ezt írta egy levélben: „Várhatóan arra az eredményre jutok, hogy egy személy két tárgyban készített írása inkább megegyezik egymással, mint két személy azonos tárgyban írt írása”.

Mi történik akkor, ha csak egyetlen kérdéses írásmű vagy egyetlen szöveg minta áll a rendelkezésünkre az összehasonlító vizsgálat elvégzésekor? Ebben az esetben nem tudhatjuk biztosan, hogy az általunk különös sajátosságnak értékelt nyelvi marker az adott szövegé (vagyis a fogalmazó által az adott műfajban, az adott szöveggörnyezetben, az adott címzethez írt, az adott körülmények között létrehozott stb. textusé), vagy a különös sajátosság a szöveget fogalmazó személyé (vagyis a fogalmazó esetében **individuais stílusjellemző**, minden körülmények között megjelenő nyelvhasználati jegy). A kérdéses írásmű elemzésekor szem előtt kell tartanunk a fogalmazó nyelvi variabilitásának lehetőségét is, aki ezt az adott, fenti tulajdonságokkal rendelkező szöveget éppen így fogalmazta meg, egyébként máshogy (is) tudja/szokta alkalmazni a nyelvet. Axióma, hogy nyelvhasználatunk igazodik a beszédhelyzethez (HÁMORI 2006: 161), egyéni választás eredménye, vagyis a nyelvhasználat szubjektív jelenség: függ az egyén akaratától, hangulatától, társadalmi pozíciójától, de függ a beszédhelyzettől, a műfajtól, a beszédpartnertől, a témától stb. is. Ezekkel a természetes változókkal számolni kell, és több írásmű beszerzésére kell törekedni az objektív véleményalkotáshoz.

A vizsgálatok eredményessége szempontjából szerencsés, ha az inkriminált szöveg spontán írásbeli megnyilatkozás. A spontán produktumok ugyanis sokkal jobban árulkodnak az alkotóról, hiszen ezekben tükröződik a legjobban, hogyan használja a nyelvet, és milyen a saját nyelvhez való viszonya. Ez nem jelenti azt, hogy ne kellene a szakértőnek tervezett írott szövegekkel is dolgoznia: ezek közös jellemzője, hogy átesnek a szerző részéről – akár többszörös – ellenőrző, javító folyamatokon, mely a spontaneitástól eltávolítja a szöveget. A legnagyobb kihívást a hivatalos, különösen a jogi nyelvezetű textusok jelentik, ugyanis ezekben olyan mértékű a formai és nyelvi kötöttség, hogy a fogalmazó–szerző egyéni stílusjellemzői csak ritkán jelennek meg; ezért formális stílusú szövegek elemzésekor a szakértői megállapítások valószínűsítőek.

### 4.3. Nyelvészeti mintavizsgálatok a Longitudinális korpuszon

A Longitudinális korpuszon korábban nem végeztünk kísérletet, így a most bemutatott vizsgálat az első, melyen nyelvész szakértőként lehetőségünk nyílt a spontánbeszéd-stílusok stílusjegyeinek megfigyelésére. Ebben a nyelvészeti szempontú kísérletben azt vizsgáltuk, hogy két, közel azonos nyelvi paraméterrel (csoporttulajdonsággal) rendelkező személy beszédsajátossága (beszélt nyelvhasználata) között milyen hasonlóságok és különbségek figyelhetők meg, vagyis el lehet-e különíteni a nyelvhasználati jegyek alapján a két beszélő személyt egymástól, ha a csoportjegyeik egyébként azonosak.

#### 4.3.1. A mintavizsgálat anyaga

A vizsgálat elvégzéshez a BEA adatbázisban, illetve a Longitudinális korpuszhoz tartozó adatbázisban a Nyelvtudományi Intézet kutatói által rögzített adatok alapján, az alábbi szempontoknak megfelelően kiválasztottunk két beszélőt:

- az adatközlő személyektől (A és B) közel azonos mennyiségű beszédleirat álljon rendelkezésre mind a 10 évvel ezelőtti (A1, B1), mind a mostani (A2, B2) felvételeket illetően (kiegénylített korpusz);
- a vizsgált beszélő megnyilatkozása legyen spontán, hogy egyedi jegyek vizsgálhatóak legyenek (interjúrészt); illetve
- a beszélők – az összehasonlító nyelvészeti vizsgálati szempontból releváns – csoportjegyei is legyenek nagyon hasonlóak (úgy mint nem, életkor, iskolai végzettség).

A fenti szempontok alapján létrehozott minikorpusz és a beszélők paramétereit a 3. táblázat foglalja össze.

3. táblázat

*A nyelvész szakértői elemzéshez választott két adatközlő jellemzői*

	<b>A szavak száma a beszédleiratban (db)</b>	<b>Nem</b>	<b>Életkor (év)</b>	<b>Végzettség</b>
A1	451	férfi	25	egyetem/főiskola
A2	823		35	
	<b>1274</b>			
B1	581	férfi	24	egyetem/főiskola
B2	604		34	
	<b>1185</b>			

A felvételekhez tartozó beszédleiratok készítésekor az elhangzottak szó szerinti leírására (tehát tartalomrögzítésre) törekedtünk, így a beszédleirat a beszédet kísérő – szupraszegmentális – jellemzőket (pl. szünet, hezitálás, elcsukló hang stb.) nem tartalmazza, ahogy a „köhint”, „nevet” stb. megjegyzéseket sem.

A vizsgálati korpusz méretét tekintve természetesen nem reprezentatív; a két beszélőtől egy 10 évvel ezelőtti és egy mostani szöveg vizsgálata (összesen kicsit több, mint négy A/4-es oldal terjedelemben) nem alkalmas arra, hogy a beszélők egyedi beszédsajátosságait (itt: beszélt idiolektusát) teljeskörűen feltérképezzük, és kategorikus kijelentéseket tegyünk a beszélők nyelvhasználati jegyeinek összehasonlításakor, ugyanakkor a korpusz alkalmas az általános beszéd- vagy nyelvhasználati jellemzők detektálására. A kiértékeléskor a beszédleiratok rövid terjedelme ellenére konzekvensen előfordul és/vagy kifejezetten a beszélt nyelvre jellemző tartalmas és funkciószavak megjelenítésére koncentráltunk.

Néhány nyelvi jellemzőnél a vizsgálati eredmények kiértékelésekor figyelembe kellene venni, hogy a kísérletben nem írott, hanem lejegyzett beszélt nyelv elemzése történik. Ilyen nyelvi jellemzők például a mondatokra vonatkozó adatok, mint a hosszúság, a mondatkezdő kötőszó előfordulása, a tagmondatok száma, a mondatok telítettsége stb. A beszéd lejegyzése ugyanis bizonyos szempontból szubjektív, például a mondathatárok megállapítása egyebek között a szünet hosszúsága alapján történik (a megnyilatkozásokra tagolásról összefoglalóan l. MARKÓ 2005: 30–34, a szünetészlelésről pedig KASSAI 1998). Nem függetleníthető a folyamat a beszédet lejegyző személytől. Mindazonáltal a jelen kísérletben a beszédleiratokat írott szöveggént értelmeztük, ahogy tesszük egyébként a nyelvész szakértői gyakorlatban is.

A beszédleiratokat – mint korábban írtuk – morfológiai, szintaktikai, szemantikai és pragmatikai jellemzőkkel írjuk körül, majd vetjük össze a feltárt nyelvi markereket. Ezen felül a statisztikailag kimutatható nyelvhasználati jegyek összehasonlítására is sor kerül.

A kísérlet során a szövegeket percepció útján és konkordanciaprogramok segítségével elemeztük (pl. AntConc 3.5.8 [Windows] 2019), a kvantitatív elemzések nagyrészt NAGY Ferenc (1972) *Kvantitatív nyelvészet* című könyve alapján készültek, az abban tárgyalt nyelvi markerek figyelembevételével, manuális és gépi feldolgozással – ez utóbbiban Vincze Veronika volt segítségünkre.

### 4.3.2. Nyelvhasználati eltérések

Az A és B beszélő beszédstílusa közötti statisztikai, morfológiai, szintaktikai jellemzőket érintő releváns eltéréseket a 4. táblázat foglalja össze. Az A és B adatközlők szókincsének gazdagságában nincs számottevő különbség (lemmák előfordulási aránya), akárcsak a szövegek telíttségében (lemmaszám / tagmondatok száma). Ugyanakkor azt láthatjuk, hogy a B beszélő mondatai jelentősen hosszabbak, mint az A-é, és a mondatok hosszúságának megfelelően a tagmondatok száma is magasabb B-nél, mint A-nál. Azt is láthatjuk a táblázatban, hogy A kétszer annyi egyszerű mondatot használ beszéde során, mint B. Ez némileg meglepő adat; az interjúkból kiderül, hogy az A beszélő bölcsész (magyar–esztétika szakos), a B pedig műszaki végzettségű (mérnök). A mondat szerkesztési eljárásokra vonatkozó elvárás épp a kapott eredmények ellenkezője lenne (ti. a bölcsészek hajlamosak a túlbonyolított mondatokban beszélésre, „írói véna” és egyéb sztereotípiák alapján).

Mivel mindkét adatközlő személynél a spontán vagy részben irányított interjú rész került feldolgozásra, az volt az előzetes feltételezésünk, hogy ezt az igei személyragok is tükrözik

(az interjúkészítő szabta feladat szerint az adatközlőnek magáról, a vele történekről kellett beszélnie). Ezzel szemben az A beszélő sokkal többször beszélt önmagáról, mint a B, aki viszont lényegesen többször használt T/1 igeragos szóalakot a beszédében, mint az A. Az utóbbi jelenség magyarázata az lehet, hogy B az interjúban egy projektről beszél, amelyet többen együtt valósítottak meg (pl. *fáradozunk, próbálkozunk, csináltuk*).

A modalitást illetően nincs jelentős különbség a két szövegcsoport között (a kijelentő, kérdő és felszólító mondatok előfordulási aránya közel azonos), és a két beszélő szófajhasználatára is szinte teljesen megegyezik (főnév, melléknév, ige, határozószó, kötőszó). Egyedül a névmások előfordulásában mutatkozik különbség az A beszélő javára; különösen a vonatkozó névmások előfordulási arányait tekintve (*aki-, ami-*).

Az A beszélő több múlt idejű igét használt (ez a tényező erősen témafüggő is, hiszen az interjú során az is feladat volt, hogy hasonlítsa össze a jelenlegi életét a 10 évvel ezelőttivel), viszont több feltételes módú igét is – mindkét igeidőben (pl. *adnám a fejem, sokáig tartana, beépültem volna, szerette volna*).

A szókészlet vizsgálata során azt tapasztaltuk, hogy az A beszélő sokkal gyakrabban használta a *hát* szócskát mind mondat belseji, mind mondatkezdő helyzetben (összesen 27 előfordulás a B beszélő 10-ével szemben), ugyanígy az *ugye* módosítószót (az A beszélő 7-szer, míg a B 1-szer), és a két szó összekapcsolódására is csak az A-nál láttunk példát: *hát ugye* (4-szer). *A na*, a *nos* és a *ja* indulatszavak is csak A-nál jelentek meg, de nem számottevő a használatuk (1-1 alkalom). Az A beszélő stílussajátosságai közé tartozik a szókészletet illetően pl. a *nem is tudom* szerkezet (3 előfordulás, míg a B-nél 1), melynek beékelődésszerű használatával 1-1 esetben talákoztunk mindkét adatközlőnél: *és nagyon, nem is tudom, informatívan tudta előadni; és akkor ott próbálgattam, nem is tudom, a merre tekerjek biciklis alkalmazást*. Egyedinek mondható még A beszédében a *kedv* szó kapcsolódásainak (*elveszik a kedvet, lenne megint kedvem, ehhez semmi kedvem*), valamint a *tekintetében/tekintetben* névutó (3-szor), illetve gyakori a *mondjuk* formula (10-szer, míg B-nél 2-szer), az *elég, eléggé* határozószók (8-szor, míg B-nél 2-szer) és az *illetve* kötőszó megjelenése (5-ször, míg B-nél 2-szer).

Az A beszélő stílusáról elmondható, hogy nagy mennyiségű negatív érzelmi töltetű szó jelenik meg benne (a bizonytalanság, undor, nyugtalanság érzelmekhez kapcsolódóan): ő az A1 felvétel készítésekor (10 éve) államvizsgázott, és a beszélgetés során a frissen szerzett keserű tapasztalatait osztotta meg az interjúkészítővel. A negatív élménnyel kapcsolatos kifejezések és szavak döntő többsége ennek köszönhető (*nem is volt a feleletem túlzottan, hogy is mondjam, túlzottan nagy teljesítmény; nem volt egy túlzottan pozitív élmény; valamint az elvárások egészen irreálisak voltak; baromira inkorrektnek gondoltam; csalódás volt; rettenetesen zavaró volt; nem szimpatizálnak ezzel a témával stb.*). Az A2 felvételnél szintén több a negativitás az A beszélő megnyilvánulásaiban, ekkor a téma a tanulmányok és a munkahely volt (*nem vagyok elég elhivatott; küzdelmes lenne; meg kéne küzdeni a figyelmükért; egyáltalán nem szeretek; nem vagyok én eléggé ott ebben a dologban stb.*). Több adatra lenne szükség annak megállapításához, hogy ez a negatív szemlélet témától függetlenül jelentkező alapbeállítottság-e a beszélőnél.

## 4. táblázat

## A nyelvhasználati markerek összehasonlítása

	A beszélő	B beszélő
A mondatok átlagos hosszúsága (szószám/mondat)	20,64	28,55
A tagmondatok aránya (tagmondat/mondat)	3,26	4,69
Az egyszerű mondatok aránya	0,29	0,12
A 2 tagmondatból álló mondatok aránya	0,15	0,12
A 3 tagmondatból álló mondatok aránya	0,16	0,18
A névmások aránya	0,11	0,08
A vonatkozó névmások aránya	0,014	0,004
A múlt idejű igealakok aránya	0,43	0,37
A feltételes módú igék aránya	0,08	0,01
Az E/1 igei személyrag aránya	0,30	0,24
A T/1 igei személyrag aránya	0,05	0,13

A B beszélő szóhasználatának jellegzetessége például a *tehát* kötőszó mondatkezdő helyzetben (9-szer, míg A-nál 4-szer), az *úgyhogy* kötőszó (9-szer; érdekes, hogy A-nál is előfordul a szó 3-szor, de csak mondatkezdő helyzetben, míg mondatkezdőként B-nél egyáltalán nem), a *viszonylag* határozószó (3-szor, míg A-nál egyszer sem), valamint a *végül is* szókapcsolat (5-ször, míg A-nál egyszer sem). Az általános értelmű *ember* szót inkább B preferálja (9-szer, míg A-nál 3-szor), de például a *dolog* szót mindkét beszélő használja (3-3 előfordulás). A szintén általánosító *mindig* névmás csak a B beszélőre jellemző, ugyanakkor mindkét adatközlő nyelvhasználatában megtalálható – és szinte azonos arányban – a *meg* kötőszó, és mondatkezdő helyzetben az *és* kapcsolóelem (7, ill. 8 előfordulás).

A B beszélő stílusáról általánosságban elmondható, hogy pozitívabb attitűdöt sugall A-énál, mind a 10 évvel ezelőtti, mind a mostani beszédleiratokban, legalábbis a beszélgetés során felmerült témákat illetően mindenképpen (*nagyon-nagyon megtetszett; érdemes lenne; kiderült, hogy felvettek; biciklizek egy jót; nagyon szerettem régen; nekünk ez könnyebben ment; szerencsénk van* stb.). Természetesen a B beszélőnél is vannak bizonytalanságra utaló kifejezések (*ez még nem egészen derült ki; valamennyit segít a dolgon; viszonylag jónak mondott iskola;* illetve határozatlan névmások, pl. *valamennyire, valahogy, valamilyen, valamelyik, valamennyit*), ám ezek nem uralkodók a stílusában.

A helyesírás vizsgálata a kísérletben elmaradt (beszédleiratokról lévén szó).

A spontán beszédre jellemző ismétléseket mindkét beszélő produkálja: A-nál sokkal gyakrabban találkozunk szóismétléssel, főleg a kötőszavakat és névmásokat érintően. Az A beszélő összesen 56-szor, a B beszélő 36-szor ismételt (pl. *hogy hogy, mint mint, és és, de de, ez ez, egy egy, nem nem, ilyen ilyen, akkor akkor, ami ami, ilyesmi ilyesmi, aki akivel, amik amik*). A hármas ismétlés csak A-nál fordult elő, összesen 7 alkalommal (pl. *hogy hogy hogy, ami ami*

*aminek*). Szintén az A beszélőnél találkozhatunk afféle névmási megerősítéssel (*a nyelvészeti tárgyakat azokat nagyon korrekten tanították; szerintük a képregények azok elveszik a kedvet az olvasástól*). Mindkét beszélőnél megjelenik az utólagosan raggal pontosítás (A: *elvállal valami olyan munkát, ami ami aminek mondjuk vajmi kevés köze van*; B: *jó lenne ... matematika, fizika, vagy ilyesmi ilyesmit tanulni; de azt hiszem aki akivel beszéltem*).

#### 4.3.3. A nyelvész szakértői elemzés összegzése és kitekintés

A bemutatott elemzés kis mintán készült, így a beszélők azonosságára vagy különbözőségére vonatkozóan nem tehetnénk egy valós kriminalisztikai vizsgálatban megállapítást, azt azonban látjuk, hogy kimutathatóak voltak a beszélők nyelvhasználata között olyan eltérések, amelyek mentén, további anyagok bevonásával, elindulhat a szakértő.

A Longitudinális korpusz nagyszerű kutatási és tapasztalatszerzési lehetőségeket rejt, nemcsak a hangtechnikai szakértő, hanem a nyelvész szakértő számára is. A bemutatott kísérlettel – mely beszédleiratokat dolgozott fel a beszélők nyelvhasználati sajátosságainak feltárása és összehasonlítása céljából – azt kívántuk bemutatni, hogy a kis méretű korpusz ellenére vannak olyan stílussajátosságok, melyek a beszélőket megkülönböztethetik egymástól a – nyelvészeti szempontból – szinte teljesen azonos csoportjegyeik mellett.

Érdeemes lenne a most bemutatott elemzést kiterjeszteni: több (hasonló paraméterű) beszélőt bevonni, vagy egy adott beszélőtől az adatbázisban rögzített egyéb spontán szövegrészt is elemezni (pl. véleménykifejtés), vagy egy adott beszélőnél a spontán blokkokat (interjúrészs vs. véleménykifejtés) nyelvészetileg összehasonlítani stb.; így sokkal pontosabb adatokhoz juthatunk, a korpusz kiterjesztésével pedig objektívebb elemzést tudunk megvalósítani egy adott beszélő csoportjegyeire és individuális stílusjellemzőire vonatkozóan. Érdekes lehet olyan aspektusból is vizsgálni a Longitudinális korpuszt, hogy egy spontán / részben irányított kommunikációs helyzetben (l. interjú) adott személy(ek) beszédstílusában / beszélt nyelvhasználatában bekövetkezik-e valamilyen változás 10 év alatt, és ha igen, milyen mértékű? Ebben az esetben a 10 évvel ezelőtti és a mostani interjúblokkok beszédleirata lenne az összehasonlítás alapja, mely összevetéssel a beszélt nyelvben bekövetkező egyéni (nem feltétlenül életkori) változások is detektálhatók lehetnének.

A jelenlegi kísérlet igazolta, hogy a beszédleiratok szövegnyelvészeti szempontú elemzése értékes információval szolgálhat a beszélőazonosítás és a beszélőprofil-készítés során a hangszakértők számára, illetve a beszélők nyelvhasználati jegyeinek összehasonlítására is alkalmas lehet, amikor a beszélőazonosítás műveletébe a nyelvész szakértő is bekapcsolódik.



## 5. Általános összegzés és kitekintés

A tanulmányban a kriminalisztikai beszélőazonosítás módszerét mutattuk be egy folyamatban lévő kutatás kis mintáján. Az NBSZ Szakértői Intézetében dolgozó hangszakértők és nyelvész szakértők gyakorlati eljárásait szemlélítettük. Természetesen a mindenkori vizsgálat további elemzések bevonását is szükségessé teheti (pl. eredetiségvizsgálat, zajszűrés), amennyiben az adott ügy hanganyagai vagy kérdésfeltevése ezt megköveteli.

A jelen tanulmányban hosszabb idő elteltével rögzített hanganyagok összevetésével mutattuk be a módszereinket. Az ilyen felvételek vizsgálatának a kereteit és a véleményalkotás lehetőségeit a folyamatban lévő kutatás fogja körüljárni.

### FOGALMAK

*beszélőazonosítás; spektrogram/hangszínekép; FFT; LTA; LPC; biometrikus; beszélőprofil; hangtechnikai szakértő; nyelvész szakértő; formáns; alapfrekvencia; Likelihood Ratio (LR) vagy valószínűségi arányszám; csoportbehatárolás (nyelviprofil-alkotás); szerzőazonosítás; idiolektus*

### IRODALOM

- ANTHONY, Laurence 2019. *AntConc (Version 3.5.8)* [Computer Software]. Waseda University, Tokyo, Japan. <http://www.antlab.sci.waseda.ac.jp/>
- CAO, Honglin – DELLWO, Volker 2019. The role of the first five formants in three vowels of mandarin for forensic voice analysis. In CALHOUN, Sasha – ESCUDERO, Paola – TABAIN, Marija – WARREN, Paul (eds): *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*. Australasian Speech Science and Technology Association Inc., Canberra. 617–621.
- CRYSTAL, David 1997. *A nyelv enciklopédiája*. Osiris Kiadó, Budapest.
- DRYGAJLO, Andrzej – JESSEN, Michael – GFROERER, Stefan – WAGNER, Isolde – VERMEULEN, Jos – NIEMI, Tuija 2015. *Methodological Guidelines for Best Practice in Forensic Semiautomatic and Automatic Speaker Recognition. Including Guidance on the Conduct of Proficiency Testing and Collaborative Exercises*. Verlag für Polizeiwissenschaft, Frankfurt am Main. [https://enfsi.eu/wp-content/uploads/2016/09/guidelines\\_fasr\\_and\\_fsasr\\_0.pdf](https://enfsi.eu/wp-content/uploads/2016/09/guidelines_fasr_and_fsasr_0.pdf) (A letöltés ideje: 2021. március 14.)
- GÓSY Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- GRÁCZI Tekla Etelka – KREPSZ Valéria – HUSZÁR Anna 2021. Egyéni variabilitás a beszédben. In MARKÓ Alexandra (szerk.): *Tanulmányok a beszédtudomány alkalmazásainak köréből*. ELTE Eötvös Kiadó, Budapest, 201–227.
- HÁMORI Ágnes 2006. A társalgási műfajokról. In TOLCSVAI NAGY Gábor (szerk.): *Szöveg és típus*. Tinta Könyvkiadó, Budapest. 157–181.

- JESSEN, Michael 2010. The forensic phonetician. Forensic speaker identification by experts. In COULTHARD, Malcolm – JOHNSON, Alison (eds): *The Routledge Handbook of Forensic Linguistics*. Routledge, New York. 378–394.
- KASSAI Iлона 1988. A szünet kérdésköre a szöveglejegyzésben. In KONTRA Miklós (szerk.): *Beszélt nyelvi tanulmányok*. Linguistica, Series A, Studia et Dissertationes 1. MTA Nyelvtudományi Intézet, Budapest. 22–43.
- LAZICZIUS Gyula 1979. *Fonetika*. Tankönyvkiadó, Budapest.
- MAGYAR IGAZSÁGÜGYI SZAKÉRTŐI KAMARA 2020. 3/2020. *Módszertani levél a beszélőazonosítási vizsgálatokról*. [https://igazsagugyiinformaciok.kormany.hu/download/d/13/b2000/MISZK%20Módszertani%20levél%203\\_2020.pdf](https://igazsagugyiinformaciok.kormany.hu/download/d/13/b2000/MISZK%20Módszertani%20levél%203_2020.pdf) (A letöltés ideje: 2021. március 4.)
- MARKÓ Alexandra 2005. *A spontán beszéd néhány szupraszegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hűmmögés vizsgálata*. PhD-értekezés. ELTE BTK, Budapest.
- MCMENAMIN, Gerald R. 2002. *Forensic Linguistics Advances in Forensic Stylistics*. CRC Press, Washington D.C.
- NAGY Ferenc 1972. *Kvantitatív nyelvészet*. Tankönyvkiadó, Budapest.
- NAGY Ferenc 1980. *Kriminalisztikai szövegnyelvészet*. Akadémiai Kiadó, Budapest.
- PERLMAN, Alan 2018. *What Is Forensic Stylistics?* <https://www.language-expert.net/category-stylistic-analysis/> (A letöltés ideje: 2021. január 8.)
- ROSE, Philip 2002. *Forensic Speaker Identification*. Taylor & Francis, London.
- SZILÁK Jolán 1980. Az írásszokások néhány formai jegyének háttéréről. *Belügyi Szemle* 16/8. 67–68.
- TATÁR Zoltán 2013a. Beszélőprofil-alkotás lehetőségei a kriminalisztikai fonetikában. *Alkalmazott Nyelvtudomány* 13/1–2. 121–130.
- TATÁR Zoltán 2013b. A spontán beszéd egyes jellemzői különböző felnőtt korcsoportokban. In VÁRADI Tamás (szerk.): *Alknyelvdok7 Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből*. MTA Nyelvtudományi Intézet, Budapest. 185–193. <http://www.nytud.hu/alknyelvdok13/proceedings13/proceedings13.pdf> (A letöltés ideje: 2021. január 15.)
- TATÁR Zoltán – OSZTERTÁG János – MORBER Szilárd Krisztián 2019. Technikai szakértői szakterület fejlődése a feladatok tükrében. *Belügyi Szemle* 67/1. 108–117.
- TOLNAINÉ KABÓK Zsuzsanna 2015. Interdiszciplináris kapcsolatok a rendészettudományok és az alkalmazott nyelvészet között – különös tekintettel a törvényszéki nyelvészetre. *Magyar Rendészet* 15/5. 131–145.
- ÜRMÖSNÉ SIMON Gabriella 2019. Miben segítik a nyelvi ujjnyomok a nyomozást? *Magyar Rendészet* 19/1. 65–75.
- VINCZE Veronika – ÜVEGES István – SZABÓ Martina Katalin – TAKÁCS Károly 2021. A magyar beszélt és írott nyelv különböző korpuszainak morfológiai és szófaji vizsgálata. In BEREND Gábor – GOSZTOLYA Gábor – VINCZE Veronika (szerk.): *XVII. Magyar Számítógépes Nyelvészeti Konferencia*. SZTE TTIK, Informatikai Intézet, Szeged. 177–182.
- WARDHAUGH, Ronald 1995. *Szociolingvisztika*. Osiris–Századvég, Budapest.

## **KÖSZÖNETNYILVÁNÍTÁS**

Köszönettel tartozunk Vincze Veronikának (MTA–SZTE Mesterséges Intelligencia Kutatócsoport Nyelvtechnológiai Csoport) a beszédfelvételek gépi elemzéséhez nyújtott segítségéért. A publikációban szereplő kutatást az Innovációs és Technológiai Minisztérium és a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal támogatta a Mesterséges Intelligencia Nemzeti Laboratórium projekt, valamint az FK128814 számú pályázat keretében.

# Fonetika és társalgáskutatás: a beszélőváltások kérdésköre

Horváth Viktória

NYELVTUDOMÁNYI KUTATÓKÖZPONT

Bóna Judit

ELTE EÖTVÖS LORÁND TUDOMÁNYEGYETEM

Dér Csilla Ilona

KÁROLI GÁSPÁR REFORMÁTUS EGYETEM

Gyarmathy Dorottya

NYELVTUDOMÁNYI KUTATÓKÖZPONT

Hámori Ágnes

NYELVTUDOMÁNYI KUTATÓKÖZPONT

Huszár Anna

NYELVTUDOMÁNYI KUTATÓKÖZPONT

Krepsz Valéria

NYELVTUDOMÁNYI KUTATÓKÖZPONT

Weidl Zsófia

ELTE EÖTVÖS LORÁND TUDOMÁNYEGYETEM

## 1. A társalgások vizsgálata és a fonetikai kutatások

A fonetikai kutatások középpontjában sokáig az egyéni beszélő beszédprodukciója állt, amit többnyire monologikus helyzetben, mesterséges környezetben – laborkörülmények között – rögzítettek. A múlt század utolsó évtizedeiben azonban a nyelvtudományban előtérbe került a mindennapi, spontán nyelvhasználat és a verbális társas interakciók (pl. beszélgetések, társalgások) vizsgálata, számos ezekkel foglalkozó (pl. szociolingvisztikai, nyelvészeti etnográfiai, diskurzuselemző stb.) kutatás indult meg, és ezek által egyre fontosabbá vált az empirikus – valós nyelvhasználatból származó – adatokból való kiindulás is. A magyar **társalgás** kifejezés

az angol *conversation* vagy *talk-in-interaction* megfelelőjeként jelölhet tágabb értelemben mindenféle spontán, két- vagy többszereplős beszélgetést, szűkebb értelemben pedig egy bizonyos (a magyar „csevegés”-hez hasonló) kötetlen beszélgetéstípust. A társalgáselemzésben általában a tágabb értelmezés használatos (SCHEGLOFF 1968 nyomán), mely mindenféle, valós társas nyelvhasználati helyzetben létrejövő verbális interakciót magában foglal, a rendőri kihallgatásoktól kezdve a baráti csevegéseken át a tanórai diskurzusokig vagy az orvosi vizsgálatokon elhangzó párbeszéddekig (HÁMORI 2006).

A társalgások elemzése kiemelt területté vált, és több ponton fonetikai elemzésekkel is összekapcsolódott. Több nagyobb irány vagy részterület is kialakult, attól függően, hogy mi volt az egyes kutatások fő érdeklődési fókuszja (például a szóátadás szerveződése, a társalgások szerkezete vagy egyes társalgásbeli jelenségek, mint például az udvariasság, a köszönések vagy a társas jelentések), milyen háttérelmélet segítségére támaszkodtak (pl. szociolingvisztika, pragmatika, pszichológia) és milyen módszertan, milyen kutatási módszerek jellemezték őket.

A kutatások nagy része az egyik legmeghatározóbb társalgáskutatási irányzathoz, a nyelvészeti konverzációelemzéshez, más néven társalgáselemzéshez kapcsolódik. A konverzációelemzés (*conversation analysis*, CA) mint elmélet és módszertani keret az 1970-es évek végén alakult ki a társas interakciók általános szerveződésének kutatása nyomán, elsősorban SACKS, SCHEGLOFF és JEFFERSON munkái alapján (SACKS–SCHEGLOFF–JEFFERSON 1974; áttekintésként I. ATKINSON–HERITAGE 1984; SIDNELL–STIVERS 2014; magyarul: IVÁNYI 2001; HÁMORI 2006; BORONKAI 2009; HÁMORI–HORVÁTH 2019). Ennek központjában a társalgások szerkezete, szerveződése áll, amelyet nem előre rögzített, fix szabályokkal leírható dolognak, hanem rugalmasan alakuló, a résztvevők által közösen létrehozott jelenségnek tekintenek; a konverzációelemzés ennek a rugalmas, dinamikus közös alakításnak a folyamatát vizsgálja elsősorban. A konverzációelemzéshez kapcsolódó fonetikai elemzések is ebbe a megközelítésbe épülnek be, és a beszélgetések prozódiai jelenségeit – például a hangsúly, dallam, artikulációs tempó jellemzőit – főleg a társalgások szerveződésében játszott szerepük, illetve a társas cselekvéssel való viszonyuk keretében vizsgálják. Ezekből később önállóan tekinthető részterület is kialakult, amit **társalgásfonetikának** vagy **interakcionális fonetikának** is szokás nevezni (*phonetics of conversation*, *phonetics of talk-in-interaction*), mint például a *Prosody in Conversation* (COUPER–KUHLEN–SELTING 1996), a *Prosody in Interaction* (COUPER–KUHLEN–FORD 2004), illetve *Sound Patterns in Interaction* (BARTH–WEINGARTEN–REBER–SELTING 2010) című kötetek tanulmányai is mutatják. Ezekben különböző nyelvek társalgásainak fonetikai elemzése is megtalálható, valamint különböző témák is megjelentek: ilyen például a vizuális jelzések szerepe a beszélőváltásban, a fonetikai jellemzők és a társas viszonyulás (*stance*) kifejezése a társalgásban vagy a prozódia és a társalgási műfajok kapcsolata.

Természetesen nemcsak konverzációelemzési keretben, hanem számos más megközelítésben is folynak fonetikai vizsgálatok az élőnyelvi beszélgetések kapcsán. Így lényeges irányt képviselnek azok a kutatások, amelyek a fonetika és a pragmatika (pontosabban valamelyik pragmatikai elmélet) összekapcsolásával dolgoznak, és többnyire egy-egy szűkebb diskurzus- vagy pragmatikai jelenséget vizsgálnak. A kutatások egy része a **diskurzusjelölőkkel** foglalkozik: ezek olyan nyelvi egységek, amelyek a társalgásban ismertető jegyei lehetnek

a **beszédforduló**knak, ezért hozzájárulhatnak a beszélőszegmentáláshoz és a diskurzus működéséhez. Átmeneti pontokat jelölnek a diskurzusban, jelzik a diskurzusstruktúrát, illetve segítenek abban, hogy a beszédpartnerek koherens mentális reprezentációt alakítsanak ki az egymással közölt információkról (FRASER 1999). BRYANT és FOX TREE (2005) szintén a pragmatika és a fonetika összekapcsolásával, rádiós showműsorok akusztikai elemzésével vizsgálták, hogy milyen prozódiai jelzéseknek van részük az ironia értelmezésében, illetve mi a nem nyelvi jellegű, kontextuális információk szerepe. De említhető a *Where Prosody Meets Pragmatics* című kötet is, amelynek tanulmányai a társalgások számos más jelenségét és azok fonetikai aspektusait vizsgálják, többféle pragmatikai vagy interakciós elmélet segítségével (relevanciaelmélet, udvariassági elméletek, társas pozicionálás), például az irreguláris zöngeminőség szerepét a beszélői attitűdök jelzésében (BARTH-WEINGARTEN–DEHÉ–WICHMANN 2009).

Egy további lényeges irányt képviselnek a társalgások pszicholingvisztikai és kultúra-összehasonlító (*cross-cultural*) keretű kutatásai is. Ezekben központi téma a beszélőváltásokban megfigyelhető univerzális (minden nyelvben egyformán jellemző) vonások és kulturális különbségek vizsgálata, például a váltások gyorsasága vagy az átfedő beszéd (más néven egyszerre beszéls, párhuzamos megnyilatkozás) terén (STIVERS et al. 2009), vagy az ezek mögött álló pszicholingvisztikai folyamatok (DE RUITER–MITTERER–ENFIELD 2006).

Más elméleti hátterű, de szintén igen lényeges irány a társalgások fonetikai elemzésében a szociolingvisztikai alapú kutatások köre, az ún. szociofonetika (pl. PRESTON–NIEDZELSKI 2010). Itt korábban (főleg LABOV és a múlt századi nyelvváltozat-szociolingvisztika nyomán) az egyes nyelvi változatoknak a makroszociológiai vagy biológiai tényezőkkel való összefüggései álltak a középpontban: tehát például az, hogy miként hat a lakóhely, a társadalmi csoport, a nem, az életkor vagy akár a beszélgetőpartner személye a beszéd vagy társalgás fonetikai jellemzőire. A kortárs szociolingvisztikában ugyanakkor már nem jellemző a nyelvi változatok ilyen merev felfogása, sem a beszélői csoportok ilyen homogén és stabil kategóriaként való elképzelése, ehelyett a dinamikus és konstruktivista szemlélet a meghatározó, így a jelen szociofonetikai kutatásai középpontjában már inkább a fonetikai jellemzőknek az identitás és a társas kategóriák (pl. kor, nem) aktív megalkotásában játszott szerepe áll.

Fontos említeni továbbá a korpusznyelvészetet is: a technikai fejlődés nyomán egyre több beszélt nyelvi, társalgási korpusz jött létre, így a társalgások fonetikai aspektusainak a vizsgálatára korpusznyelvészeti módszerekkel is egyre több lehetőség nyílik. A nagy adatbázison, nagyszámú társalgás elemzéséből kapott adatok számos érdekes kérdésre választ adnak (pl. a beszélőváltás gyorsasága, a prozódiai jelzések szerepe különböző helyzetekben vagy témák kapcsán); ezek kiemelten lényegesek a számítógépes, a pszicholingvisztikai és a szociofonetikai kutatásokban is. Egyre nagyobb szerepe van végül a társalgások fonetikai vizsgálataival dolgozó alkalmazott kutatásoknak is, ezek jelentőségére alább részletesebben is kitérünk.

Összefoglalóan tehát elmondható, hogy a társalgás, vagyis a *talk-in-interaction* fonetikai elemzése számos elméleti és módszertani megközelítéssel kapcsolódhat össze. Ezek nagy része ugyanakkor komoly elméleti és módszertani kihívásokkal is jár, mint például a kvalitatív és kvantitatív módszerek kérdései; a funkcionális és formális megközelítések különbözősége vagy

a természetes beszélgetések fonetikai elemzésének technikai nehézségei (pl. a felvétel minősége, a háttérzaj) (l. ODGEN 2012). Különösen érvényes ez a fonetikai és a társalgáselemzési vizsgálat összekapcsolására, ahol a fonetikában jellemző, hagyományosan formális nyelvszemlélet, valamint a társalgáselemzés interakcionális és funkcionális felfogása, illetve a kvantitatív és kvalitatív módszertani hagyományaik sok ponton ellentétben állnak egymással. Ezekre a kihívásokra az egyes, fent említett vagy itt nem érintett kutatási irányok és alterületek többféle választ adnak, ugyanakkor eltérő módszertanukkal és kutatási céljaikkal együtt is mindegyik fontos eredményeket kínál akár nyelvészeti, akár alkalmazott tudományi szempontból.

## 2. Empirikus eredmények a beszélőváltásokkal kapcsolatban

A társalgás az egyik legtermészetesebb és legáltalánosabban megjelenő/használt beszédhelyzet a mindennapokban. A konverzációelemzés mint átfogó elmélet és módszer kialakulását elsősorban SACKS, SCHEGLOFF és JEFFERSON (1974) kiinduló munkának számító tanulmányától szokás számítani, amelyben megfogalmazták ennek a megközelítésnek a fő fogalmait és gondolatait. A beszélői szerep váltakozásából adódóan jön létre az interakció alapegysége, a **forduló** (*turn*, más megközelítés szerint társalgási egység vagy [beszéd]lépés): ez olyan diskurzusegységet jelent, amely a beszélőnek a megszólalásától a szóátadásig tart. A fordulók megnyilatkozásait tehát ugyanazon beszélő hozza létre anélkül, hogy közben átadná a szót. A forduló lehet egy hosszabb megnyilatkozás, de akár egyetlen szó is. A fordulók határán szerveződnek a beszélőváltások a résztvevők közös létrehozásának eredményeként (SACKS–SCHEGLOFF–JEFFERSON 1974). A **beszélőváltás** (vagy szóátadás-átvétel, angolul *turn-taking*) tehát a fordulók váltakozása, amelynek eredményeképpen megváltozik a beszélő személye olyan módon, hogy a beszédhez való jog is átkerül az újonnan megszólalóhoz (SACKS–SCHEGLOFF–JEFFERSON 1974). A **beszédjog** a megnyilatkozás joga: kinél van a szó aktuálisan a társalgásban részt vevők közül. A társalgásnak az a résztvevője veheti át a szót az előző beszélőtől, aki a leggyorsabban reagál vagy válaszol az előző beszélő kérdésére, megszólítására (DÉR 2012).

Spontán társalgásokban sem a beszélők sorrendje, sem a fordulók hossza vagy azok szerkezete, sem pedig a beszélők közti megoszlásuk aránya nem előre rögzített, mindez rugalmasan alakul számos tényező (például a beszélői szerepek, a téma, a beszélők közötti viszony, a beszédhelyzet) hatására, nem véletlenszerűen, hanem egyfajta interakcionális rendezettség szerint (SACKS–SCHEGLOFF–JEFFERSON 1974; magyarul pl. IVÁNYI 2001; BORONKAI 2009; HÁMORI–HORVÁTH 2019). A fordulók között történő beszélőváltás megvalósulása szintén a résztvevők interakciójától függően, rugalmasan szerveződik; ehhez az alapot a **lehetséges szóátvételi pontok**, az ún. „beszélőváltásra alkalmas pontok” (angolul *transition-relevance place[s]*, TRP) adják (SACKS–SCHEGLOFF–JEFFERSON 1974). A beszélőváltásra alkalmas pontok a társalgás számos síkjának együttműködésében jönnek létre: a társalgópартnerek ezek azonosításához egyidejűleg

figyelembe veszik a szemantikai, pragmatikai, szintaktikai, prozódiai jelenségeket (pl. az intonáció típusa, a frázisvégi nyúlás, a hangerő változása, vö. pl. DUNCAN 1972; FORD–THOMPSON 1996; WENNERSTROM–SIEGER 2003), valamint a nonverbális jelzéseket (testbeszéd, mimika, szemmozgás). A szóátadás-átvétel több technikával történhet (vö. SACKS–SCHEGLOFF–JEFFERSON 1974; BORONKAI 2009): 1. **külválasztással**, amikor az aktuális beszélő kijelöli a következő megszólalót (pl. *És Kitti, te mit gondolsz erről?*), 2. **önkiválasztással**, amikor a következő beszélő önként szólal meg (pl. *Ilyen már velem is történt, amikor Horvátországba utaztunk tavaly...*). Amennyiben a társalgópartner(ek) nem veszi(k) át a szót, az eredeti beszélő folytathatja a saját megnyilatkozását.

Ahogy a bevezetőben már szó volt róla, a nemzetközi fonetikai kutatások körében az elmúlt évtizedekben egyre nagyobb szerepet kapott a társalgások akusztikai-fonetikai szempontú elemzése. A következőkben ezekből mutatunk be néhány empirikus elemzésre épülő eredményt.

A társalgások szünetezése az időzítés egyik kulcskérdése; a jelenség megközelítése azonban nem egységes. A beszéd folytonosságának megszakadása fonetikai tekintetben többféleképpen osztályozható és számos szempont szerint vizsgálható. Ilyenek például a forma (néma szünet, kitöltött szünet vagy más hezitációs formák, vö. MACLAY–OSGOOD 1959), a szintaxissal való kapcsolat (pl. határjelző vagy hezitációs szünet, vö. VAN DONZEL–VAN BEINUM 1996), az időtartam (GOLDMAN–EISLER 1972), esetleg a fordulószerveződésben betöltött szerep (*pause, gap, lapse* – néma szünet, [közös] hallgatás; SACKS–SCHEGLOFF–JEFFERSON 1974; HELDNER–EDLUNG 2010) vagy a beszélői szándék (LOCAL–KELLY 1986). Mindez befolyásolja a használt terminológiát is, amely tekintetében sokszor az egyes szakirodalmi források sem egységesek.

Az eredeti elmélet szerint (SACKS–SCHEGLOFF–JEFFERSON 1974) a társalgások során fontos szervezőelv a megfelelő időzítés betartása, azaz, hogy a beszélőváltások olyan módon menjenek végbe, hogy se hosszú hallgatás, se átfedő beszéd ne jelenjen meg (*no gap, no overlap*) – az első kommunikációs modellek ezt az időzítési sajátosságot hangsúlyozták. Mindemellett a beszélőváltások időzítése társas jelentéshordozó funkcióval is bírhat (például a hétköznapi társalgásokban többletinformációt hordozhat, ha a partner csak nagyon hosszú idő után vagy épp azonnal válaszol a feltett kérdésre), illetve ezek szerveződésében a kulturális különbségeket is figyelembe kell venni, vagyis azt, hogy mi számít normatív (megfelelő, elvárt) viselkedésnek, például mi számít szokásos vagy hosszú hallgatásnak a különböző kultúrákban (SIDNELL 2001). A két beszélő megnyilatkozása között eltelt idő az ún. **FTO-érték** (*Floor Transfer Offset*, vö. DE RUITER–MITTERER–ENFIELD 2006), amely háromféle lehet: amennyiben a pozitív tartományban van, akkor a szóátvétel néma szünettel valósult meg, negatív érték esetén átfedő beszédként realizálódott, míg az azonnali beszédjogátadás esetén az FTO értéke 0 ms. Feladatorientált dialógusok alapján mérték különböző nyelvekben a szünetek átlagos időtartamát, amely az angol nyelvben 380 ms, a német nyelvben 363 ms, a japánban pedig 389 ms volt (vö. WEILHAMMER–RABOLD 2003). HELDNER és EDLUNG (2010) három különböző telefonos korpusz (Dutch Dialog Corpus, English Map Task, Swedish Map Tasks) eredményeinek összevetése alapján szintén nagyon hasonló értékeket adatolt. A beszélőváltások körülbelül 60%-a néma szünettel, 40%-uk pedig átfedő beszéddel valósult meg. A néma szünetek több mint 70%-a rövidebb volt, mint 500 ms, és a váltások közel fele rövidebb volt, mint 200 ms, tehát azt mondhatjuk, hogy a beszélőváltások



nagyon gyorsan mentek végbe. Az átfedő beszéddel történő szóátvételek esetén a szóátvételi idő átlagértéke 610 ms-os (medián 470 ms) volt a holland Spoken Dutch Corpus alapján, amelyben nem telefonos, hanem személyes beszélgetéseket rögzítettek.

Mind a szünetek időtartama, mind a megjelenésük igen jelentős különbségeket mutat a nyelvek, a beszéd típusok vagy a beszédtempó szerint, valamint az egyes beszélők között (RUDER-JENSEN 1972). A kutatások arra is rámutattak, hogy azok a szünetek, amelyek szintaktikai tagolásra szolgálnak, gyakran egybeesnek az intonációs frázisokat jellemző prozódiai határként funkcionáló szünetekkel. És nem csupán az előfordulásuk, de az időtartamuk is összefüggésben áll a határjelző funkcióval: a szünetek időtartama korrelációt mutatott a prozódiai határjelzés erősségével, valamint az adott frázis hosszával is (ZELLNER 1998).

WENNERSTROM és SIEGEL (2003) vizsgálatának eredményei azt valószínűsítik, hogy a váltások megjelenésének valószínűsége a szünet első 500 ms-án belül csökken, majd hosszabb szünet időtartamára (kb. 1500 ms) nő. A beszélőváltás lehetősége rövid és hosszú szünetek esetén a legnagyobb, míg a közepes időtartamú szüneteknél a legkisebb.

Az átfedő beszéd fontosságát számos tanulmány hangsúlyozta (vö. TEN BOSCH–OOSTDIJK–BOVES et al. 2005 és SHRIBERG–STOLCZKE–BARON 2001). Az utóbbi a több résztvevős párbeszédet vizsgálva az összes szó 17%-ában és a szünet nélküli intonációs egységek 54%-ában talált átfedéseket.

A nagy arányú átfedő beszéd és a hasonló beszélőváltási értékek alapján ma már egyértelműen kijelenthetjük, hogy a beszélőváltások szerveződésében nem kötelező elv a hosszú néma szünetek vagy az átfedő beszéd mint kommunikációs jelenségek elkerülése, hanem ezek – valamint a váltások más kivitelezési módjainak jellegzetességei is – mint jelentéssel bíró nyelvi jellemzők valósulhatnak meg a társalgásokban.

Amennyiben a szóátvétel kérdés-válasz helyzetben túl hosszú időtartamú (kulturánként eltérő az az érték, amelyet az adott nyelv beszélői is hosszúnak tekintenek), a nempreferált válasz adásával, az egyet nem értéssel, a megbízhatatlansággal, az őszintétlenséggel, a beszélő önbizalomhiányával, a beszédprodukciónak nehézségével vagy a következő beszélő kiválasztásának nehézségével magyarázható. A fonetikai és pragmatikai kutatások közös eredménye szerint tehát minél később érkezik, annál kevésbé tűnik őszintének a válasz, és annál kevésbé tűnik hitelesnek a beszélő (ROBERTS–FRANCIS–MORGAN 2006).

LEVINSON és TORREIRA (2015) a Switchboard Corpus-on végzett elemzésükben, összesen közel 350 két résztvevős társalgás alapján azt találták, hogy ezek közel 20%-át tette ki a fordulók közötti szünettartás, és csupán közel 4%-át átfedő beszéd. A beszélőváltások közel kétharmada néma szünettel, míg egyharmaduk átfedő beszéddel realizálódott. A néma szünetek időtartama a legtöbb váltás esetében 500 ms-nál rövidebb volt, vagyis ez is alátámasztja a váltások sokszor nagyon gyors megvalósulását.

Egy másik vizsgálatban (WEINGARTOVÁ–CHURAŇOVÁ–ŠTURM 2014) négy, átlagosan körülbelül 20 perces cseh párbeszéd vizsgálata alapján azt találták, hogy a fordulók körülbelül kétharmada valósult meg néma szünettel és egyharmada átfedő beszéddel. A különböző típusú váltások időtartama nem különbözött egymástól jelentősen (a néma szünetes váltások mediánja

333 ms, az átfedő beszédé 353 ms volt). A típusok szerint nem, míg a beszélők között szignifikáns különbség volt kimutatható, ami a jelenségek erősen individuális voltára utal.

Más vizsgálatok kizárólag a kérdés-válasz szekvenciák esetében mérhető FTO-értékek alakulására fókuszáltak. STIVERS és kollégái (2009) 10 különböző nyelv adatait vetették össze, és nagyon gyors váltási időtartamokat mértek a kérdés-válasz szekvenciák esetében (7–468 ms, módusz: 0–200 ms), noha ezek az értékek különböző mértékben eltértek egymástól az egyes nyelvek esetében. Mindezek mellett elmondható, hogy a fordulók vizsgálata sokszor nehezen összevethető, ugyanis számos tényező eltérést okozhat a módszertanban (LEVINSON–TORREIRA 2015). Ismert, hogy a kiegészítendő kérdésekre adott válaszok esetében átlagosan nagyobbak az FTO-értékek, mint az eldöntendő kérdések esetében (STIVERS et al. 2009), ugyanis azok valószínűsíthetően több kognitív előkészületet igényelnek, illetve összefüggést feltételeznek a légzés és a válaszadás minősége között is (TORREIRA–BÖGELS–LEVINSON 2015; CASILLAS–BOBB–CLARCK 2014). Emellett szerepet játszhat az FTO-értékek alakulásában a beszélőalkalmazkodás jelensége (TEN BOSCH–OOSTDIJK–DE RUITER 2004; TEN BOSCH–OOSTDIJK–BOVES 2005), valamint a beszédhelyzet is (a telefonon keresztül rögzített társalgások esetében nagyobb FTO-értékeket adtak, mint a személyes társalgások esetében (LEVINSON 1983; TEN BOSCH–OOSTDIJK–BOVES 2005)). A beszélőváltással kapcsolatos normákat és FTO-értékeket továbbá a kontextus és a szűkebb műfaj (pl. osztálytermi, orvosi interakció) is befolyásolhatja. Számos elemzés mutat rá egyfajta konklúzióként arra, hogy a társalgások elemzése nem szűkíthető le csupán egy-egy terület eszköztárára a problémakör interdiszciplináris jellegéből adódóan, az átfogó leíráshoz szükséges például a pszicholingvisztika, a pragmatika vagy akár a szociolingvisztika módszertanának bevonása is.

### 3. Fonetika és társalgáskutatás a magyar beszédtudományban

A fonetikai szempontú társalgáskutatás a magyar beszédtudományban viszonylag rövid múltra tekint vissza. A kutatók saját gyűjtésű beszédfelvételei és a nagyobb beszédkorpuszok, mint például a BEA beszélt nyelvi adatbázis (vö. GÓSY és mtsai 2012) és a multimodális anyagokat tartalmazó HuComTech (HUNYADI és mtsai 2018) tették lehetővé az egyes társalgási jelenségek korpuszalapú, fonetikai fókuszú vizsgálatát.

A kutatások foglalkoztak például a háttérsatorna-jelzések megvalósulásával (pl. MARKÓ–GÓSY–NEUBERGER 2014, GYARMATHY és mtsai 2020), a diskurzusjelölőkkel (1. 4. fejezet), a társalgások szünetezésével (pl. MARKÓ 2005a; BÓNA 2013; HORVÁTH et al. 2020), a nonverbális hangjelenségekkel a társalgásokban (pl. NEUBERGER 2012; GOSZTOLYA–BEKE–NEUBERGER 2019), gépi beszélődetektálással, azaz annak automatikus jelölésével, hogy mikor melyik résztvevő beszél a társalgáson belül (BEKE 2014), illetve a prozódiai jellemzők automatikus jelölésének és kinyerésének kidolgozásával (SZEKRÉNYES 2019).

A továbbiakban néhány olyan kutatást ismertetünk részletesebben, a teljesség igénye nélkül, amelyek fonetikai aspektusból (is) foglalkoztak a társalgások szerveződésével, a beszélőváltás jelenségével.

MARKÓ Alexandra kutatásai (2005a, b; 2015) saját gyűjtésű felvételen alapultak, és úttörőnek számítanak a magyarországi fonetikai szempontú társalgáskutatásban. Disszertációjában (2005b), amely négy fővel rögzített, kétórányi spontán beszéden alapul, egyrészt összevetette monologikus és dialogikus közlések szupraszegmentális jellemzőit, másrészt részletesen foglalkozott a hűmmögés mint különleges társalgási elem kísérletes vizsgálatával. Saját korpuszán vizsgálta továbbá a társalgások szerveződésén belül a szüneteket és hallgatásokat, az átfedő beszédet és a beszélőváltási pontok szintaktikai jellemzőit (2006). Megállapításai szerint a társalgási egységeken belüli szünetek szignifikánsan rövidebbek, mint a társalgási egységek határán lévő hallgatások; a vizsgált társalgásban a résztvevők számára a közös hallgatás elfogadható időtartama maximum 5 másodperc volt, az ennél hosszabbra nyúlt hallgatást kényelmetlennek érezték. Eredményei szerint továbbá az átfedő beszéd létrejöttének oka nagyrészt az egyszerre megszólalás, vagyis amikor egy időben történik önkiválasztás két beszélő részéről is; illetve amikor az aktuális beszélő fordulójának vége előtt szólal meg egy másik beszélő. A beszélőváltási pontok elemzése azt mutatta, hogy ezeknek kb. 10%-a szintaktikailag befejezetlen egységet követett, és ezek majdnem háromnegyed része kötőszóval végződött (*úgyhogy, vagy, de, akkor stb.*). Ez összecseng azokkal a nemzetközi kutatásokkal (pl. FORD–THOMPSON 1996), amelyek feltárták, hogy a fordulóváltásnak nem kötelező feltétele a szintaktikai lezártság, hanem abban szemantikai és pragmatikai tényezők is szerepet játszanak. A szóátvételben, a beszédjog-megtartásban és a szóátadásban részt vevő elemeket és a fonetikai síkkal való összefüggéseiket későbbi kutatások részletesebben is vizsgálták (pl. az *úgyhogy* kapcsán GYARMATHY 2015; DÉR 2012; HÁMORI–DÉR 2020). Mint ezek is mutatják, a diskurzusjelölők fontos szerepet játszanak a társalgások szerveződésében, a beszélőváltásokban (l. 4. fejezet).

A beszélőváltásokban a megszólalások jellemzőit elemezték a BEA adatbázisból származó hat darab háromfős társalgásban (MARKÓ–GÓSY 2015). Vizsgálták a szóátadások szerveződését, ezek típusait; illetve összefüggéseit a hallgatások/szünetek időtartamával és bizonyos lexikai elemek megjelenésével. A beszélőváltások többsége önkiválasztással valósult meg. A hallgatások időtartamának és a beszélőváltás típusának összefüggés-elemzése azt igazolta, hogy tendenciaszerűen rövidebb volt a hallgatási periódus önkiválasztás esetén, vagyis amikor a beszélő önmaga akart megszólalni. A kiválasztásnál tendenciaszerűen hosszabb hallgatások arra utalnak, hogy a megszólított résztvevő nem feltétlenül akart megszólalni, így a kiválasztást követően kell megterveznie, mit és milyen formában mondjon.

A társalgások és a társalgási jelenségek időbeli változását kevés kutatás vizsgálta a hazai szakirodalomban. Az egyik kutatásban a BEA adatbázis négy felvétele alapján elemezték, hogy miként alakulnak bizonyos beszédparaméterek és társalgási jelenségek a társalgások első és utolsó négy percében, illetve attól függően, hogy a beszédpartnerek ismerték-e egymást korábban, vagy nem (GRÁCZI–BATA 2010). Az eredmények azt mutatták, hogy az ismerősökkel készült felvételen gyakoribb volt az átfedő beszéddel történő szóátvétel; illetve a felvételek

utolsó négy percében gyakoribbak voltak az ilyen jellegű beszélőváltások, mint az első négy percben.

Egy másik kutatásban elemezték, hogy az idő előrehaladtával a társalgásokban az egyes szóátvételtípusok milyen gyakorisággal fordulnak elő (MARKÓ–GÓSY 2015). A vizsgált hat BEA-társalgás első fordulói, amelyeket az adatközlők hoztak létre, jellemzően külválasztások voltak (az interjúkészítő kérdéseire adott válaszok), majd egyre gyakoribbá váltak az önkiválasztásos megszólalások. Ezt a tendenciát a szerzők az egymásra hangolódással, a kezdeti feszültség oldódásával magyarázták. Egy további pilotkutatás a BEA adatbázis egy 13 perces társalgását elemezte részletesen, a társalgáselemzési szempontokat ötvözve a fonetikaiakkal (HÁMORI–HORVÁTH 2019). Ennek során a szerzők nemcsak a fordulók és beszélőváltások (típusok, időzítés, dallammenet) jellemzőit elemezték, hanem vizsgálták a társalgás globális szerkezetét és a beszélőváltások azzal való összefüggéseit is, valamint bevonták a beszélőváltást magyarázó legújabb pszicholingvisztikai elméletet, a „bejólás” vagy „jóló feldolgozás”-elméletet is (DE RUITER–MITTERER–ENFIELD 2006; LEVINSON–TORREIRA 2015; LEVINSON 2016). Ez az elmélet a beszélőváltások gyors és gördülékeny lebonyolításának magyarázatához kapcsolódik, és azt állítja, hogy a beszélőváltás nem az aktuális beszélő szóátadó jelzéseinek hatására történik, hanem annak előkészítése már a társalgás korábbi szakaszában megkezdődik a beszédpartnerrel a forduló várható lezárásának megjósolásával és a beszédprodukciónak tervezésének korábbi megindulásával. Ez a korai predikció teremt meg annak a lehetőségét, hogy a beszélőváltások gördülékenyen és sokszor igen gyorsan, 400 ms alatti szünettel vagy rövid átfedő beszéddel végbemehessenek. Az elemzés eredményei azt is mutatták, hogy a beszélőváltás többségében önkiválasztással történt; a váltások nagyjából fele 400 ms-osnál rövidebb néma szünetet követően ment végbe (akár 0 ms-os szünettel, vagyis azonnali szóátvétellel). Ez is arra utal, hogy a résztvevők figyelik egymás nonverbális és verbális jelzéseit a társalgás során a gördülékeny beszélőváltás lebonyolításának érdekében. Az eredmények tehát nemcsak összhangban vannak a korábbi magyar mérések adataival (pl. MARKÓ–GÓSY 2015), hanem a jóló feldolgozás működését is alátámasztják. A dallamra vonatkozó eredmények szerint a beszélőváltást nem feltétlenül előzi meg intonációs lezárás (ereszkedő vagy eső dallam); majdnem ugyanolyan gyakran lebegő dallam adathozható a váltás előtt. Ez megfelel azon korábbi szakirodalmi megállapításoknak, miszerint a spontán beszédben az emelkedő dallam közlészáró helyzetben is gyakori (MARKÓ 2009). A társalgás szerkezetének elemzése azt mutatta, hogy a globális struktúra a bevezetés és a lezárás között további nagyobb tematikai és funkcionális egységekre tagolódott; váltakoztak a narratív jellegű és a dialogikus szakaszok. Ezekben a nagyobb egységekben eltérő fordulójellemzők, eltérő szóátvételi arányok és mintázatok valósultak meg. A beszélgetés bevezetésében és zárásában, valamint egyes dialogikus szakaszaiban rövidebb fordulók és gyakoribb beszélőváltás volt jellemző, más, narratív szakaszokban a fordulók jóval hosszabbak, a beszélőváltások pedig ritkábbak voltak.

Multimodális felvételek alapján elemezték kétfős társalgásokat (ABUCZKI 2011), amelyekben a beszélőváltásokat megelőző szünetek leggyakrabban 400 ms és 700 ms között realizálódtak (a minimumérték 270 ms, a maximum pedig 3777 ms volt – ez utóbbit a partnerek már túl hosszúnak ítélték, a videofelvételek alapján kissé „lefagyott a társalgás”). A szóátadás egybeesett

azzal, hogy az aktuális beszélő szünetet tartott, közben tekintetét a másik fél felé irányította, szemkontaktust felvéve, gyakran kérdéssel kísérve. Testhelyzetváltások leggyakrabban diskurzusszegmensek határán, illetve beszélőváltásnál fordultak elő (ABUCZKI 2011). (A **diskurzusszegmens** változó hosszúságú, megnyilatkozások- [sorozatá]ból álló, a tagmondattól a bekezdésig tartó terjedelmű írott vagy beszélt nyelvi szövegegység.)

A társalgások szerveződését és a társalgási jelenségek, valamint a fonetikai paraméterek megvalósulását befolyásolhatja az is, hogy a résztvevők egyetértenek vagy nem értenek egyet az adott témában. A kérdéskört fonetikai szempontból elemezték saját korpuszon (WEIDL 2019). Középsiskolai diákoknak kellett megvitatni adott témákat, ellentétes álláspontot képviselve az egyik kondícióban, a másikban pedig azonos véleményt. A diákok az egyet nem értések során az egyet-értésekben megjelenő paraméterektől eltérő alaphérfrekvenciával és artikulációs tempóval beszéltek. Az alaphérfrekvencia alacsonyabb, az artikulációs tempó pedig a diákok többségénél magasabb volt egyet nem értés során; de az egyéni különbségek szerepe meghatározó. Az egyetértés / egyet nem értés nonverbális jelzéseit, gesztusait vizsgálták a HuComTech korpusz multimodális felvételein. A megállapítások között szerepelt például, hogy a verbális egyet nem értés előtt és után is megfigyelhető volt a fejrázás gesztusa, illetve az egyetértést pislogás kísérte (HUNYADI 2019).

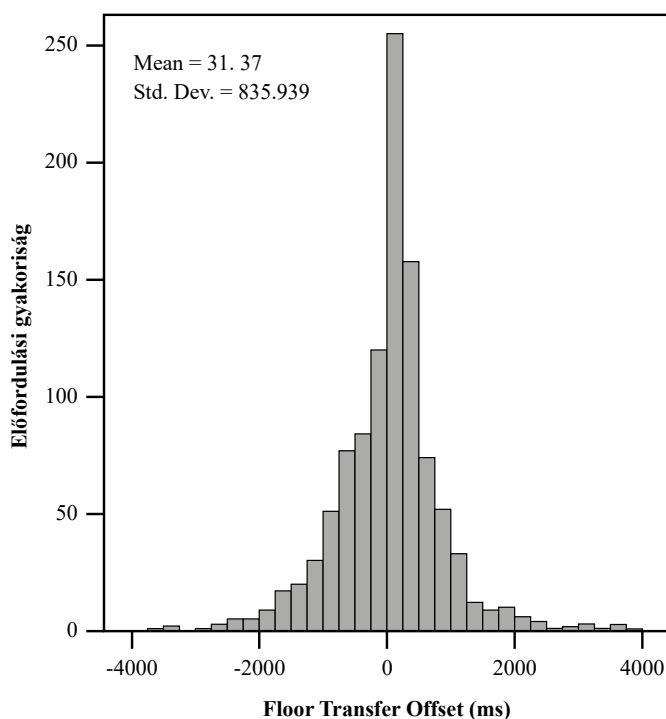
A *Beszédegységek fonetikai jellemzőinek összefüggései* című projektben a jelen tanulmány szerzői a BEA beszélt nyelvi adatbázis (GÓSY és mtsai 2012) társalgási alkorpuszának anyagát (HORVÁTH és mtsai 2019) használják a társalgáselemzések során. A BEA beszélt nyelvi adatbázis hét részből épül fel, melyben a társalgás az ötödik feladat. A teljes BEA-felvételek kb. 45 percesek, amelyekből a társalgás 15–20 percet tesz ki. A társalgások három résztvevősek, az interjúkészítőn és az adatközlőn kívül egy társalgási partner (az interjúkészítő munkatársa) vesz részt még a beszélgetésben. A témákat minden esetben az interjúkészítő vezeti be (ahogy a BEA többi feladatában is), a másik két résztvevőnek azonban nincs felkészülési ideje, a társalgás spontán alakul. A társalgások során a mindennapokhoz kapcsolódó, aktuális társadalmi, közéleti témák megvitatására kerül sor (pl. éppen aktuális ünnepek megünneplése; nyaralási élmények; mobiltelefon kisgyermekeknek; új KRESZ és a biciklisek; dohányzás; állattartási szokások; tehetségkutató műsorok). A témák kiválasztásánál az interjúkészítő figyelembe veszi az adatközlő életkorát, érdeklődési területeit, hobbját, melyekről a korábbi spontánbeszéd-részekben tájékozódott. A felvételek annotációja és az adatok kinyerése a Praat programban történik (BOERSMA–WEENINK 2018).

A globális (a teljes társalgást leíró) elemzéseken kívül dinamikus elemzések is készülnek a pályázat keretén belül. Ezen elemzések célja, hogy a társalgásokon belüli változásokat írják le, tendenciákat vázolja fel a különböző fonetikai és társalgásszervező jelenségek alakulásáról mind a teljes társalgásokban az idő előrehaladtával, mind pedig a váltásokhoz viszonyított pozíciójuk függvényében. A dinamikus elemzések többféle megközelítésből vizsgálják a jelenségek alakulását. Az egyik elemzés során a társalgásokat automatikusan több egyenlő időtartamú részre osztjuk, így az egyes részekben számolt gyakoriság összevethető. A másik elemzés során a társalgásban betöltött pozíciót korreláljuk az egyes jelenségek időzítési jellemzőivel. A beszélőváltásoktól való távolság esetén szintén dolgozunk csoportokba sorolással automatikus klaszteranalízis alapján is

(közel vagy távol helyezkedik-e el az adott jelenség a váltástól), illetve folytonos változókkal is, amikor a váltásoktól való távolságértékekkel számolunk.

Egy kutatássorozatban a BEA beszélt nyelvi adatbázis társalgási alkorpuszának 20 felvételén elemeztük a beszélőváltások megvalósulását (a korpusz bemutatását bővebben l. HORVÁTH et al. 2019). A felvételek társalgásai három résztvevősek, amelyekben az interjúkészítő és a társalgó partner személye mindig állandó, a beszélők személye változik, ők 20 és 40 év közöttiek. A korábbi kutatások többségében statikus, leíró szempontrendszer alapján vizsgálták a beszélőváltásokat, ezért a jelen vizsgálatsorozat legfőbb célkitűzése annak megválaszolása volt, hogy hogyan változnak bizonyos paraméterek a társalgásban dinamikusan, az idő előrehaladtával, illetve a belső szerkezet, pl. a beszélőváltásokhoz képest elfoglalt pozíció függvényében (HUSZÁR és mtsai 2021).

A 20 felvételen összesen több mint ezer esetben történt beszélőváltás, percenként átlagosan háromszor. Az 1. ábra hisztogramja az FTO-értékek eloszlását mutatja: a váltások leggyakrabban  $-250$  és  $+500$  ms közötti tartományban realizálódtak. Ez az eloszlás összhangban van mind az angol Switchboard korpuszban adatolt értékekkel, mind a holland nyelvre irányuló kutatások adataival (STIVERS et al. 2009; LEVINSON–TORREIRA 2015), ami alapján a beszélőváltások ezen paramétere nyelvtől függetlennek, univerzálisnak tűnik.

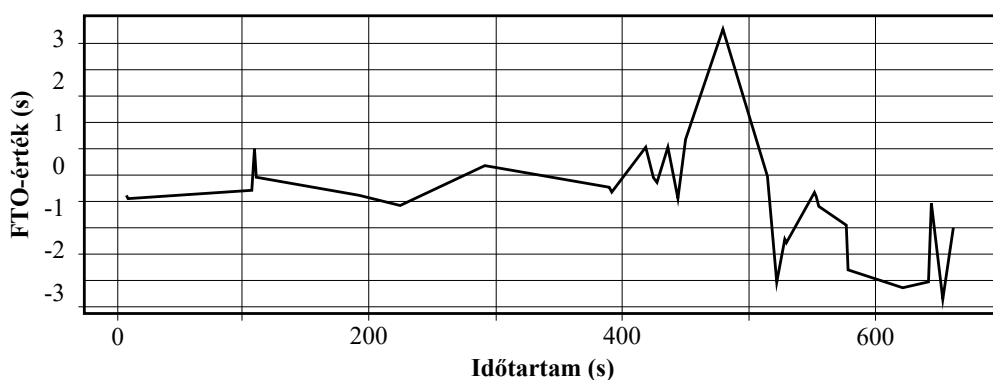


1. ábra

Az FTO-értékek gyakorisági eloszlása (HUSZÁR et al. 2021)

A beszélőváltások kicsit több mint fele pozitív FTO-értékkel, vagyis néma szünettel valósult meg, valamivel kevesebb, mint a fele negatív FTO-értékkel, vagyis átfedő beszéddel. Csupán az esetek 5%-a realizálódott 0 ms-os FTO-val, tehát amint az egyik beszélő abbahagyta a közlését, a másik beszélő azonnal elkezdte a megnyilatkozását. Az eredmények szerint a beszélőváltás helye a társalgásban meghatározó. Minél jobban közeledünk a társalgás végéhez, annál valószínűbb, hogy a beszélőváltás negatív FTO-értékkel, vagyis átfedő beszéddel valósul meg.

A 2. ábrán az FTO-értékek időbeli alakulása látható: az idő előrehaladtával egyre többször megy negatív tartományba az FTO-érték. Az átfedő beszéddel megvalósuló beszélőváltások gyakoriságának növekedése azzal magyarázható, hogy az idő előrehaladtával már nemcsak világosak a társalgásbeli szerepviszonyok, hanem a partnerek egyre jobban megismerik egymás kommunikációs stratégiáját, továbbá megfigyelhető egyfajta közeledés, szinkronizáció is.



2. ábra

Az FTO értékeinek alakulása a társalgásokban (HUSZÁR et al. 2021)

## 4. Diskurzusjelölők a beszélőváltásokban

A társalgások fonetikai szempontú elemzéseiben fontos területet alkot a diskurzusjelölők kutatása; ugyanez fordítva is igaz, a diskurzusjelölő-kutatásban nagy jelentősége van a fonetikai aspektusok vizsgálatának, valamint a korpuszalapú elemzéseknek is, amelyek egy része társalgási korpuszokkal dolgozik. A diskurzusjelölők társalgásbeli használata számos szempontból vizsgálható, ebben a fejezetben egy szűkebb jelenségkört, a társalgás szerkezeti szerveződésében és a beszélőváltásban játszott szerepüket tekintjük át.

A diskurzusjelölők beszélőváltásokban betöltött szerepéről már az e pragmatikai elemeket tárgyaló legelső monográfiában is olvashatunk. SCHIFFRIN (1987) szerint az általa felvett öt beszédszint közül csak a dialógusokra jellemző a cserestruktúra (*exchange structure*), amelynek az

esetében a beszédfordulók a beszédegységek, valamint az e szinthez tartozó diskurzusjelölők (például *oh, well, and, so, but*) felelnek a szerepváltásokért, vagyis a szóátadásért és -átvételért.

SCHIFFRIN (1987) megkülönböztet egymástól kooperatív és nem kooperatív fordulókezdést. Kooperatív a fordulókezdés akkor, ha az előző beszélő be tudta fejezni a mondandóját. Nem kooperatív fordulókezdés esetén az előző beszélő még nem adta volna át a másik félnek a szót, de az átvette tőle. Ezeknek megfelelően a szerző kétféle beszédforduló-kezdő pozíciót vesz fel a diskurzusjelölők esetében: lehetséges beszélőváltásnál (*transition space entry*) vagy nem lehetséges váltásnál fordulókezdő pozíciót (*non-transition space entry*) (vö. *transition-relevant place*, TRP, lehetséges beszélőváltási pont: SACKS–SCHEGLOFF–JEFFERSON 1974; IVÁNYI 2001; HÁMORI 2006).

A diskurzusjelölők tehát a szóátvételi kísérleteknél és a sikeres szóátvételeknél is meghatározó elemek. Ezt a későbbi témába vágó kutatások a diskurzuskoherecia pragmatikai szintjéhez, azon belül pedig a szekvenciális struktúrához kötve diskurzusstrukturáló szerepnek nevezik (GONZÁLEZ 2004). Ebbe a szerepbe beletartozik GONZÁLEZ szerint a diskurzusszegmensek határainak jelzése, illetve a diskurzushálózat fenntartása, ezeken belül pedig a szegmenskezdés és -zárás, az idézés jelzése, valamint a cselekményfejlődés felvezetése.

Azok az annotációs sémák, amelyeket természetes nyelvi szövegekben előforduló diskurzusjelölők számítógépes nyelvészeti feldolgozásához készítettek, műfajfüggetlen, általános diskurzusjelölő-funkciókat tartalmaznak, amelyek ötvözik a korábbi funkciórendszerezéseket. CRIBLE (2018) a diskurzusjelölők szekvenciális funkciói között említi a tagolást, a nyitó és a záró szegmenshatár jelzését, a témaváltást és a korábbi témához való visszacsatolást, az idézés, a hozzáadás, valamint a felsorolás jelölését (l. még FURKÓ 2019; DÉR 2020).

Változó, hogy a téma (topik) kezelését – témaindítás, témaváltás, korábbi témához való visszatérés – a szekvenciális vagy a retorikai funkciókörhöz rendelik-e a szerzők: mint láttuk, CRIBLE (2018) – magasabb szintű diskurzusszervező funkcióként értelmezve azt, elsősorban REDEKER (1990) alapján – az előbbihez, GONZÁLEZ (2004) pedig – az illokúciós szándékokat kifejező diskurzusjelölőket sorolva ide – az utóbbihoz.

A szekvenciális szerepek nem mindegyike releváns a beszélőváltásokban, elsősorban a szegmensnyitás és -zárás kapcsolható hozzájuk, de a következőkben bemutatásra kerülő diskurzusjelölői jellemző miatt az empirikus elemzésekben mégsem csak ezeket célszerű figyelembe venni. Az egyes szekvenciális funkciók ugyanis egymással és más, nem szekvenciális (ideációs, retorikai, interperszonális) funkciókkal együttesen jelenhetnek meg a diskurzusjelölők aktuális használata során, e vonásukat nevezik multifunkcionalitásnak (MÜLLER 2005; DÉR 2017). Fokozottan igaz ez a társalgásokban megjelenő változataikra, hiszen ezekben a – szűken értett – szekvenciális szerepek igen meghatározóak, de nem kizárólagosak az előbbi műfajra: idézést például más beszélt és természetesen írott nyelvi regiszterekben is megfigyelhetünk, így az ezeket felvezető diskurzusjelölőket is (vö. DÉR–CSONTOS–FURKÓ 2020).

Az erős multifunkcionalitás különösen igaz a spontán beszélt nyelv leggyakoribb diskurzusjelölői közé tartozó *hát-ra* (DÉR–MARKÓ 2007), amely mind a négy fent említett funkciókörben (ideációs, retorikai, szekvenciális, interperszonális) feladatot láthat el, jellemzően két-háromban



egyszerre (SCHIRM 2011, 2017; DÉR 2017, 2020). A következő társalgásbeli példában (1) a *hát* egyszerre nyitja a szegmenst, az interjúkészítő (IK) fordulóját, vezet be új altémát, illetve kapcsolja azt egy korábbi szegmensben elhangzottakhoz (a rúzs kifejezett megemlítésével, amit a társalgópártner [TP] a *de azt már ne* közlésrészhez is köti), valamint jelez következtetést (az evéstől lejön a rúzs). (A példákban a B a beszélő/adatközlő jele, a zárójelben lévő közlés pedig átfedő beszéd.):

- (1) TP: *akkor elmész a kozmetikushoz és a fodrászhoz az körülbelül három óra de az arra minimum számolj akkor hogyha mondjuk tízre mentél már egy óra van és akkor utána már ugye hogy hazamész akkor valamit mondjuk eszel de azt már ne*  
 B: *tudok enni?*  
 TP: *valamennyit muszáj különben elájulsz*  
 IK: ***hát*** *úgy hogy a rúzsod le ne jöjjön*

Az erős multifunkcionalitás lehetővé teszi a diskurzusjelölők és társulásaik rugalmas alkalmazását: a *hát* esetében a többféle (szekvenciális, retorikai és interszónális) funkció együttes használata és aluldefiniáltsága elősegíti a megnyilatkozók közös jelentéskonstruálását, egyezkedését (HÁMORI–DÉR 2020). A következő példában (2) az interjúkészítő és a beszélő a hűsvéti locsolkodást illetően ugyan eltérő állásponton vannak, de a *hát* használata lehetővé teszi, hogy a beszélgető felek ennek révén fejezzenek ki új szekvenciát, új altémát, megengedést, a saját véleményét és annak finomítását, valamint konklúziót:

- (2) IK: *és a locsolkodást is szereted*  
 TP: *hát most már azt csinálja a papám meg most már a kornélékat is rávettük hogy azzal a kölnivel ami a sajátunk*  
 IK: *(aha)*  
 TP: *és akkor azzal na melyikkel fújhatom és kó így kezébe nyomom és akkor az nem annyira durva mintha ilyen mindenféle lenne*  
 IK: ***hát*** *igen én is azt utáltam amúgy mindig benne hogy olyan izolat iszonyatos pacsul szaga lett az embernek*  
 B: ***(hát de az tök jó)*** *az a lényeg az a lényeg hogy*  
 IK: *(de ezt csak) azért mondd mert mert fiú vagy*  
 B: ***(hát igen)***  
 TP: *(igen)*  
 IK: *rajtad hogy nem volt még az hogy egész nap így szagolgattad magad hogy már átöltöztél hatszor és hogy hol (lehet még ma fúj)*  
 B: *(igen)*  
 TP: *(és megörülsz tet)*  
 B: *(vagy hát utána) utána jöhetnek azok a fiúk akik a zuhannyal locsolnak és akkor kész*  
 TP: ***(ühüm)***

IK: *á hát én mondjuk inkább annak vagyok a híve hogy maximum szódával egy picit*

B: *(és akkor megtisztított)*

IK: *de tehát*

B: *na de hát azt nem te döntöd el*

TP: *(hát nem de ato füg)*

IK: *(hát igen) fiúk jönnek és locsolnak hát*

Fontos látnunk, hogy a szegmensnyitó és -záró szerep magában foglalja a beszédjog (azé, akinél a szó van) irányítását: ezek a jelölőhasználatok mind az addigi beszélő külválasztását, vagyis amikor ő jelöli ki az új beszélőt, mind az új megszólaló önkiválasztását (IVÁNYI 2001) példázhatják, de főként az utóbbit; a (spontán) társalgásokban ugyanis nem előre adottak ezek a szerepek, de gyakoribbak az önkiválasztások (BORONKAI 2013; SCHIRM 2017).

Úgy tűnik, nincs olyan diskurzusjelölő, amelyiknek a kizárólagos szerepe a szóátvétel vagy a szóátadás lenne, tehát csakis a szegmensnyitást vagy -zárást jeleznék; ilyen értelemben tehát minden társalgásbeli fordulókezdő, illetve -záró, vagyis beszédirányító szerepű diskurzusjelölő multifunkcionális.

A társalgásokon belül használatos diskurzusjelölők két nagy típusa különíthető el aszerint, hogy az elem egyes funkcióiban szokott-e társulni más diskurzusjelölőkkel, vagy sem, például a következtetést kifejező *így igen* (pl. *és így, de így*), míg figyelemirányító, hangsúlytalan(abb) változatában nem (3), akkor sem, ha olyan jelölők állnak mellette, amelyekkel más szerepében össze tud kapcsolódni, lásd a (3)-beli két *és akkor/így* előfordulást a önálló *így*-használat mellett:

(3) B: *(és az olasz kempingek najon) jók tehát mindegyiknek van (úszómedencéje amit ingyen)*

IK: *(mhm mhm)*

B: *lehet (használni)*

IK: *(mhm)*

B: *korrekt fürdőszobaa meg ilyenn le vannak ilyen fogkefeautomaták meg fogkrémautomaták és akkor így mindig meg egyéb (automaták)*

IK: *teát*

B: *és akkor így mindent lehet ööö csinálni amit csak (szeretnél)*

IK: *(aha)*

B: *nem kell feltétlenül így így (szenvedned)*

A *hát* ebből a szempontból is igen figyelemreméltó, mert beszélt nyelvi előfordulásai több mint felében más diskurzusjelölőkkel társul, leginkább a (szegmens)kezdő *de* és *és* jelölőkkel, követve azokat (DÉR 2020), például:

(4) IK: *én azt elkezdtem olvasni em emlékszem és így a nemtom hányadik oldal után nem tudtam hogy most hol vagyok meg most mi meg most egyáltalán ki az akinek a tehát*

*ugye egyes szám harmadik személyben íródott ki az egyes szám harmadik személy mer hol a Steve vagy ki az isten volt ugye Steve-nek hívják*

A: *aha asszem igen*

IK: *Ste Stephen vagy valami ilyesmi*

B: *aha Stephen*

IK: *hol a izé am Mollynak a férje*

B: *a másik nem tudom de igen*

IK: *most ezt remélem [...] tanár úr nem hallgatja meg*

TP: *elég kicsi az esély hogy beszédelemzésnek szeretne nekiállni*

IK: *de de hát ez így katasztrófa volt viszont úgy ahogy össze volt foglalva tartalmilag mer ugye hát mihez folyamodik végül a vizsga előtt a magyar szakos úgy halál érdekes volt*

A diskurzusjelölők tehát szóátvételkor nemcsak akkor tölthetnek be szerepet, ha abszolút fordulókezdő helyzetben vannak, hanem akkor is, ha relatív fordulókezdő pozíciójuk. Az utóbbiak közé tartozik, ha – mint fentebb láttuk – egy fordulókezdő társulásban nem az első, hanem a többedik helyet foglalják el. Ettől elkülönítendő, amikor fordulón belül – de nem annak a legvégén – jelentkeznek szegmenskezdő szerepben, akár önálló elemként, akár egy diskurzusjelölő-társulás első vagy többedik tagjaként, hiszen ott már nem hordozhatnak szegmenshatár-jelző szerepet. A kezdő pozíció ugyanis szegmenskezdést jelent, szegmens egy propozíciót tartalmazó (tulajdonképpen tagmondatnyi) nyelvi egység lehet (vö. FRASER 1999).

Innen nézve a beszédforduló speciális szegmens(ek sorozata): mivel fordulók csak dialogikus műfajokban léteznek, érdemes megnézni, hogy megfigyelhető-e valamilyen sajátossága a fordulókezdő (szóló vagy valamelyik társulásbeli pozíciót felvevő) diskurzusjelölőknek az ezektől eltérő jelölökhöz képest. Az eddigi vizsgálatok (DÉR 2020) alapján úgy tűnik, hogy kevéssé: ugyanazok a jelölők és társulásaik jelentkeznek a fordulók, mint más szegmensek elején, de e témában további célzott vizsgálatokra van szükség.

A magyar nyelvű társalgások beszélőváltásaikor megjelenő diskurzusjelölők az eddigi kutatások szerint (DÉR 2012; DÉR–MARKÓ 2007, 2010; HÁMORI–DÉR 2020) mind gyakoriságukat, mind variabilitásukat tekintve szóátvételkor meghatározóbbak. Egy, a BEA 30 társalgását, csaknem 8 órányi beszélgetést tartalmazó alkorpuszra vonatkozó vizsgálat eredményei szerint (DÉR 2012) 1102 diskurzusjelölő jelent meg a fordulók elején (önállóan vagy diskurzusjelölő-társulás valamelyik tagjaként), 189-féle típust reprezentálva, de ezek kétharmada csak egyszer szerepelt. Lényeges, hogy a jelöltípusok 81%-a társult diskurzusjelölő volt (pl. *hát igen, és akkor, de hát*), ugyanakkor az összes előfordulás több mint felét a már fentebb is említett *hát* (az összes fordulókezdő diskurzusjelölő 26,59%-a), *de* (13,34%) és *és* (12,79%) adta ki (a további leggyakoribb egyszavas jelölő a *meg*, a *mondjuk* és a *tehát* volt). Az összes beszédforduló (2540 darab) csaknem 44%-ában jelentkezett szóátvételkor diskurzusjelölő elem (percenként 2,34 darab). Szóátadásnál összesen 44 darab (14-féle) diskurzusjelölőt (pl. *úgyhogy, tehát, vagy*) talált az elemzés, ez a beszédfordulók összmennyiségéhez viszonyítva 1,73%. Nemek közötti eltérés nem mutatkozott egyik esetben sem a diskurzusjelölők gyakoriságában.

A beszélőváltással kapcsolatos más kutatásokat is bevonva azonban azt látni, hogy a diskurzusjelölőknek a fordulók végén is fontos szerepe lehet. Noha ennek a pontos jellege nehezen ragadható meg, erre a szerepre számos korábbi kutatás utal. Így például már DUNCAN (1972) a beszélőváltás legelső modelljében is ír a beszélőváltást segítő lexikális elemekről vagy „sztereotip kifejezésekről”; későbbi korpuszalapú kutatások is alátámasztották, hogy egyes diskurzusjelölők kiemelten gyakoriak nemcsak fordulókezdő, hanem fordulóvégi helyzetben is. Szerepük lehet a fordulóváltási szándék jelzése, szóátvételre buzdítás (vö. DÉR 2012), ugyanakkor más kutatások ezeknek a beszédfolytatási szándékot jelölő, vagyis beszédjogmegtartó funkcióját mutatták be (HORVÁTH 2009). A fordulóváltás lexikális elemeit vizsgálva a BEA 10 magyar társalgásán végzett elemzés azt találta, hogy a beszélgetések összesen 443 fordulójából mintegy 90 forduló végződött valamilyen, a fordulóváltás szempontjából relevánsnak vagy tipikusnak tekinthető („sztereotip”) lexikális elemmel, és ezek között domináltak a diskurzusjelölő jellegűek. A 90 elemből 75 db, azaz 85% diskurzusjelölő volt (HÁMORI 2020a; HÁMORI–DÉR 2020). Ezek között az *úgyhogy* és a *tehát/dehát* voltak a leggyakoribbak, ami megegyezik a korábbi kutatások adataival (HORVÁTH 2009; DÉR 2012; GYARMATHY 2015). Ugyanakkor ezekhez a diskurzusjelölőkhöz, bár voltak bizonyos prozódiai jellegzetességeik, sem szintaktikai szerkezet, sem fonetikai jellemzőik alapján nem volt hozzárendelhető egyértelmű fordulózáró vagy szóátadó funkció, és sok esetben hasonlóságot mutattak a fordulók belsejében található diskurzusjelölőkkel is. A kvantitatív és kvalitatív elemzések alapján úgy tűnik, hogy ezek szerepe elsősorban nem a szóátadó jelzések jelenségek köre kapcsán, hanem inkább a fordulókonstrukció-típusok, fordulóváltási stratégiák és a lokális, közös jelentés-létrehozás felől írható le (HÁMORI 2020a; HÁMORI–DÉR 2020).

Ezeknek, illetve a beszélőváltás mechanizmusának a további vizsgálatához több fontos irányba is kiterjeszthetők a kutatások: ilyen például a magyar társalgásban jellemző mondat- és fordulókonstrukció-típusok fonetikai elemzése, a diskurzusjelölő-társulások kutatása vagy a fordulóváltással kapcsolatos metapragmatikai reflexiók.

A témában multimodális, vagyis a látható jelzésekre, például gesztusokra is kiterjedő vizsgálatok is születtek. A multimodális HuComTech korpusz 20 dialógusában elemezték a *mondjuk* diskurzusjelölő típusait, beszédlépésbeli pozícióját (elején, belsejében, végén), témairányításban betöltött szerepét, a dallammal és a kísérő gesztusokkal való összefüggését (ABUCZKI 2013). A *mondjuk* diskurzusjelölő leggyakrabban megnyilatkozás belsejében, lebegő dallammal együtt járva; illetve témakifejtő egységben jelent meg, ezzel a beszélő a már adott téma kibontását, pontosítását vezette be. A diskurzusbeli funkció és a gesztusok együtt járását elemezve az adatok azt mutatták, hogy amikor a példaadás beszédaktusát a beszélő a *mondjuk* diskurzusjelölővel vezette be, akkor a kézzel való gesztikulációt leggyakrabban a diskurzusjelölő kimondásával egy időben vagy előtte 100–2000 ms-mal hajtotta végre (az egyik vagy mindkét tenyér felfelé nézett). Ha a beszélő lexikai előhívás során ejtette a *mondjuk*-ot, azt gyakran kísérte felfelé irányuló tekintet és „felidéző” arckifejezés. Amikor a *mondjuk* valamilyen érték saccolásának bevezetésére szolgált, gyakori volt a fej oldalra, oda-vissza mozgása.

## 5. A társalgáskutatás alkalmazási lehetőségei

A társalgások vizsgálata a nyelvészeti vonatkozásokon túl is számos alkalmazás alapját teremti meg. Ilyenek például a beszéd- és nyelvi zavarok diagnosztikája, a gyógyító orvos-beteg kommunikáció, a célzottan strukturált terápiás beszélgetések, a hatékony osztálytermi kommunikáció, a nyelvtudás mérése, avagy a mesterséges intelligencia különböző alkalmazásai. Az alábbiakban a teljesség igénye nélkül bemutatunk néhány példát arra, hogy a társalgáskutatás eredményei hogyan alkalmazhatók a gyakorlatban.

Egyre több vizsgálat születik arról, hogy a különböző beszéd- és nyelvi zavarok, illetve a beszédbeli tünetekkel is együtt járó neurológiai zavarok hogyan jelennek meg a társalgások során. Jelenleg ugyanis a diagnosztikai eljárások és a nyelvészeti/logopédiai kutatások főként célzott nyelvi és kognitív képességeket vizsgáló tesztek alkalmaznak, és eddig kevés figyelem fordult a természeteshez hasonló kommunikáció során fellépő tünetek elemzésére. A társalgások során olyan pragmatikai funkciók működése is vizsgálható, amelyek nem szükségesek a mondatisméltési, szóelőhívási feladatokban, avagy narratívák létrehozása során. ROUSSEAU, DAVELUY és KOZLOWSKI (2010) például stroke-on átesett betegeknél vizsgálták a társalgásban való részvétel mikéntjét, a verbális és nonverbális kommunikáció sajátosságait. Megállapították, hogy a társalgások elemzésével olyan komplex kommunikációs problémák is vizsgálhatók voltak, amelyek nem pusztán a bal féltekei sérültek szintaktikai és lexikai nehézségeivel és a jobb féltekei sérültek pragmatikai nehézségeivel magyarázhatók. ROUSSEAU, SÈVE és munkatársaik (2010) enyhe és közepesen súlyos Alzheimer-kórban, frontotemporális demenciában és Lewy-testes demenciában vizsgálták a társalgások jellemzőit. Eredményeik nagy különbségeket mutattak a demencia különböző típusai között. A legsúlyosabb nehézségeket a frontotemporális demenciával küzdő betegek mutatták a kommunikációban (üdvözlés, figyelem, részvétel), illetve a verbális és a nonverbális pragmatikában. Az Alzheimer-kóros betegeknél a lexikai-szemantikai műveletek károsodtak jelentősen, míg a kommunikációs folyamatban való részvételük csak enyhén volt érintett. A Lewy-testes demenciával küzdők pedig csak enyhe nehézségeket mutattak a verbális kommunikációban.

A társalgások megvalósulásának módja igen komoly szerepet játszik a gyógyításban, így az orvos-beteg kommunikációban is. Az utóbbi sajátosságai hatással vannak a terápiák hatékonyságára is, ezért kutatása és oktatása az orvosi egyetemeken egyre nagyobb jelentőséggel bír (összefoglalásul lásd KUNA 2020). Egy magyar kutatás például orvos-páciens beszélgetések hangfelvételeit vizsgálta a társalgáselemzés és kognitív pragmatika módszereivel, és feltárta a metapragmatikai kifejezések és a közös megértést segítő nyelvi eszközök fő típusait (pl. megértéssel kapcsolatos explicit reflexiók vagy ismétlés, átfogalmazás), bemutatva, hogy ezeknek a megfelelő alkalmazása segíti az orvos-páciens közti kommunikáció hatékonyságát, a jobb megértést vagy a terápiahűséget (KUNA–HÁMORI 2019). A specifikus terápiás beszélgetések nemcsak a pszichológiai betegségek, de például a neurodegeneratív betegségek gyógyításában is nagy szerepet játszhatnak. Több vizsgálatban (például TAPPEN et al. 1997; TAPPEN et al. 2000;

GENTRY–FISHER 2007; TAPPEN–WILLIAMS 2009) is elemezték, hogy hogyan hat a társalgás az Alzheimer-kóros betegek kezelésére, amelynek során a speciálisan kiképzett terapeuták és gondozók előre meghatározott módon reagáltak a betegek kommunikációjára. Az eredmények szerint a terápiás beszélgetés pozitívan hatott a betegek kedélyállapotára (TAPPEN–WILLIAMS 2009), egyes stratégiák segítettek fenntartani az előrehaladott stádiumban levő betegekkel is a kommunikációt (TAPPEN et al. 1997; GENTRY–FISHER 2007), míg a társalgással kombinált járásterápia hatékonyabbnak bizonyult a mozgás kezelésében (hatékonyabbnak, mint pusztán a hagyományos járásterápia; TAPPEN et al. 2000). Az ezekhez a vizsgálatokhoz/terápiákhoz kifejlesztett stratégiák is részben a társalgáselemzés eredményein alapulnak.

Az osztálytermi kommunikáció elemzése a társalgáskutatás egy specifikus területe (például ANTALNÉ SZABÓ 2006; BORONKAI 2009; KIRÁLY 2015; SÁFRÁNYNÉ MOLNÁR 2016; ASZTALOS 2020). Jelentőségét az adja, hogy általa megismerhetjük, milyen tanári kommunikációs stratégiák a leghatékonyabbak az ismeretek átadására, a tanulók motiválására, illetve a tanulók anyanyelvi kompetenciájának fejlesztésére. A tanórai kommunikáció számos szempontból vizsgálható: elemezhető például a tanári kérdések és instrukciók (ANTALNÉ SZABÓ 2006, 2009; SÁFRÁNYNÉ MOLNÁR 2016), a szóátadás jellemzői (KIRÁLY 2015), a metakognitív kifejezések használata (HÁMORI 2020b), avagy az egyéni tanulói megnyilatkozástípusok jellemzői (ASZTALOS 2020).

A társalgás jellemzőinek az elemzése hozzájárulhat a nyelvtudás megítéléséhez is. VAN OS, DE JONG és BOSKER (2020) például percepciók tesztben vizsgálták, hogy a társalgási fordulók jellemzői hogyan hatnak a folyamatosság megítélésére. Az eredmények azt mutatták, hogy amíg az anyanyelvi beszélők esetében a túlságosan „lelkes” válaszok (azaz a kérdések megszakítása gyors válaszadással) és a túlságosan „vonakodó” válaszok (hosszú reakcióidő után lassú válaszadás) negatívan befolyásolták a nyelvi folyékonyág értékelését, addig a nem anyanyelvi beszélők beszédében csak a túlságosan „vonakodó” válaszok vezettek a folyamatosság rosszabb megítéléséhez. Mindez fontos nyelvtanítási és -értékelési szempontokat vehet fel.

A beszélt nyelvi adatbázisok alapvető szerepet töltenek be a beszédtechnológiai és ehhez kapcsolódóan a mesterséges intelligenciával kapcsolatos kutatásokban. A társalgáskutatás eredményei felhasználhatók a nagy méretű beszélt nyelvi adatbázisok automatikus feldolgozása, illetőleg az ezekben végezhető adatbányászat során (vö. BECHET et al. 2012), beszédtechnológiai és ehhez kapcsolódóan a mesterséges intelligenciával kapcsolatos kutatásokban is. A különböző társalgásokra épülő beszédtechnológiai alkalmazások nagy eredményességgel képesek detektálni a különféle társalgási eseményeket, mint például az átfedő beszédet, avagy a társalgásban részt vevő beszélőket (JURAFSKY et al. 1997; BEKE 2014).

Alan TURING 1950-ben publikálta híres tanulmányát, amelyben azt a kérdést teszi fel, hogy képesek-e a gépek gondolkodni. A Turing-teszt koncepciója szerint: ha egy gép 5 percig úgy tud kommunikálni egy emberrel, hogy az nem ismeri fel, hogy nem emberrel társalog, akkor a gép átment a teszten. Ehhez az emberi társalgás szabályait jól kell ismerni, hiszen úgy kell a gépnek kérdeznie és válaszolnia, ahogyan azt egy valódi beszélgetőpartner tenné. Ez mindeddig nem sikerült egyetlen gépnek sem (bár 2014-ben egy beszélgetőprogramnak sikerült 33%-os arányban meggyőzni a felhasználókat arról, hogy ő egy 13 éves angolul kommunikáló ukrán gyerek,

ez nem tekinthető a kívánt megoldás elérésének), de Turing tanulmánya és a teszt a mai napig nagy hatással van a mesterséges intelligencia fejlesztésére.

Ugyanakkor napjainkban az ember-gép kommunikáció már jelentős eredményeket tud felmutatni, olyannyira, hogy Sophia, a 2016-ban bemutatott humanoid robot 2017-ben Szaúd-Arábiában „állampolgárságot” is kapott. A szélesebb körben való alkalmazások szempontjából is kiemelt figyelmet kap a társas-beszélgető robotok fejlesztése, amelyeknek fontos szerepe lehet az idősek vagy az autizmussal élők mindennapjaiban (SHIMAYA et al. 2019; BEVACQUA et al. 2012; DEVAULT et al. 2014; JOHANSSON et al. 2016; LALA et al. 2017), de például különböző humán erőforrással végzett munkák kiváltásában és a szórakoztatásban is. Ilyen például a 2014-ben bemutatott Pepper robot, amely képes az érzelmek felismerésére, és magyar nyelvű honlapja (<http://pepperrobot.hu/>) szerint megállja a helyét olyan társalgást igénylő feladatokban is, mint recepció, tanácsadó, értékesítési munkatárs, felmérésvezető, vendéglátó stb. Az Erica nevű robot fejlesztésének is az a célja, hogy a robot képes legyen az emberihez hasonló társalgásra, illetve gesztusok kifejezésére. Különlegessége, hogy beszélgetés közben képes folyamatosan jelzéseket adni a beszélő számára arról, hogy figyelmesen hallgatja (LALA et al. 2017). Mindezekhez szükséges nemcsak szemantikailag megfelelő tartalom létrehozása, hanem az is, hogy a robot az adott nyelv társalgásainak fonetikai és társalgási jellemzőit (pl. a tipikus háttérzsoltorna-jelzéseket és mintázatokat, üdvözlési formákat, preferált válaszokat és azok dallamát) is magas szinten ismerje és alkalmazza.

Több magyar fejlesztés is létezik, amelyben az ügyfél egy robottal beszélget. Jelenleg a leg szélesebb körben használt fejlesztés VANDA, a T-Systems digitális üzleti asszisztense, amely képes az ügyféllel chaten keresztül vagy szóban beszélgetve kommunikálni. 2018 óta a Magyar Telekom telefonos ügyfélszolgálatán segíti az ügyintézők munkáját. Sokféle ügyféligényt képes kezelni, például önkiszolgáló GYIK, önkiszolgáló ügyfélszolgálat, konferenciaregisztráció, élő információ, digitális pénzügyi szolgáltatások (FinTech), értékesítési folyamatok, kérdőívkitöltés.

Az említett példák is azt mutatják, hogy a társalgási korpuszoknak és kutatásoknak kiemelkedő szerepük van a jövőbeni elméleti és a közvetlenül felhasználható, alkalmazott kutatások szempontjából is. Maga a társalgáskutatás is dinamikusan fejlődő tudományterület, az alkalmazási lehetőségei pedig igen sokszínűek, a mesterséges intelligenciában betöltött szerepe megkerülhetetlen.

## FOGALMAK

*társalgás; forduló (beszédlépés, turn); beszélőváltás; beszédjog; diskurzusszegmens; lehetséges szóátvételi pont; FTO-érték; diskurzusjelölő; bejósolás / jósló feldolgozás elmélete; külválasztás; önkiválasztás*

## IRODALOM

ABUCZKI Ágnes 2011. Társalgási fordulók multimodális vizsgálata a HuComTech adatbázis informális dialógusaiban. In BODA István Károly – MÓNOS Katalin (szerk.): *Az alkalmazott nyelvészet ma: innováció, technológia, tradíció*. A XX. Magyar Alkalmazott Nyelvészeti Kongresszus kötete. Debreceni Egyetemi Kiadó, Debrecen. 282–288.

- ABUCZKI Ágnes 2013. A *mondjuk* nem konceptuális használatának vizsgálata multimodális kontextusban. In VÁRADI Tamás (szerk.): *VII. Alkalmazott Nyelvészeti Doktorandusz-konferencia, AlkNyelvDok7, Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből*. MTA Nyelvtudományi Intézet, Budapest. 3–16.
- ANTALNÉ SZABÓ Ágnes 2006. *A tanári beszéd empirikus kutatások tükrében*. Magyar Nyelvtudományi Társaság, Budapest.
- ANTALNÉ SZABÓ Ágnes 2009. A tanári instrukciók grammatikája. In KESZLER Borbála – TÁTRAI Szilárd (szerk.): *Diskurzus a grammatikában – grammatika a diskurzusban*. Tinta Könyvkiadó, Budapest. 322–331.
- ASZTALOS Anikó 2020. Az egyéni tanulói megnyilatkozástípusok jellemzői az osztálytermi diskurzusban. *Anyanyelv-pedagógia* 13/2. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=837> (A letöltés ideje: 2021. február 11.)
- ATKINSON, Maxwell J. – HERITAGE, John (eds) 1984. *Structures of Social Action: Studies in Conversation Analysis*. Cambridge University Press, Cambridge.
- BARTH-WEINGARTEN, Dagmar – DEHÉ, Nicole – WICHMANN, Anne (eds) 2009. *Where Prosody Meets Pragmatics*. Emerald Group, Bingley.
- BARTH-WEINGARTEN, Dagmar – REBER, Elisabeth – SELTING, Margret (eds) 2010. *Prosody in Interaction*. Benjamins, Amsterdam.
- BECHET, Frederic – MAZA, Benjamin – BIGOUROUX, Nicolas – BAZILLON, Thierry – EL-BÈZE, Marc – DE MORI, Renato – ARBILLOT, Eric 2012. DECODA: A call-center human-human spoken conversation corpus. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*. 1343–1347. <https://pdfs.semanticscholar.org/ede5/d2d2557188c153272152c9c8117ce8b67898.pdf> (A letöltés ideje: 2021. január 23.)
- BEKE András 2014. *Beszélődetektálás magyar nyelvű spontán társalgásokban*. ELTE Eötvös Kiadó, Budapest.
- BEVACQUA, Elisabetta – COWIE, Roddy – EYBEN, Florian – GUNES, Hatice – HEYLEN, Dirk – MAAT, Mark – MCKEOWN, Gary – PAMMI, Sathish – PANTIC, Maja – PELACHAUD, Catherine – DE SEVIN, Etienne – VALSTAR, Michel – WOLLMER, Martin – SHRODER, Marc – SCHULLER, Bjorn 2012. Building autonomous sensitive artificial listeners. *IEEE Transactions on Affective Computing* 3/2. 165–183.
- BOERSMA, Paul – WEENINK, David 2018. *Praat: Doing Phonetics by Computer*. [http://www.fon.hum.uva.nl/praat/download\\_win.html](http://www.fon.hum.uva.nl/praat/download_win.html). (A letöltés ideje: 2021. január 25.)
- BÓNA Judit 2013. A beszédszünetek fonetikai sajátosságai a beszéd típus függvényében. *Beszédkutatás* 2013. 60–76.
- BORONKAI Dóra 2009. *Bevezetés a társalgáselemzésbe*. Ad Librum Kft., Budapest.
- BORONKAI Dóra 2013. A diskurzusjelölők szerepe társalgási szövegek beszélőváltásaiban a szociokulturális tényezők tükrében. In FEKETE Richárd – KURUCZ Rózsa – NAGY Janka Teodóra (szerk.): *„Szépet, jót, igazat akarva”*. *Tanulmányok N. Horváth Béla 60. születésnapjára*. Pécsi Tudományegyetem Illyés Gyula Kar, Szekszárd. 173–188.



- BRYANT, Gregory A. – FOX TREE, Jean E. 2005. Is there an ironic tone of voice? *Language and Speech* 48/3. 257–277.
- CASILLAS, Marisa – BOOB, Susan C. – CLARK, Eve V. 2014. Taking the floor on time: Delay and deferral in children’s turn taking. In INBAL, Arnon – CASILLAS, Marisa – KURUMADA, Chigusa – ESTIGARRIBIA, Bruno (eds): *Language in Interaction: Studies in honor of Eve V. Clark*. John Benjamins, Amsterdam. 101–114.
- COUPER-KUHLEN, Elizabeth – SELTING, Margret (eds) 1996. *Prosody in Conversation*. Cambridge University Press, Cambridge.
- COUPER-KUHLEN, Elizabeth – FORD, Cecilia E. (eds) 2004. *Sound Patterns in Interaction*. John Benjamins, Amsterdam.
- CRIBLE, Ludivine 2018. *Discourse Markers and (Dis)fluency*. John Benjamins, Amsterdam–Philadelphia.
- DÉR Csilla Ilona 2012. Beszélőváltások során használt diskurzusjelölők a magyar spontán beszédben. *Beszédkutatás 2012*. 130–141.
- DÉR Csilla Ilona 2017. A hát multifunkcionalitása a beszédműfajok és a diskurzusjelölő-társulások függvényében. *Beszédkutatás 2017*. 169–184.
- DÉR Csilla Ilona 2020. *Diskurzusjelölők és társulásaik a magyar nyelvben*. KRE–L’Harmattan, Budapest.
- DÉR Csilla Ilona – CSONTOS Nóra – FURKÓ Péter 2020. A mondván diskurzusjelölő variációs pragmatikai elemzése 2. rész. *Magyar Nyelv* 116/2. 156–165.
- DÉR Csilla Ilona – MARKÓ Alexandra 2007. A magyar diskurzusjelölők szupraszegmentális jelöltsége. In GECSŐ Tamás – SÁRDI Csilla (szerk.): *Nyelvelmélet – nyelvhasználat*. Segédkönyvek a Nyelvészet Tanulmányozásához 74. Kodolányi János Főiskola – Tinta Kiadó, Székesfehérvár–Budapest. 61–67.
- DÉR, Csilla Ilona – MARKÓ, Alexandra 2010. A pilot study of Hungarian discourse markers. *Language and Speech* 53/2. 135–180.
- DE RUITER, Jan – MITTERER, Holger – ENFIELD, Nick J. 2006. Projecting the end of a speaker’s turn: A cognitive cornerstone of conversation. *Language* 82/3. 515–535.
- DEVULT, David – ARTSTEIN, Ron – BENN, Grace – DEY, Teresa – FAST, Ed – GAINER, Alesia – GEORGILA, Kallirroi – GRATCH, Jon – HARTHOLT, Arno – LHOMMET, Margaux – LUCAS, Gale – MARSELLA, Stacy – MORBINI, Fabrizio – NAZARIAN, Angela – SCHERER, Stefan – STRATOU, Giota – SURI, Apar – TRAUM, David – WOOD, Rachel – XU, Yuyu – RIZZO, Albert – MORENCY, Louis-Philippe 2014. SimSensei Kiosk: A virtual human interviewer for healthcare decision support. *International Conference on Autonomous Agents and Multi-Agent Systems* 1. 1061–1068.
- DUNCAN, Starkey 1972. Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology* 23/1. 283–292.
- FORD, Cecilia – THOMPSON, Sandra 1996. Interactional units in conversation: Syntactic, intonational and pragmatic resources for the management of turns. In OCHS, Elinor – SCHEGLOFF, Emanuel – THOMPSON, Sandra (eds): *Interaction and Grammar*. Cambridge University Press, Cambridge. 134–184.

- FRASER, Bruce 1999. What are discourse markers? *Journal of Pragmatics* 31/7. 931–952.
- FURKÓ Bálint Péter 2019. *Diskurzusjelölők és egyéb diskurzuspragmatikai eszközök forgatókönyvek, tankönyvek és irodalmi szövegek párbeszédeiben*. Debreceni Egyetemi Kiadó, Debrecen.
- GENTRY, Ruth A. – FISHER, Jane E. 2007. Facilitating conversation in elderly persons with Alzheimer’s disease. *Clinical Gerontologist* 31/2. 77–98.
- GONZÁLEZ, Montserrat 2004. *Pragmatic Markers in Oral Narrative: The Case of English and Catalan*. Pragmatics and Beyond 122, New Series. John Benjamins, Amsterdam.
- GOLDMAN-EISLER, Frieda 1972. Pauses, clauses, sentences. *Language and Speech* 15/2. 103–113.
- GÓSY Mária – GYARMATHY Dorottya – HORVÁTH Viktória – GRÁCZI Tekla Etelka – BEKE András – NEUBERGER Tilda – NIKLÉCZY Péter 2012. BEA: Beszélt nyelvi adatbázis. In GÓSY Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 9–24.
- GOSZTOLYA, Gábor – BEKE, András – NEUBERGER, Tilda 2019. Differentiating laughter types via HMM/DNN and probabilistic sampling. In SALAH, Albert – KARPOV, Alexey – POTAPOVA, Rodmonga (eds): *Speech and Computer. SPECOM 2019. Lecture Notes in Computer Science*. Springer, Cham. 122–132.
- GRÁCZI, Tekla Etelka – BATA, Sarolta 2010. The effect of familiarization on temporal aspects of turn-taking: A pilot study. *Acta Linguistica Hungarica* 57/2–3. 307–328.
- GYARMATHY Dorottya 2015. Az *úgyhogy* funkcióbővülése a spontán beszédben. *Beszédkutatás 2015*. 92–106.
- GYARMATHY Dorottya – KREPSZ Valéria – DÉR Csilla Ilona – HÁMORI Ágnes – HORVÁTH Viktória 2020. Háttércsatorna-jelzések: határterület és új szempontok a diskurzusok elemzésében. In FÓRIS Ágota – BÖLCSKEI Andrea (főszerk.) – DÉR Csilla Ilona – CSONTOS Nóra (szerk.): *Nyelv, kultúra, identitás. Alkalmazott nyelvészeti kutatások a 21. századi információs térben. II. Pragmatika, diskurzuselemzés, interkulturális kommunikáció*. A MANYE Kongresszusok Előadásai 12/2. Akadémiai Kiadó, Budapest. 113–118.
- HÁMORI Ágnes 2006. A társalgási műfajokról. In TOLCSVAI Nagy Gábor (szerk.): *Szöveg és típus. Szövegtipológiai tanulmányok*. Tinta Kiadó, Budapest. 157–181.
- HÁMORI Ágnes 2020a. „Szóval, úgyhogy, tehát...” A „bizonytalan zárású” mondatok mint a társalgás egy sajátos mondat típusa: konstrukció a fordulóváltás-szerveződés és a pragmatikai stratégiák szolgálatában. Előadás az *Új nézőpontok a magyar nyelv leírásában* című konferencián, 2020. november 20., ELTE BTK, Budapest.
- HÁMORI Ágnes 2020b. „Mi lehet az oka ennek, mit gondoltok?” – A metakognitív nyelvhasználat elemzése az osztálytermi diskurzusban. In LUDÁNYI Zsófia – JÁNK István – DOMONKOSI Ágnes (szerk.): *A nyelv perspektívája az oktatásban*. Líceum Kiadó, Eger. 273–297.
- HÁMORI Ágnes – DÉR Csilla Ilona 2020. A beszélőváltások dinamikus jellemzői II.: Pragmatikai/társalgáselemzési aspektusok. Lexikális elemek a fordulók végén és kezdetén a magyar társalgásokban: a *tehát*, az *úgyhogy* és a *hát*. Előadás a *Beszédkutatás – Speech Research 2020* konferencián, 2020. december 15., Nyelvtudományi Intézet, Budapest.
- HÁMORI Ágnes – HORVÁTH Viktória 2019. Társalgás, beszélőváltás és diskurzusszerveződés új megközelítésben – fonetikai jellemzők és pragmatikai tényezők összefüggései magyar társalgásokban. *Beszédkutatás* 27. 134–153.

- HELDNER, Mattias – EDLUND, Jens 2010. Pauses, gaps and overlaps in conversations. *Journal of Phonetics* 38/4. 555–568.
- HORVÁTH Viktória 2009. *Funkció és kivitelezés a megakadásjelenségekben*. PhD-értekezés. ELTE BTK, Budapest.
- HORVÁTH Viktória – KREPSZ Valéria – GYARMATHY Dorottya – HÁMORI Ágnes – BÓNA Judit – DÉR Csilla Ilona – WEIDL Zsófia 2019. Háromfős társalgások annotálása a BEA-adatbázisban. *Nyelvtudományi Közlemények* 115. 255–274.
- HORVÁTH Viktória – GYARMATHY Dorottya – KREPSZ Valéria – HÁMORI Ágnes 2020. A társalgások szünetezése – különböző funkciók és realizációk. In MAGYARI Sára – BARTHA Krisztina (szerk.): *Nyelvi közösségek – közösségi perspektívák*. Partium Kiadó, Nagyvárad. 99–114.
- HUNYADI, László – VÁRADI, Tamás – KOVÁCS, György – SZEKRÉNYES, István – KISS, Hermina – TAKÁCS, Karolina 2018. Human–human, human–machine communication: on the HuComTech multimodal corpus. Presentation at the *CLARIN Annual Conference, 2018*. 2018. October 8–10, Pisa, Italy.
- HUNYADI, László 2019. Agreeing/disagreeing in a dialogue: multimodal patterns of its expression. *Frontiers in Psychology* 10. 1–9.
- HUSZÁR Anna – KREPSZ Valéria – GYARMATHY Dorottya – HORVÁTH Viktória 2021. Beszédjellemzők és időzítés háromfős társalgások beszélőváltásaiban. *Alkalmazott Nyelvtudomány* XXI/1. 93–112.
- IVÁNYI Zsuzsanna 2001. A nyelvészeti konverzációelemzés. *Magyar Nyelvőr* 125/1. 74–93.
- JOHANSSON, Martin – HORI, Tatsuuro – SKANTZE, Gabriel – HÖTHKER, Anja – GUSTAFSON, Joakim 2016. Making turn-taking decisions for an active listening robot for memory training. In AGAH, Arvin – CABIBIHAN, John-John – HOWARD, Ayanna M. – SALICHS, Miguel A. – HONGSHENG, He (eds): *Social Robotics: 8th International Conference, ICSR 2016*. Springer International Publishing, Cham. 940–949.
- JURAFSKY, Daniel – BATES, Rebecca – COCCARO, Noah – MARTIN, Rachel – MATEER, Marie – RIES, Klaus – SHRIBERG, Elizabeth – STOCKLE, Andreas – TAYLOR, Paul – VAN ESS-DYKEMA, Carol 1997. Automatic detection of discourse structure for speech recognition and understanding. In *Proceedings of the 1997 IEEE Workshop on Speech Recognition and Understanding*. IEEE, Santa Barbara. 88–95.
- KIRÁLY Flóra 2015. A beszédfordulók és a szóátadások vizsgálata osztálytermi kontextusban. *Anyanyelv-pedagógia* 8/1. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=552> (A letöltés ideje: 2021. január 23.)
- KUNA Ágnes 2020. Változás az orvos-beteg kommunikációban: Változó szemlélet, módszer és gyakorlat. *Magyar Nyelvőr* 144/3. 280–303.
- KUNA Ágnes – HÁMORI Ágnes 2019. „Hallgatom, mi a panasz?” A metapragmatikai reflexiók szerepei és mintázatai az orvos-beteg interakciókban. In LACZKÓ Krisztina – TÁTRAI Szilárd (szerk.): *Kontextualizáció és metapragmatikai tudatosság*. ELTE Eötvös József Collegium, Budapest. 215–241.

- LALA, Divesh – MILHORAT, Pierrick – INOUE, Koji – ISHIDA, Masanari – TAKANASI, Katsuya – KAWAHARA, Tatsuya 2017. Attentive listening system with backchanneling, response generation and flexible turn-taking. In *Proceedings of the SIGDIAL 2017 Conference*. Saarbrücken. 127–136. <https://www.aclweb.org/anthology/W17-5516>. (A letöltés ideje: 2019. június 6.)
- LEVINSON, Stephen C. 1983. *Pragmatics*. Cambridge Textbooks in Linguistics, Cambridge.
- LEVINSON, Stephen C. 2016. Turn-taking in human communication – Origins and implications for language processing. *Trends in Cognitive Sciences* 20/1. 6–14.
- LEVINSON, Stephen C. – TORREIRA, Francisco 2015. Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology* 6. 731.
- LOCAL, John – KELLY, John – WELLS, Bill 1986. Towards a phonology of conversation: Turn-taking in Tyneside English. *Journal of Linguistics* 22/2. 411–437.
- MARKÓ Alexandra 2005a. A temporális szerkezet jellegzetességei eltérő kommunikációs helyzetekben. *Beszédkutatás* 2005. 63–77.
- MARKÓ Alexandra 2005b. *A spontán beszéd néhány szupraszegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hümmögés vizsgálata*. PhD-értekezés. ELTE BTK, Budapest.
- MARKÓ Alexandra 2009. Stigmatizált hanglejtésforma a spontán beszédben. *Beszédkutatás* 2009. 88–106.
- MARKÓ Alexandra 2015. Jelkimaradások a beszédben: szünet és hallgatás. In BÁRTH M. János – BODÓ Csanád – KOCSIS Zsuzsanna (szerk.): *A nyelv dimenziói. Tanulmányok Juhász Dezső tiszteletére*. Magyar Nyelvtudományi Társaság – ELTE BTK Magyar Nyelvtudományi és Finnugor Intézet, Budapest. 515–524.
- MARKÓ Alexandra – GÓSY Mária 2015. A megszólalás stratégiái társalgásban. In BÁRDOSI Vilmos (szerk.): *A nyelvi pragmatika kérdései szinkrón és diakrón megközelítésben*. Tinta Kiadó, Budapest. 159–168.
- MARKÓ, Alexandra – GÓSY, Mária – NEUBERGER, Tilda 2014. Prosody patterns of feedback expressions in Hungarian spontaneous speech. In CAMPBELL, Nick – GIBBON, Dafydd – HIRST, Daniel (eds): *Social and Linguistic Speech Prosody: Proceedings of the 7th International Conference on Speech Prosody*. Science Foundation Ireland, Dublin. 482–486.
- MÜLLER, Simone 2005. *Discourse Markers in Native and Non-Native English Discourse*. John Benjamins, Amsterdam–Philadelphia.
- MACLAY, Howard – OSGOOD, Charles E. 1959. Hesitation phenomena in spontaneous English speech. *Word* 15/1. 19–44.
- NEUBERGER Tilda 2012. Nonverbális hangjelenségek a spontán beszédben. In GÓSY Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 215–235.
- ODGEN, Richard 2012. The phonetics of talk in interaction – Introduction to the Special Issue. *Language and Speech* 55/1. 3–11.
- PRESTON, Denis L. – NIEDZELSKI, Nancy A. 2010. *Folk Linguistics. Trends in Linguistics. Studies and Monographs*. De Gruyter, Mouton.

- REDEKER, Gisela 1990. Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics* 14/3. 367–381.
- ROBERTS, Felicia – FRANCIS, Alexander L. – MORGAN, Melanie 2006. The interaction of interturn silence with prosodic cues in listener perceptions of “trouble” in conversation. *Speech Communication* 48/9. 1079–1093.
- ROUSSEAUX, Marc – DAVELUY, Walter – KOZŁOWSKI, Odile 2010. Communication in conversation in stroke patients. *Journal of Neurology* 25/7. 1099–1107.
- ROUSSEAUX, Marc – SÈVE, Amandine – VALLET, Marion – PASQUIER, Florence – MACKOWIAK-CORDOLIANI, Marie Anne 2010. An analysis of communication in conversation in patients with dementia. *Neuropsychologia* 48/13. 3884–3890.
- RUDER, Kenneth F. – JENSEN, Paul J. 1972. Fluent and hesitation pauses as a function of syntactic complexity. *Journal of Speech and Hearing Research* 15/1. 49–58.
- SACKS, Harvey – SCHEGLOFF, Emanuel A. – JEFFERSON, Gail 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50/4. 696–735.
- SCHEGLOFF, Emanuel A. 1968. Sequencing in conversational openings. *American Anthropologist* 70/6. 1075–1095.
- SCHIFFRIN, Deborah 1987. *Discourse Markers*. Cambridge University Press, Cambridge.
- SCHIRM Anita 2011. *A diskurzusjelölők funkciói: a hát, az -e és a vajon elemek története és jelenkori szinkrón státusa alapján*. PhD-értekezés. Szegedi Tudományegyetem Nyelvtudományi Doktori Iskola, Szeged.
- SCHIRM Anita 2017. A diskurzusjelölők és a szövegtípusok viszonyáról. *Magyar Nyelv* 113/3. 330–341.
- SHIMAYA, Jiro – YOSHIKAWA, Yuichiro – KUMAZAKI, Hirokazu – MATSUMOTO, Yoshio – MIYAO, Masumoto – ISHIGURO, Hiroshi 2019. Communication support via a tele-operated robot for easier talking: Case/laboratory study of individuals with/without autism spectrum disorder. *International Journal of Social Robotics* 11/1. 171–184.
- SHRIBERG, Elizabeth – STOLCKE, Andreas – BARON, Dan 2001. Observations on overlap: Findings and implications for automatic processing of multi-party conversation. In DALSGAARD, Paul – LINDBERG, Børge – BENNER, Henrik – ZHENG-HUA, Tan (eds): *Proceedings of Eurospeech 2*. ISCA, Aalborg, Denmark. 1359–1362.
- SIDNELL, Jack 2001. Conversational turn-taking in a Caribbean English Creole. *Journal of Pragmatics* 33/8. 1263–1290.
- SIDNELL, Jack – STIVERS, Tanya (eds) 2014. *The Handbook of Conversation Analysis*. Blackwell, Malden–Oxford.
- STIVERS, Tanya – ENFIELD, Nick J. – BROWN, Penelope – ENGLERT, Christina – HAYASHI, Makoto – HEINEMANN, Trine – HOYMAN, Gertie – ROSSANO, Federico – DE RUITER, Jan – YOON, Kyung-Eun – LEVINSON, Stephen C. 2009. Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America* 106/26. 10587–10592.

- SÁFRÁNYNÉ MOLNÁR Mónika 2016. A verbális tanári instrukciók az osztálytermi diskurzusban I. *Anyanyelv-pedagógia* 9/2. <http://anyanyelv-pedagogia.hu/cikkek.php?id=624> (A letöltés ideje: 2021. január 6.)
- SZEKRÉNYES István 2019. *Prozódiai jellemzők gépi feldolgozása és hasznosítása élőnyelvi korpuszok elemzésében*. PhD-értekezés. Debreceni Egyetem Nyelvtudományok Doktori Iskola, Debrecen.
- TAPPEN, Ruth M. – ROACH, Kathryn E. – APPLGATE, E. Brooks – STOWELL, Paula 2000. Effect of a combined walking and conversation intervention on functional mobility of nursing home residents with Alzheimer disease. *Alzheimer Disease and Associated Disorders* 14/4. 196–201.
- TAPPEN, Ruth M. – WILLIAMS, Christine L. 2009. Therapeutic conversation to improve mood in nursing home residents with Alzheimer’s disease. *Research in Gerontological Nursing* 2/4. 267–275.
- TAPPEN, Ruth M. – WILLIAMS-BURGESS, Christine – EDELSTEIN, Jackie – TOUHY, Theris – FISHMAN, Sarah 1997. Communicating with individuals with Alzheimer’s disease: examination of recommended strategies. *Archives of Psychiatric Nursing* 11/5. 249–256.
- TEN BOSCH, Louis – OOSTDIJK, Nelleke – DE RUITER, Jan 2004. Turn-taking in social talk dialogues: temporal, formal and functional aspects. Presentation at the *SPECOM’2004. 9th Conference Speech and Computer*, 2004. September 20–22. St. Petersburg, Russia.
- TEN BOSCH, Louis – OOSTDIJK, Nelleke – BOVES, Lou 2005. On temporal aspects of turn taking in conversational dialogues. *Speech Communication* 47/1–2. 80–86.
- TORREIRA, FRANCISCO – BÖGELS, Sarah – LEVINSON, Stephen C. 2015. Breathing for answering: the time course of response planning in conversation. *Frontiers in Psychology* 6. 284. <https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00284/full> (A letöltés ideje: 2021. január 26.)
- TURING, Alan M. 1950. Computing machinery and intelligence. *Mind* 49/236. 433–460.
- VAN DONZEL, Monique E. – VAN BEINUM, Florian J. 1996. Pausing strategies in discourse in Dutch. In *Proceedings of ICSLP ’96*. Philadelphia, PA, USA. 1029–1032.
- VAN OS, Marjolein – DE JONG, Nivja H. – BOSKER, Hans Rutger 2020. Fluency in dialogue: Turn-taking behavior shapes perceived fluency in native and nonnative speech. *Language Learning* 70/4. 1183–1217.
- WEIDL Zsófia 2019. Egyet nem értések fonetikai jellemzői középiskolások beszédében. In VÁRADI Tamás – GRÁCZI Tekla Etelka – LUDÁNYI Zsófia (szerk.): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019. XIII. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. MTA Nyelvtudományi Intézet, Budapest. 138–142.
- WEILHAMMER, Karl – RABOLD, Susen 2003. Durational aspects in turn taking. In SOLÉ, Maria-Josep – RECASENS, Daniel – ROMERO, Joaquin (eds): *Proceedings of the 15th International Conference of Phonetic Sciences*. Causal Productions, Barcelona. 2145–2148.
- WEINGARTOVÁ, Lenka – CHURAŇOVÁ, Eliška – ŠTURM, Pavel 2015. Transitions, pauses and overlaps: Temporal characteristics of turn-taking in Czech. *Proceedings of the 7th International Conference on Speech Prosody*. Dublin, Ireland. 2014–87.

- WENNERSTROM, Ann – SIEGEL, Andrew F. 2003. Keeping the floor in multiparty conversations: Intonation, syntax and pause. *Discourse Processes* 36/2. 77–107.
- ZELLNER, Brigitte 1998. Temporal structures for fast and slow speech rate. In *ESCA/COCOSDA. Third International Workshop on Speech Synthesis*. November 1998, Jenolan Caves, Australia. 143–146.

### **KÖSZÖNETNYILVÁNÍTÁS**

A „Beszédegységek fonetikai jellemzőinek összefüggései” című pályázat keretében készült kutatásokat az NKFI K-128810 pályázat, a Tématerületi Kiválósági Program, valamint „A diskurzusjelölők nyelv- és műfajközi kutatása (multidiszciplináris diskurzuskutatások)” kutatócsoport 20633B800 témaszámú, a Károli Gáspár Református Egyetem által finanszírozott pályázata támogatta.

---

A  
**BESZÉD • KUTATÁS • ALKALMAZÁS**

című sorozat eddig megjelent kötetei:

MARKÓ ALEXANDRA:

**Az irreguláris zöngé funkciói a magyar beszédben**

ISBN 978-963-312-195-5

BÓNA JUDIT:

**A spontán beszéd sajátosságai az időskorban**

ISBN 978-963-312-199-3

HORVÁTH VIKTÓRIA:

**Hezitációs jelenségek a magyar beszédben**

ISBN 978-963-312-205-1

NEUBERGER TILDA:

**A spontán beszéd sajátosságai gyermekkorban**

ISBN 978-963-312-204-4

BEKE ANDRÁS:

**Gépi beszélődetektálás magyar nyelvű spontán társalgásokban**

ISBN 978-963-312-234-1

DEME ANDREA:

**Magánhangzók ejtése és észlelése a szopránéneklésben**

ISBN 978-963-312-261-7

TAR ÉVA:

**Fonológiai fejlődés, variabilitás, beszédhanghibák**

ISBN 978-963-312-289-1

AUSZMANN ANITA:

**Magyar gyermekek magánhangzóinak akusztikai-fonetikai jellemzői**

ISBN 978-963-312-295-2

KOHÁRI ANNA:

**Időzítési mintázatok a magyar beszédben**

ISBN 978-963-312-296-9



BÓNA JUDIT – HORVÁTH VIKTÓRIA (szerk.):  
**Az anyanyelv-elsajátítás folyamata hároméves kor után**  
ISBN 978-963-312-310-2

MARKÓ ALEXANDRA (szerk.):  
**Tanulmányok a beszédtudomány alkalmazásainak köréből**  
ISBN 978-963-489-358-5  
ISBN 978-963-489-359-2 (pdf)

*A Beszéd – Kutatás – Alkalmazás* című tudományos könyvsorozat 11. kötetként közrebocsátott gyűjtemény célja az, hogy az oktatásban felhasználható friss, aktuális áttekintést adjon a beszédtudománnyal kapcsolatos alkalmazásorientált kutatásokról, ezek módszertanáról, olykor módszertani nehézségeiről, eredményeiről és az alkalmazások köréről. A témák között szerepel a hallás, a beszédészlelés, a beszédmegértés működése és vizsgálata; a beszéd mint biomarker; a neurodegeneratív kórképek hatása a beszédre; az orofaciális miofunkcionális diszfunkció jelensége; az akcentus; a gyermekekhez és a felnőttekhez szóló beszéd eltérései; a beszédnek az életkorral és a nemmel összefüggő változatossága; a beszélőazonosítás; a beszélőváltások működése a társalgásban. A legfrissebb tudományos eredmények is megjelennek: a tanulmányok túlnyomórészt most zajló tudományos projektekbe vagy éppen csak lezárult kutatásokba engednek betekintést az olvasónak. Az eredmények egy része itt olvasható először magyarul.

ISBN 978-963-489-358-5



9 789634 893585