

Színusz hullámú és amplitúdó modulált beszéd észlelése: perceptuális tanulás és összefüggés nyelvi képességtesztekkel¹

JAKAB ZOLTÁN – MOHAI KATALIN – VIG JULIANNA

jakab.zoltan@barczy.elte.hu

mohai.katalin@barczy.elte.hu

vig.julianna@barczy.elte.hu

ABSZTRAKT

Háttér és célok: A színusz hullámú és amplitúdó modulált beszéd észlelésének vizsgálata több évtizedes múltra tekint vissza; célja kezdetben a beszédészlelés és nyelvi feldolgozás alapvető kérdéseinek vizsgálata volt, majd később, ezzel összefüggésben, a beszédészlelés és fonológiai feldolgozás gyermekkori fejlődésének jobb megértése. Jelen vizsgálat a fejlődés két jelenségére fókuszál: a gyors perceptuális tanulás meglétére, és e két fajta beszédinger észlelésének nyelvi képességekkel való kapcsolatára óvodás- és iskoláskorban.

Módszer: Három korcsoport (óvodások, kisiskolások és felnőttek, összesen 134 személy) vett részt a vizsgálatokban. Az alkalmazott beszédészlelési feladatot a *Hallás/ beszédértés zajban (Hearing in noise)* teszt gyermekváltozatában használt mondatok egy részének (HINT-C) hozzávetőleges magyar fordításait használta fel színusz hullámú (SWS) és amplitúdó modulált (AM) beszédingerek szintézisére. Összesen húsz, egyenként négy szóból, illetve nyolc szótagból álló mondatból három formánst tartalmazó színusz hullámú és négy sávós amplitúdó modulált változatokat szintetizáltunk, és ezeket használtuk kísérleteinkben. A nyelvi képességek felmérésére több magyar nyelvű tesztet, valamint a Woodcock-Johnson tesztből a fordított számerjedelem próbát használtuk. Kérdéseink megválaszolására korreláció- és regresszióelemzéseket végeztünk adatainkon.

Eredmények: Sikertült kimutatnunk a beszédészlelési teljesítmény határozott csoportszintű növekedését egy 20 szintetizált mondat (SWS vagy AM) bemutatásából álló kísérleti ülés során; e hatást gyors perceptuális tanulás hatásaként értelmeztük. A nyelvi tesztek közül a kétféle szintetizált beszéddel a fonológiai feldolgozást mérő alskálák (fonémamegkülönböztetés; gyorsított beszéd észlelése), a mondatismétlés és a produktív szókinés korreláltak leginkább. Ezenkívül egy morfémaproduktív szubteszt is elég szoros kapcsolatban állt az SWS és AM beszéd észlelésével.

Következtetések: A hangtani tudatosság és az álszóismétlés kapcsolata a beszédészlelési feladatunkkal a fonológiai feldolgozás fontos szerepét jelzi e feladatokban. A morfémaproduktívval való összefüggés a nyelvtani feldolgozás támogató, restaurációs hatásának következménye lehet. A kapott eredmények alapján Pszichológiai Intézetünk arra esélyt látunk arra, hogy a színusz hullámú és amplitúdó modulált beszéd észlelése nyelvi diagnosztikai, illetve fejlesztő eszközök kidolgozásában felhasználható legyen.

Kulcsszavak: a beszédészlelés fejlődése, színusz hullámú beszéd, amplitúdó modulált beszéd, perceptuális tanulás, nyelvi képességtesztek

¹ A tanulmány az ELTE Bárczi Gusztáv Gyógypedagógiai Kar Gyógypedagógiai Pszichológiai Intézete és a Magyar Pszichológiai Társaság Gyógypedagógiai Pszichológiai Szekciója által a Magyar Tudomány Ünnepe, 2019 és a Ranschburg Pál emlékévé, 2020 alkalmából rendezett *Szisztematikus kutatások és evidencia-alapú gyakorlatok a humán fogvatékosságok és az atipikus fejlődés pszichológiájában* című konferencián (Budapest, 2019. november 21., ELTE BGGYK) elhangzott előadáson alapul (Jakab Zoltán és Vig Julianna: A színusz hullámú és amplitúdó-modulált beszéd észlelésének diagnosztikai lehetőségei). (A tematikus szám szerkesztői.)

HÁTTÉR ÉS CÉLOK

A szinuszhullámú és amplitúdómodulált beszéd kutatásának alapvető kérdései

Az emberi beszéd redundáns inger, vagyis jóval több információt tartalmaz annál, mint amire szükség van ahhoz, hogy a hallgatók számára érthető legyen. Beszédértés zajos környezetben is lehetséges, és több, egyszerre hallható beszédhelyzetből is ki tudunk választani egyet, amelyet figyelemmel követünk. Sőt, ha a természetes beszédben lévő akusztikus információt különböző módszerekkel mesterségesen csökkentjük, sok esetben akkor is képesek vagyunk az így keletkező elszegényített beszédinger részben vagy egészben megérteni. A beszédingernek ilyen átalakításával közelebb kerülhetünk annak megértéséhez, hogy agyunk hogyan dolgozza föl a beszédben lévő információt, és ennek mely összetevői szükségesek, illetve elégségesek a beszéd megértéséhez.

A Yale Egyetem Haskins Laboratóriumaiban kezdődött el a természetes beszéd egyszerűsített, redukált formáinak – elsősorban a *szinuszhullámú beszédnek* (sine-wave speech; SWS) és *amplitúdómodulált beszédnek* (vocoded speech; VOC; amplitude envelopes; AE) – a rendszeres vizsgálata (Remez, Rubin, Pisoni, & Carrell, 1981; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995; Remez, Rubin, Nygaard, & Howell, 1987; Remez, Rubin, Berns, Pardo, & Lang, 1994; Remez & Rubin, 1990; Rosner és mtsai, 2003). A szinuszhullám-transzformáció a természetes beszédből kiemeli az alsó három formánst, majd ezeket változó frekvenciájú szinuszhullámokkal helyettesíti, és az eredeti beszédinger egyéb frekvenciaösszetevőit elhagyja. Az eredmény egy sípolásra, madárfüttyre emlékeztető hang, mely általános jellemzőiben nem emlékeztet a természetes beszédre. Az amplitúdómodulált beszéd előállításánál során először a természetes beszéd spektrogramját néhány (jellemzően 2-8) frekvenciasávra osztják. Ezután minden frekvenciasávban a szintetizáló eljárás megméri az amplitúdó idői változásait, és csak ezt az információt tartja meg az egyes sávokból. Végül az amplitúdóváltozásokat az eredeti sávokkal azonos frekvenciatartományú homogén zajra ültetik rá (amplitúdó-moduláció). Az eredmény egy szélsusugásra, suttogásra emlékeztető hangzás, mely, szemben a szinuszhullámú beszéddel, nem tartja meg a beszéd formánsainak pontos mozgását, mégis elég jól érthető. Mindkét redukált beszédforma esetén a naív személyek figyelmét fel kell hívni arra, hogy szokatlanul hangzó beszédet hallanak, egyébként nem beszédként érzélik őket.

Ezek a kutatások reflektáltak a hallási szerveződés más megközelítéseire, így például a hallásijelenet-elemzés igen alaposan – de elsősorban nem nyelvi ingerek alapján – kidolgozott elméletére (Bregman, 1990). Kritikájuk szerint a beszédészlelés alapját képező perceptuális integráció más természetű, mint a nem beszédingerké. Utóbbit a hallásijelenet-elemzés az alaklélektan elveiből kiindulva, azokat a hallás esetére továbbfejlesztve írta le, míg Robert Remez és munkatársai részletesen érveltek amellett, hogy ezek az elvek a beszédészlelés folyamatának megértéséhez nem megfelelőek – és érveik jelentős része éppen a redukált beszédformák, elsősorban a szinuszhullámú beszéd észlelésén alapult (Remez és mtsai, 1994; Remez, 2005). A szinuszhullámú beszéd észlelését olyan elvek, mint a hasonlóság, harmonicitás, egy irányba tartás, jó folytatás nem képesek megmagyarázni, annak ellenére, hogy ezek az elvek pl. a zenei formák észlelésében észrevehető szerepet játszanak. A beszéd észlelése eszerint más elvek alapján, sőt a hallási feldolgozó rendszeren belül egy másik alrendszer útján valósulna meg. Később a redukált beszédformák észlelésének vizsgálatában egyre nagyobb hangsúlyt kapott a fejlődési szempont; a továbbiakban ennek az iránynak a legfontosabb alapfogalmait vezetjük be.

A statikus és dinamikus koartikuláció hatásai a beszédészlelés fejlődésében

A beszédészlelés kutatásának egyik alapvető kérdése, hogy a beszéd akusztikai ingerének folytonos, folyamatosan változó jellemzőit hogyan bontjuk fel és soroljuk be diszkrét, invariáns fonémaosztályokba. Egy régebbi felfogás szerint erre a célra a beszéd-folyam rövidebb, spektrálisan invariáns szakaszait használjuk, és az ezek között lévő átmeneteket figyelmen kívül hagyjuk (Blumstein & Stevens, 1979). Ezt a felfogást később számos adat cáfolta, így például az a megfigyelés is, hogy szótagokba ágyazott magánhangzókat könnyebb azonosítani, mint izoláltan ejtetteket (Strange, Verbrugge, Shankweiler, & Edman, 1976). A későbbi kutatások során világossá vált, hogy a beszéd észlelésében a *dinamikus koartikulációs struktúra*² alapvető szerepet játszik (Jenkins et al., 1983; Strange, Jenkins, & Johnson, 1983). Az említett korábbi felfogásban alapvető *statikus koartikulációs struktúra* a szótagokba ágyazott magánhangzók viszonylag állandó középső szakaszán, illetve – mássalhangzó-kontextustól függően – egy egész szótagon végighúzó, többé-kevésbé állandó formánsmintázatnak felel meg. Ezzel szemben a *dinamikus koartikuláció* éppen a mássalhangzó-magánhangzó határokon létrejövő formáns-szerkezet-változásokat jelenti. A mássalhangzó-kontextus hatását a statikus és dinamikus komponens arányára jól illusztrálja a /b/ és a /d/ fonémák közötti különbség. A /b/V/b/ típusú szótagok (két b hang közé elhelyezett valamely magánhangzó) kiejtése során a nyelv és az állkapocs pozíciója végig viszonylag stabil, az akadályt az ajkak képezik. Ezért a mássalhangzóhoz időben közel nem változik meg lényegesen a formáns-szerkezet. Ezzel szemben a /d/V/d/ szótagokban a /d/ képzése során az akadályt a nyelv csúcának az alveoláris gerinchez való hozzáértése képezi, ennek során felemelkedik a nyelv egésze, ami lényegesen megváltoztatja a szájüreg akusztikáját, így a formáns-frekvenciák változékonyabbak a szótagok képzése során.

A szinusz hullámú beszéd azáltal, hogy élesen kirajzolja a formánskontúrokat (bár a felharmonikusokat és zörejkomponenseket megszünteti), a statikus és dinamikus koartikulációval kapcsolatos információt is megőriz. Ezzel szemben az amplitúdómoduláció azáltal, hogy viszonylag széles frekvenciasávokra bontja a természetes beszéd spektrogramját, jobban lerontja a dinamikus koartikuláció (formáns-változások) jeleit, és valamivel több statikus koartikulációs információt őriz meg. Ebben a megközelítésben a beszédészlelés fejlődése a következőképpen értelmezhető (Nitttrouer & Lowenstein, 2014). 4-5 éves kortól a gyerekek már meglehetősen jók a szinusz hullámú és amplitúdómodulált beszéd észlelésében is, ám a felnőtteknél jobban támaszkodnak a dinamikus koartikulációs jelekre. Ennek oka, hogy figyelmük iskoláskor előtt elsősorban teljes szavakra, illetve szótagokra irányul. A felnőttek, fejlettebb szegmentációs képességüknek köszönhetően, már képesek kiemelni rövidebb ideig fennálló állandó formánsmintázatokat is, így a statikus koartikuláció jeleire érzékenyebbek. Ezt igazolja az az adat is, hogy felnőttek teljesítménye amplitúdómodulált szótagok észlelésében gyengébb /d/V/d/ típusú szótagok, mint /b/V/b/ típusúak esetén. E folyamat

² Általánosságban a koartikuláció azt jelenti, hogy a beszéd képzésekor az egymást követő beszédhangok egymást befolyásolják; egymás számára artikulációs kontextust képeznek. Hangképző szerveink állapota beszéd közben folyamatosan változik, így például egy magánhangzó ejtése közben már a szájüreg, a nyelv, az ajkak már „rákészülnek” a következő mássalhangzóra, és ezzel kicsit módosítják az adott magánhangzó akusztikus sajátosságait.

következménye, hogy gyerekeknél a szinuszbeszéd feldolgozási előnye nagyobb az amplitúdómodulálttal szemben; felnőtteknél kiegyenlítettebb a teljesítmény (Nittrouer, Lowenstein, & Packer, 2009; Nittrouer & Lowenstein, 2014). Úgy tűnik tehát, hogy az amplitúdómodulált beszéd észlelési teljesítménye gyorsabban javul a korral, mint a szinusz hullámúé. Ez pedig azért van, mert a dinamikus struktúrára gyerekek és felnőttek támaszkodnak (bár ebben is lehet némi javulás a korral), míg a statikus struktúrát a gyerekek nem, vagy alig használják, a felnőttek viszont sokkal inkább.

Gyors perceptuális tanulás

A vizuális észlelésben és a beszédpercepcióban is megfigyelhető az azonnali perceptuális tanulás jelensége, melynek lényege, hogy egy elszegényített vizuális ingert kezdetben nem vagyunk képesek értelmezni, majd az eredeti, részletgazdag ábra megnézését követően rögtön látjuk, felismerjük ugyanazt a tartalmat a nehezített változatban. A vizuális modalításban ennek van némi irodalma (Mooney, 1957; Kovács & Eisenberg, 2005). A gyors perceptuális tanulás jelensége ugyanakkor a látás kutatásában összekapcsolódik a *perceptuális integráció* fogalmával. A tér és a tárgyak felnőttekre jellemző vizuális észlelése hosszú fejlődés eredménye, mely például korai depriváció esetén súlyosan sérülhet (Kovács, 2000; Maurer, Lewis, & Mondloch, 2005). A vizuális modalításban az alacsony szintű perceptuális integráció egyik jellegzetes példája a kontúrintegráció, melynek megjelenése nem igényel hosszú tapasztalatszerzést, és már néhány hónapos csecsemőknél is megnyilvánul (Kovács, 2000). A magas szintű vizuális integráció ugyanakkor időben kiterjedt perceptuális tanulást feltételez (Hudák, Jakab, & Kovács, 2013), és a gyors perceptuális tanulás jelenségei feltételezik e korai perceptuális tanulás meglétét (pl. gyermekkor után látóvá vált született vak személyeknél a fent említett gyors vizuális tanulás nem figyelhető meg).

A szinusz hullámú, illetve amplitúdómodulált beszéd észlelése során is tapasztalható hasonló jelenség: amikor először hallunk egy szintetizált beszédingert, általában nem értjük, azonban a természetes kiejtés meghallgatása után a szintetizált változat azonnal teljesen érthetővé válik, ami egy jellegzetes aha-élménnyel is jár³. Ezt a jelenséget próbáltuk meg mérhetővé tenni egy korábbi vizsgálatunkban (Jakab, Nagyné Ringer, Víg, & Szabó, 2018), amelyben arra voltunk kíváncsiak, hogy viszonylag rövid gyakorlás során milyen mértékben javul a szintetizált beszédformák észlelési teljesítménye. Ezt két okból tartottuk érdekes kérdésnek. Egyrészt, mert a beszédészlelés fejlődésével és atipikus vonatkozásaival kapcsolatos eddig bemutatott vizsgálatok nem foglalkoztak rövidtávú perceptuális tanulással, csak az életkori csoportok közti különbségekkel. Másrészt Nittrouer és munkatársai újabb vizsgálatai a diszlexia hátterére vonatkozóan fogalmaztak meg és támasztottak alá egy perceptuális integrációs hipotézist, mely a beszédhangok feldolgozását érinti (Nittrouer et al., 2009; Nittrouer & Lowenstein, 2010; Miyawaki, Strange, Verbrugge, Liberman, Jenkins, & Fujimura, 1975). Ezért felmerült bennünk, hogy a beszédészlelésben megnyilvánuló gyors tanulásnak a diagnosztika és a lehetséges fejlesztési eljárások szempontjából is lehet jelentősége.

³ Egy korábbi vizsgálatunkban ingerként használt mondatunk természetes kiejtésben, valamint szinusz hullámú és amplitúdómodulált változatban is meghallgatható a következő internetes oldalon: <http://barczy.elte.hu/content/kutatas.t.1250>

A kutatás kérdései

Az alábbi bemutatandó kutatás a következő két kérdést vizsgálja.

(1) *Kimutatható-e rövidtávú tanulási hatás a szinusz hullámú és amplitúdómodulált ingerek észlelésében tipikus mintán? Hogyan függ ez össze az életkorral?* Már említett korábbi vizsgálatunkban (Jakab és mtsai, 2018) ezzel kapcsolatban ellentmondó adatokat kaptunk, ezért a megismétlés mellett döntöttünk, kissé megváltoztatott vizsgálati tervvel.

(2) *Hogyan függ össze a két redukált beszédforma észlelése a nyelvi képességtesztek eredményeivel tipikus gyermekmintán?* Itt az elemzésben fordított utat követtünk, mint Nittrouer és Lowenstein (2013), akik a redukált beszéd észlelési eredménye mint független változó segítségével magyarázták bizonyos nyelvi képességtesztek eredményeinek varianciáját. Mi bevált képességteszteket használunk prediktorként, és ezek segítségével próbáljuk értelmezni a szinusz hullámú és amplitúdómodulált beszéd észlelésének hátterében álló folyamatokat.

Míg az első kérdés vizsgálatához mindhárom (óvodás, iskolás és felnőtt) csoportunk adataira támaszkodtunk, nyelvi képességteszteket csak az óvodás- és iskoláskorú gyermekekkel vettünk föl, így a második kérdéssel kapcsolatban e gyermekkorú kísérleti személyek adataira hagyatkoztunk. Mindkét kérdés megválaszolásához alapvetően korreláción, illetve regresszióan alapuló elemzési módszereket használtunk. A rövidtávú tanulás esetében arra voltunk kíváncsiak, hogy a kísérleti ülés során végzett gyakorlás mennyisége mint egyetlen független változó, mennyiben segít előrejelezni az észlelési teljesítményt. A második kérdés megválaszolásához többváltozós elemzéseket végeztünk, ezzel igyekeztünk tisztázni, hogy mely nyelvi részképességek játszanak elsősorban szerepet a redukált beszéd két vizsgált típusának észlelésében.

MÓDSZER

Személyek

Az itt leírt vizsgálatban 134 személy vett részt, három korcsoportból: óvodások (középső és nagycsoportosok), kisiskolások (1-3. osztály) és fiatal felnőttek.

Az óvodás csoport adatai a következők voltak: korátlag 77 hónap, szórás 7.1 hónap, minimum: 61, maximum: 90; 31 lány, 18 fiú.

Az iskolások adatai: korátlag: 100 hónap, szórás: 6.9 hónap, minimum: 87, maximum: 123; 40 lány, 21 fiú.

Végül a felnőtteké: átlag: 275 hónap, szórás: 36.9 hónap, minimum: 235, maximum: 390; 23 nő, 1 férfi.

Felnőtt személyeink valamennyien az ELTE BGGYK hallgatói voltak. Az óvodások és iskolások különböző fővárosi és vidéki intézményekbe jártak. Az óvodás és iskolás csoport életkori terjedelme kissé nagyobb a szokásosnál (1-3 év); ennek hatásait az elemzés során igyekeztünk felmérni.

Ingerek

A redukált beszéd észlelésének vizsgálata

Szintetizált beszédingereinkhez a mondatokat a Hearing in Noise Test gyermekváltozatának (HINT-C; Nilsson, Soli, & Gelnett, 1996; Nilsson, Soli, & Sullivan, 1994) 72 angol mondatából válogattuk, ezek mindegyike 5 szóból állt. Ugyanazt a 72 mondatot használtuk kiindulásként, mint egyik vizsgálatukban Nittrouer és Lowenstein (2010, p. 1628, 1633). A fordítás során nem törekedtünk a teljes tartalmi hűségre, arra azonban igen, hogy az általunk végül felhasznált mondatok mindegyike négy szóból álljon (ebbe a névelőket is beleszámoltuk), és szókincsét tekintve semmilyen nehézséget ne jelentsen egy tipikusan fejlődő 5-6 éves gyermek számára. A magyar változatok között 25 olyan mondatot találtunk, amelyek pontosan nyolc szótag hosszúak voltak; ezekből végül hármat használtunk a gyakorló próbákban, és húszat a kísérleti próbákban. A felhasznált magyar mondatok és angol eredetjük listáját a Függelék (A) tartalmazza.

Mondatainkról a hangfelvételeket az ELTE Bárczi Gusztáv Gyógypedagógiai Kar Ecseri úti hallásvizsgáló laborjában, teljesen hangszigetelt körülmények között készítettük. Ehhez egy Audio Technica AT2010-es kondenzátor mikrofont használtuk (Focusrite scarlett 2i4 hangkártyával), amely elé pop filtert helyeztünk a torzítások kiküszöbölésére. A hangrögzítést az Audacity program vezérelte. A nyers felvételek hangerejét a Praat programmal egységesen 70 decibelre (az amplitúdó négyzetes középértéke: 0.0632 Pascal) állítottuk be, és e változathoz szintetizáltuk a szinuszhullámú és amplitúdómodulált mondatainkat. A szinuszhullámú és amplitúdómodulált beszédet is e természetes kiejtés alapján készítettük. A szintézist a Praat rendszerre írt programokkal végeztük.⁴ E programok eredetijét Chris Darwin készítette⁵; a jelen kutatásban csak kisebb módosításokat végeztünk rajtuk: az amplitúdómodulált beszéd frekvenciasávjait, illetve a szinuszbeszéd formánsainak keverési paramétereit módosítottuk. A szinuszhullámú szintézis az első három formánst (F1, F2, F3) helyettesítette egy-egy frekvencia- és amplitúdómodulált szinuszhullámmal, és ezeket adta össze. Az amplitúdómodulált beszédingereknél négy frekvenciasávot használtunk (50-800 Hz; 800-1500 Hz; 1500-2500 Hz és 2500-6000 Hz), melyeken belül csak az amplitúdóváltozások őrződtek meg.

Nyelvi képességetesztek

A beszédészlelési vizsgálatba bevont gyermekcsoport egy részével a SZÓL-E tesztet (Kas, Lőrík, Bogáth, & Mályi, 2012) végeztük el. Ez egy 33 fős óvodás alcsoport volt (átlagéletkor: 79.8 hónap, szórás: 5.9 hónap, minimum: 67, maximum: 90; 20 lány; 13 fiú). E csoportból 16-an a szinuszhullámú, 17-en az amplitúdómodulált ingereket használó beszédészlelési

⁴ A Praat hang-, illetve beszédelemző számítógépes rendszer készítői Paul Boersma és David Weenink (Phonetic Sciences, University of Amsterdam, Spuistraat 210, 1012VT Amsterdam, Hollandia). A rendszer szabad felhasználású (<http://www.fon.hum.uva.nl/praat/>).

⁵ A szinuszhullámú beszédre használt eljárás forrása: http://www.lifesci.sussex.ac.uk/home/Chris_Darwin/Praatscripts/SWS; Az amplitúdómodulált beszéd szintézisé: http://www.lifesci.sussex.ac.uk/home/Chris_Darwin/Praatscripts/Shannon.

vizsgálatban vettek részt. A teljes minta egy másik, óvodásokból és iskolásokból álló része a beszédészlelési vizsgálat után a következő nyelvi tesztekkel töltötte ki: a LAPP aktív szókincs tesztet (Lőrinc és mts., 2015), a Magyar Mondat-utánmondási Teszt (MAMUT) rövid változatát (Kas és Lukács, 2011), a Magyar Álszóismétlési tesztet és a Számterjedelem tesztet (Racsmány, Lukács, Németh, & Pléh, 2005), Gósy Mária GMP tesztjének öt altesztjét (GMP2, 3, 4, 5, és 17: Mondatazonosítás zajban; Szóazonosítás zajban; Sávszűrt mondatok azonosítása; Gyorsított mondatok; A beszédhang-differenciálás vizsgálata), valamint a Woodcock-Johnson Kognitív Képességek tesztjének Fordított számterjedelem altesztjét (Rueff, Furmann, & Muñoz-Sandoval, 2003). E második alminta 76 fős volt (átlag: 94.4 hónap; szórás: 13.3 hónap; minimum: 61 hónap, maximum: 123 hónap; 50 lány, 26 fiú). A csoportból 38-an a szinusz hullámú, 38-an az amplitúdómodulált ingereket hallották.

Eljárás és kísérleti terv

A beszédészlelési vizsgálatot és a nyelvi tesztek két külön ülésben vettük föl a személyeinkkel. A redukált beszéd észlelését vizsgáltuk először. A redukált beszéd típusa csoportközi változó volt, minden gyerek vagy csak szinusz hullámú, vagy csak amplitúdómodulált beszédet hallott. Mindenki ugyanazt a 20 ingermondatot hallotta véletlenszerű sorrendben, vagy az egyik, vagy a másik szintézisben. A 20 kísérleti próba bemutatását három gyakorló próba előzte meg; ezek is ugyanazokat a mondatokat tartalmazták minden személy számára, és sorrendjük mindenkinél azonos volt (lásd az (A) függelék). A kísérleti eljárást az E-Prime kísérlettervező rendszerben írt forgatókönyv vezérelte, de az instrukciót minden személy számára a vizsgálatvezető mondta el. Minden személy egy AKG K-240 MK II típusú félig zárt fejhallgatón át hallotta az ingereket; a fejhallgatót viselve jól érthető volt a kísérletvezető beszéde. A fejhallgató hangerejét egy Voltcraft SL100 típusú hangnyomásmérővel 65 dB csúcsintenzitásra állítottuk be. A vizsgálat minden esetben megfelelően csendes helyiségben (pl. logopédiai, fejlesztő pedagógiai szobában) történt. Az instrukcióban arra készítettük föl a személyeket, hogy robothangszerű beszédet (a szinusz hullámú beszéd esetében), illetve szélsőségesre emlékeztető beszédet (az amplitúdómodulált beszéd esetében) fognak hallani, és arra kértük őket, hogy amit megértenek belőle, azt ismételjék meg. A bemelegítő próbákban a szintetizált mondat első lejátszása után is megkértük a személyt, hogy ismétlje meg, amit megértett, és ezt követően lejátszottuk a természetes kiejtést. Ennek megismétlésére is megkértük a személyt, majd a gyakorló próba utolsó részéként ismét a szintetizált változatot hallotta, amit másodsorra is elismételt. Ezzel az volt a célunk, hogy a személyek számára demonstráljuk a *Gyors perceptuális tanulás* c. részben említett aha-élményt, ezáltal motiválva őket arra, hogy a kísérleti próbákban, ahol a természetes kiejtést már nem hallották, szintén próbálkozzanak a visszamondással. A kísérleti próbák során minden ingermondatot kétszer játszottunk le, kivéve, ha a személy már az első hallgatás után tökéletesen elismételte azt. A próba eredménye a két hallgatás során pontosan elismételt szótagok kumulatív eredője volt; ha a második hallgatáskor a személy a hallott mondat olyan részeit is megismételte, amit az elsőben nem, akkor ezt hozzáadtuk az első ismétlés eredményéhez. A vizsgálat során a gyerekek egy társasjáték táblát is kaptak, melyen minden próba után léphettek egyet előre egy bábuval, s az utolsó próbát követően célba értek. A

vizsgálat után minden óvodás és iskolás személy egy plüss játékmackót is kapott ajándékba, így motivációs nehézség nem merült föl az adatfelvétel során.

A nyelvi képességeteszteket az előírt módon, külön ülésben vettük föl a gyerekekkel. Az adatgyűjtést a két vizsgálatvezető és négy szakdolgozó gyógypedagógus hallgató végezte.

Az adatok értékelése

Az eredményesség mérőszáma minden próbában az adott mondatból sikeresen elismételt szótagok százalékos aránya volt; mivel minden mondat 8 szótagból állt, ezért a próbák eredményei 0-tól 8-ig voltak skálázhatók. Ezeket az értékeket 8-cal osztva kaptunk relatív értékeket. Egyrészt minden személyre átlagoltuk a 20 próba eredményét, és ezzel jellemeztük teljesítményüket; ezek az adatok szolgálták a nyelvi képességetesztekkal való összefüggés vizsgálatára. Másrészt minden próbára vonatkozóan (1-től 20-ig) kiszámoltuk a hat alcsoport (3 korcsoport X 2 ingertípus) tagjainak átlagát, és ezeken az adatokon végeztük el a rövid távú tanulás vizsgálatát.

EREDMÉNYEK

A redukált beszédek észlelése

Eloszlási jellemzők

Mint korábbi vizsgálatunkban (Jakab és mtsai, 2018) is jeleztük, a redukált beszéd észlelésének teljesítménymutatói nem mindig normális eloszlásúak. A jelen tanulmány alapjául szolgáló adataink esetében a szinuszhullámú beszéd észlelésének teljesítményadatai mindhárom vizsgált korcsoportban eltértek a normálistól, az amplitúdómodulált beszéd észlelési adatai azonban egyikben sem. A szinuszhullámú beszéd észleléséből származó adatok eloszlását az 1. táblázat foglalja össze.

	ferdeség		csúcosság		Kolmogorov-Szmirnov	Shapiro-Wilk
	érték	sztdev. hiba	érték	sztdev. hiba		
Óvodások	-1.014	0.472	-0.518	0.918	$p < 0.001$	$p < 0.001$
Iskolások	-2.191	0.421	3.627	0.821	$p < 0.001$	$p < 0.001$
Felnőttek	-1.981	0.637	4.22	1.232	$p < 0.01$	$p < 0.01$

1. táblázat. A szinuszhullámú beszéd teljesítménymutatójának eloszlási jellemzői.

Mint a táblázatból látható, mindhárom korcsoportban határozott negatív ferdeség jellemzi a szinuszhullámú beszéd teljesítménymutatójának eloszlását, valamint a két idősebb csoportban csúcosság is. Felnőtteknél igen erős plafonhatás volt megfigyelhető: egyetlen személy kivételével a teljesítményadatok 0.94 és 1 között voltak. Az óvodás csoportban bimodális volt az eloszlás: a gyerekek mintegy harmada a 0-0.4-es teljesítménytartományban, kétharmaduk pedig a 0.6 és 1 közötti tartományban helyezkedett el. Az iskolások esetében három eloszlási

csúcspont volt megfigyelhető, bár ebből egy erősen dominált: a személyek túlnyomó többsége a 0.8-1 tartományban helyezkedett el, hárman a 0-0.1 tartományban, ketten pedig a 0.4-0.6 közötti tartományban. Az amplitúdómodulált beszéd teljesítményadatai egyik csoportban sem tértek el a normális eloszlástól, egyik fenti kritérium szerint sem. Az eloszlások különbsége alapján úgy tűnik, hogy a kétféle redukált beszéd megértése, illetve annak javulása eltérő módon történhet, legalábbis annyiban, hogy a szinuszhullámú ingerek esetében inkább egy „minden vagy semmi” hatás érvényesül, míg az amplitúdómodulált ingereknél fokozatosság figyelhető meg; a felismerési hányad mutató jobban differenciálja egymástól a személyeket.

Teljesítmény az alcsoportokban és a nemek összevetése

A személyek teljesítményét ingerek szerint átlagolva kapjuk meg az egyes alcsoportok, illetve a teljes minta átlagteljesítményét. Minden személy teljesítményét a húsz mondat felismerési arányának átlaga adta. A teljesítményátlagokat a 2. táblázat foglalja össze.

			Korcsoport (jelen adatok)			Korcsoport (2018)	
			Óvodások	Iskolások	Felnőttek	Óvodások	Felnőttek
Ingertípus	SWS	átlag	0.623	0.805	0.987	0.504	0.788
		szórás	0.343	0.285	0.021	0.135	0.104
	VOC	átlag	0.250	0.487	0.745	0.558	0.919
		szórás	0.162	0.131	0.100	0.213	0.059

2. táblázat. A szintetizált beszéd észlelésének teljesítményátlagai a jelen vizsgálatban, illetve egy korábbi vizsgálatban. (SWS: szinuszhullámú ingerek; VOC: amplitúdómodulált ingerek.)

Az adatok elemzése előtt érdemes egy gyors összehasonlítást végezni korábbi eredményeinkkel (Jakab és mtsai, 2018), ebből a célból a 2. táblázatban megjelenítjük az akkori teljesítményadatokat is. A két vizsgálat óvodás és felnőtt csoportját csak leíró statisztikai mutatóik alapján vehetjük össze – mivel a kísérleti tervek, az ingermondatok és az ingerbeállítási paraméterek is különböztek a két esetben, statisztikai próbát nem lehetett végezni az összehasonlításukra. Így az alábbi néhány megfigyelést, illetve magyarázatot ennek fényében, kellő óvatossággal lehet csak felhasználni. Az óvodások teljesítménye a szinuszhullámú mondatokra most valamivel magasabb volt, mint korábban. Ennek az lehet az oka, hogy míg most egységesen nyolc szótagot tartalmazó, négy szóból álló mondatokat használtunk, addig a korábbi vizsgálat mondatai átlagosan csaknem kétszer olyan hosszúak voltak (10-18 szótag; átlag: 14.6). Ugyanakkor az amplitúdómodulált ingerek esetében a két óvodás csoport viszonya fordított: a múltkori vizsgálatban jóval magasabb teljesítmények adódtak, mint most. Ennek viszont az a legvalószínűbb oka, hogy korábbi vizsgálatunkban hat független frekvenciasávot tartottunk meg a természetes beszéd spektrogramjából, most viszont csak négyet (mivel a hatsávosságú ingereket túl könnyűnek ítéltük). A felnőttek teljesítménye is tükrözi ezt a két hatást. A mostani egyszerűbb mondatok és szinuszhullámú ingerek esetében plafonhatás mutatkozott, míg korábban az azonos módon szintetizált, de hosszabb mondatok esetében nem. Korábbi vizsgálatunkban a felnőttek éppen az amplitúdómodulált mondatok megértésében mutattak plafonhatást, vélhetőleg a valamivel kisebb mértékű információ-redukció hatásának köszönhetően.

Teljesítmény tekintetében a jelen mintában összevetettük a nemeket is; sem a teljes mintában, sem az egyes ingertípusok esetében nem találtunk szignifikáns különbséget, és a négy gyerekcsoport egyikében sem tért el egymástól a két nem teljesítménye (Welch próba). (A két felnőtt csoportban összesen egy férfi volt). Így a továbbiakban a nemmel mint független változóval nem foglalkozunk.

A korcsoport és az ingertípus hatása a teljesítményre

A 2. táblázatban bemutatott adatok alaposabb vizsgálata a fent említett normalitásvizsgálat miatt trimmelt varianciaanalízissel és sztochasztikus homogenitásvizsgálattal (rang-varianciaanalízissel; Vargha, 2020) történt. A korcsoport és az ingertípus hatását kétszemponos trimmelt varianciaanalízissel vizsgáltuk meg (10%-os trimmelést alkalmazva). A korcsoport szerinti főhatás szignifikáns volt (Welch próba: $F(2, 44.5)=49.001$; $p<0.001$; $\eta^2=0.330$); az ingertípus szerinti szintén (Welch próba: $F(1,47.0)=92.091$, $p<0.001$; $\eta^2=0.303$). Az interakció szintén szignifikánsnak adódott, bár mérete sokkal kisebb volt a főhatásokénál (Johansen próba: $\chi^2(2)=6.341$, $p<0.05$; $\eta^2=0.011$). A páronkénti összehasonlítás eredményei (Games-Howell próba) szerint szinuszhullámú ingerek esetén az óvodások és iskolások nem különböztek szignifikánsan ($T12(3; 44)=2,98$, NS), de az iskolások és felnőttek igen ($T23(3; 31)=4,99$, $p<0.01$); az óvodások és a felnőttek szintén ($T13(3; 23)=7,33$, $p<0.01$). Amplitúdómodulált ingerek esetén mindhárom páros összehasonlítás szignifikáns eredményt adott (óvodások vs. iskolások: $T12(3; 37)=8,54$, $p<0.01$; iskolások vs. felnőttek: $T23(3; 21)=9,75$, $p<0.01$; óvodások vs. felnőttek: $T13(3; 26)=16,16$, $p<0.01$). Az eredményeket sztochasztikus homogenitásvizsgálattal ellenőrizve a két főhatás szignifikánsnak adódott, az interakció azonban nem (korcsoport: $F(1.4, 128)=152.043$, $p<0.001$; ingertípus: $F(1.0, 128)=151.844$, $p<0.001$; interakció: $F(1.1, 128)=0.945$, NS). Az értelmezéshez a Megvitatásban térünk vissza.

Az eredmények elemzése során még egy szempontot kellett figyelembe vennünk. Gyermek korcsoportjaink meglehetősen tág életkori tartományokat ölelnek fel, így pl. az óvodás csoportok legfiatalabb és legidősebb tagja között két-két és fél év a korkülönbség, míg az egyik iskolás alcsoport életkori terjedelme a 3 évet is eléri (3. táblázat). Ilyen széles tartományon belül már az életkor is összefügghet a teljesítménnyel, tehát felmerülhet, hogy a csoportok ebből a szempontból nem homogének. Ennek vizsgálatára mind a hat alcsoportban megvizsgáltuk a korrelációt az életkor és a teljesítmény között. Az összefüggés egyik alcsoport esetében sem lett szignifikáns (3. táblázat), így életkori hatások esetünkben csak a korcsoportok közti különbségben érhetőek tetten.

Csoport	Korátlag	szórás	minimum	maximum	r(kor, telj.)	df
ovi SWS	75.792	8.140	61	90	-0.122	22
isk SWS	100.194	6.123	88	111	0.007	29
feln SWS	266.500	30.655	235	342	0.104	10
ovi VOC	78.200	5.881	67	88	0.219	23
isk VOC	99.767	7.740	87	123	0.210	28
feln VOC	283.333	41.940	237	390	-0.086	10

3. táblázat. Életkor és teljesítmény közötti korrelációk az alcsoportokban (SWS: szinuszhullámú beszédet hallgató alcsoport; VOC: amplitúdómodulált beszédet hallgató alcsoport).

Az itt bemutatott vizsgálat legérdekesebb eredménye egy határozott csoportszintű teljesítményjavulás a redukált beszédingerek észlelésében az egymást követő kísérleti próbák során, mindkét ingertípus, illetve a hat alcsoport közül öt esetében. A kivételt a felnőtt szinusz csoport jelenti, ahol a plafonhatás miatt nem volt lehetőség a teljesítmény növekedésére. Az egyes alcsoportok személyeinek teljesítményét próbánként átlagolva, a próba sorszáma és a személyek átlagos teljesítménye között korrelációt kapunk, vagyis a teljesítmény az egymást követő próbák hatására javul. Véleményünk szerint ezt legalább részben perceptuális tanulás lehet megmagyarázni; ehhez a Megvitatás részben térünk vissza. Az egyes személyek adatait már jobban terheli a különböző okokra visszavezethető hiba. Először a csoportszintű eredményeket mutatjuk be. A 4. táblázat a releváns korrelációs adatokat foglalja össze.

		CI95			df = 18
		együttható	alsó	felső	p
SWS óvodások	Pearson	0.843	0.638	0.936	<0.001
	Wilcox r_{pb}	0.879	0.731	0.948	<0.001
	Kendall	0.731			<0.001
SWS iskolások	Pearson	0.729	0.423	0.886	<0.001
	Wilcox r_{pb}	0.728	0.452	0.877	<0.001
	Kendall	0.519			<0.002
SWS felnőttek	Pearson	0.279	-0.186	0.642	>0.1
	Wilcox r_{pb}	0.127	-0.301	0.512	>0.5
	Kendall	0.147			>0.4
VOC óvodások	Pearson	0.758	0.474	0.899	<0.001
	Wilcox r_{pb}	0.775	0.533	0.900	<0.001
	Kendall	0.565			<0.001
VOC iskolások	Pearson	0.747	0.455	0.894	<0.001
	Wilcox r_{pb}	0.745	0.481	0.886	<0.001
	Kendall	0.578			<0.001
VOC felnőttek	Pearson	0.618	0.242	0.833	<0.01
	Wilcox r_{pb}	0.605	0.257	0.814	<0.01
	Kendall	0.392			<0.02

4. táblázat. Korrelációk a próbaszám és a teljesítmény között az alcsoportokban.

Két megjegyzést kell tennünk a 4. táblázat értelmezése előtt. Először, mivel a szinusz hullámú beszéd teljesítményadatai nem bizonyultak normális eloszlásúnak, ezért Wilcox-féle robusztus korrelációs együtthatót is számoltunk (az esetleges kilógó adatpárok hatásának kiszűrésére), valamint Kendall-féle rangkorrelációs együtthatót is. Mint a 4. táblázatból látható, a Wilcox-féle együttható a mi adatainkra csak minimális mértékben tér el a Pearson-félétől, a

rangkorreláció ellenben kicsit jobban. Ezért további korrelációs táblázatainkban már csak a Pearson együtthatót és a Kendall-féle tau értéket adjuk meg. Ezen eredmények alapján annyit biztosan állíthatunk, hogy a próbaszám (vagyis a rövid távú gyakorlás mennyisége, az egymást követő „gyakorlási egységek” száma) és a teljesítmény közötti összefüggésnek van egy lineáris korrelációval, illetve monotonitási mutatóval jól megragadható komponense. Második megjegyzésünk, hogy a próbaszámot is tekinthetjük arányszámnak (és nem csak ordinálisnak), hiszen minden próbában adunk egy „egységnyi” feladatot a személyeknek, és feltehetően a teljesítményük ennek hatására emelkedik (csoportszinten) – tehát a teljesítmény változása esetünkben a rövidtávú gyakorlás mennyiségével függ össze.

A tanulási hatás eddigi változóorientált elemzése után, Vargha (2016) javaslatát követve, tehetünk egy gyors kitérőt az individuumentált elemzés irányába is. Az 5. táblázatban az látható, hogy az egyes alcsoportokban az egyéni teljesítmények próbaszámmal való korrelációi milyen eloszlást követnek. Ha feltesszük, hogy pl. egy húsz fős alcsoportban a teljesítmény független a próbaszámtól, akkor a húsz egyéni korrelációs mutató közül várhatóan egy lenne szignifikáns (5%) és az együtthatók mintegy fele, tehát 10 körüli lenne pozitív. Látható azonban, hogy minden alcsoportban, ahol a csoportátlag szintjén megmutatkozik a próbaszám-teljesítmény korreláció, ott az egyéni korrelációs együtthatóknak jóval több mint 5%-a szignifikáns és több mint 50%-a pozitív. Az 5. táblázat tartalmazza az erre a két arányra vonatkozó binomiális próbák elsőfajú hibáit is.

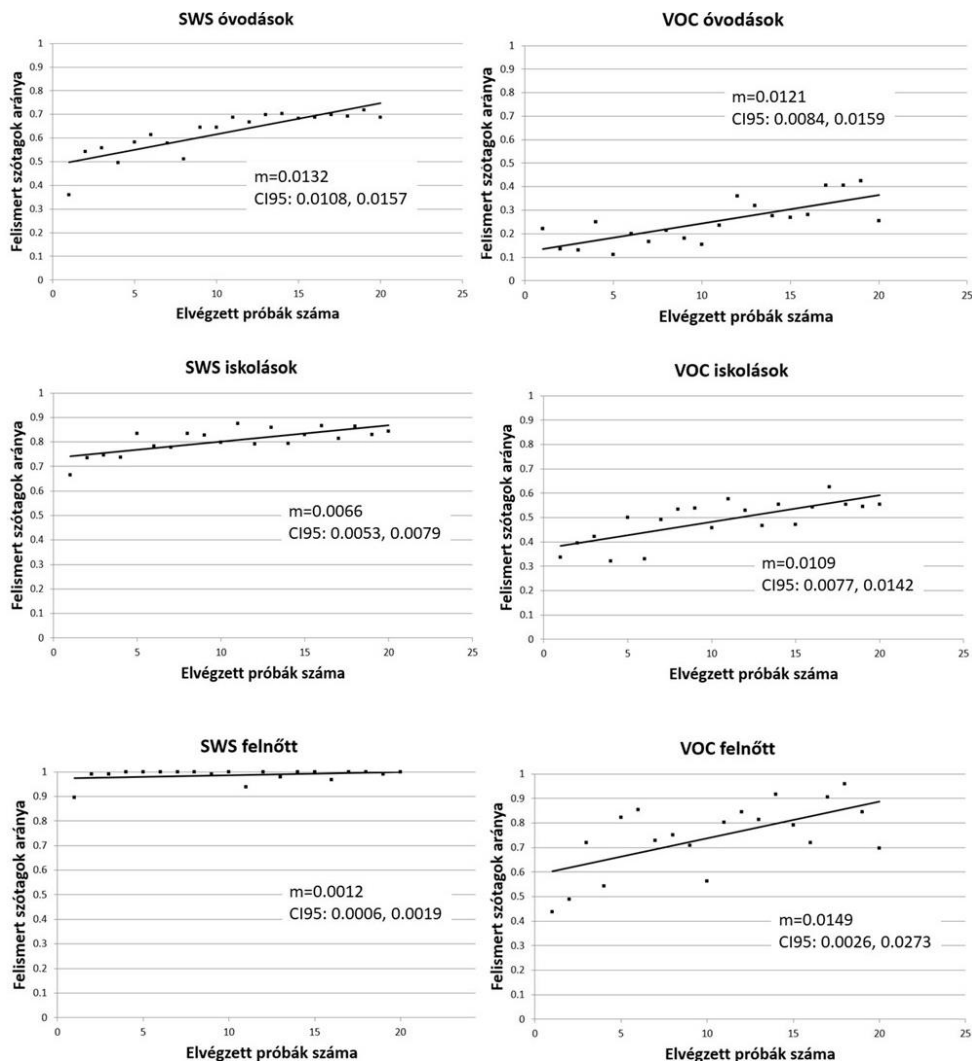
	SWS			VOC		VOC
	óvodás	SWS iskolás	SWS felnőtt	óvodás	VOC iskolás	felnőtt
Szignifikáns	28.6% (6/21)	9.7% (3/31)	0% (0/5)	25% (6/24)	26.7% (8/30)	16.7% (2/12)
Pozitív	81% (17/21)	67.7% (21/31)	40% (2/5)	66.7% (16/24)	70% (21/30)	83.3% (10/12)
Binom.próba (szign.) p érték	0.0004***	0.2008	1	0.001**	1E-04***	0.1184
Binom.próba (pozitív) p érték	0.0501+	0.0708+	1	0.1516	0.0161*	0.0396***

Szignifikanciaszintek: + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

5. táblázat. Az egyéni teljesítménymutatók összefüggése a próbasorszámmal. A felső két sor a megfelelő alcsoportban kapott összes egyéni próbaszám-teljesítmény korreláció közül a szignifikánsak, illetve a pozitívak százalékos arányát (zárójelben a tényleges arányszámokat) mutatja. Az SWS felnőtt csoportban a teljes csoportlétszám azért 5, mert a 12-ből 7 személy valamennyi értéke 1 volt, így korrelációt csak a maradék 5 személyre lehetett számolni. Az alsó két sor ennek a két mutatónak a nulla csoportszintű elméleti korreláció esetén várható aránytól (5% szignifikáns egyéni korreláció; 50% nullánál nagyobb korreláció) való eltérésére vonatkozó binomiális próbák elsőfajú hibáit tartalmazza.

Az 5. táblázatban látható, hogy ahol csoportszintű korreláció van jelen, ott a két binomiális próba legalább egyike legalább tendenciaszinten eltérést jelez. Összefoglalva tehát, a csoportszintű összefüggés nyomai, hibákkal terheltlen bár, de az egyéni próbaszám-teljesítmény korrelációkban is megmutatkoznak.

Végül vessünk egy pillantást a lineáris regresszió eredményeire is (1. ábra).



1. ábra. A teljesítmény változásának regressziós egyenesei a hat alcsoportban. SWS: szinuszhullámú ingerek; VOC: amplitúdómodulált ingerek. A grafikonokon a regressziós egyenes meredeksége (m), és annak 95%-os konfidenciaintervalluma látható.

Az 1. ábra regressziós egyenesein is látható, hogy a hat alcsoportból ötben a teljesítmény javul az egymást követő próbák során, legalábbis csoportszinten. Az első ránézésre alacsony meredekségek valójában jól értelmezhetők. Az óvodás szinuszhullámú csoportban például a 0.0132 meredekségi érték azt jelenti, hogy egy próba hatására várhatóan 1.3 százalékkal emelkedik a teljesítmény. Hús próba esetén azonban a várható emelkedés $20 \cdot 1.3 = 26$ százalékkal, ami már elméleti szempontból is jelentősnek tekinthető, különösen, ha figyelembe vesszük a rövid gyakorlási időt.

Nyelvi képességek

A SZÓL-E teszt és a redukált beszéd észlelése közötti összefüggés

A SZÓL-E eredmények alapján valamennyi, a vizsgálatban résztvevő óvodás tipikus nyelvfejlődésének tekinthető; szókincsük, mondatfűzésük, beszédfejlődési adataik a megfelelő szinten vannak. A teszt alskálái közül az *Artikuláció* nominális változóként nem volt bevonható korrelációs vizsgálatba. Ugyanakkor a vizsgálatvezető a 33 gyermekből tizenöténél tett észrevételeket az artikulációval kapcsolatban. Ennek alapján (ti. van-e kisebb artikulációs probléma vagy nincs) a mintát két részre osztottuk, és a szintetizált beszéd észlelési teljesítményét kétszemponos varianciaanalízisben hasonlítottuk össze, ahol a másik független változó az ingertípus (SWS ill. VOC) volt. Az eddigiekkel összhangban egyedül az ingertípus szerinti főhatás lett szignifikáns ($F(1,29)=10.306$, $p<0.01$, $\eta^2=0.262$), így az *Artikuláció* alskálát tovább nem elemeztük. Szintén kihagytuk az alábbi korrelációs táblázatból (6. táblázat) a *Nyelvtani morféimák megértése* alskálát, mivel ezen igen erős plafonhatás mutatkozott: a 33 személyből 30 a maximális 12 pontot érte el. Valamivel enyhébb plafonhatás volt megfigyelhető a *Gyors megnevezés – pontosság* skálán is, ahol a 33-ból 21 személy (64%) érte el a maximális 36 pontot, ennek ellenére ezt a skálát szerepeltettük a 6. táblázatban. Ezek alapján a SZÓL-E alskáláknak a redukált beszéd észlelésével való korrelációi a következőképpen alakultak.

Pearson r	SZE	ÁM	FE	MP	ÁI	MIT	MISZ	HT	GyMNp	GyMNI
SWS	0.394	0.024	-0.027	0.517*	-0.027	0.327	0.247	0.433	0.210	-0.064
VOC	0.312	0.372	0.110	0.417	0.332	0.521*	0.165	0.379	0.209	-0.207
Kendall tau										
SWS	0.270	-0.009	0.010	0.450*	0.066	0.355	0.199	0.267	0.190	0.060
VOC	0.152	0.245	0.131	0.317	0.289	0.350	0.097	0.298	0.259	-0.037

Szignifikancia: * $p<0.05$

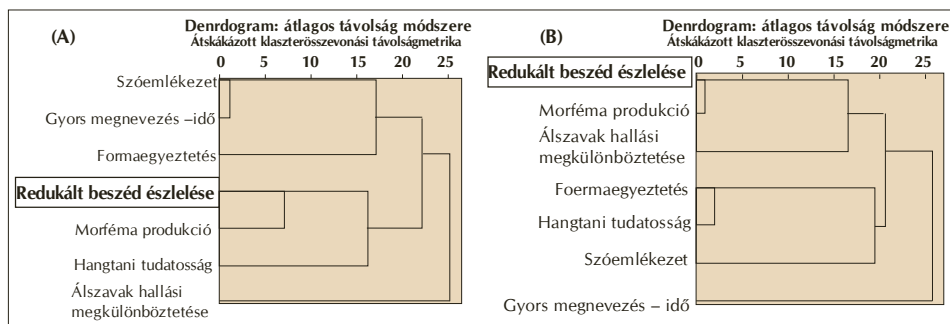
SZE	Szóemlékezet (15)
ÁM	Álszavak hallási megkülönböztetése (10)
FE	Formaegeztetés (5)
MP	Nyelvtani morféimák - produkció (12)
ÁI	Álszóismétlés (36)
MIT	Mondatismétlés - tartalom (6)
MISZ	Mondatismétlés - szerkezet (6)
HT	Hangtani tudatosság (8)
GyMNp	Gyors megnevezés - pontosság (36)
GyMNI	Gyors megnevezés - idő/mp

6. táblázat. A SZÓL-E teszt vizsgálatba bevont alskáláinak korrelációi a szintetizált beszédformák észlelési teljesítményével. Alul zárójelben az alskálákon elérhető maximális pontszámok.

Különösebben nem meglepő módon, egyetlen SZÓL-E alskála sem mutat kiugró korrelációt a redukált beszédformák észlelésével, hiszen utóbbi feladat számos nyelvi részképességet dolgoztat meg egyszerre, míg a SZÓL-E teszt e részképességeket egymástól elkülönítve igyekszik vizsgálni. Így adataink alaposabb elemzéséhez szükséges összetettebb többváltozós módszereket használni, ebben viszont némi korlátot jelent ennek az almintának az alacsony

elemszáma. Többszörös regressziós elemzéssel (ahol a redukált beszéd észlelése a függő változó, s a SZÓL-E alkálák a független változók) ezen okból nem próbálkoztunk. Ehelyett hierarchikus klaszteranalízist végeztünk többféle módszerrel, és ezek eredményeiből próbáltunk néhány használható következtetést levonni.

A hierarchikus klaszteranalízis elvégzéséhez azonban ingertípus szerint ketté kellett osztanunk a 33 fős mintát, és ezen a két alcsoporton elvégezni az elemzést. (Az eddig leírtak alapján nem volt ugyanis okunk feltételezni, hogy a szinuszhullámú és amplitúdómodulált beszéd észlelése azonos módon viszonyul a nyelvi részképességekhez.) Így a további vizsgálatokra egy 16 fős szinuszhullámú és egy 17 fős amplitúdómodulált minta áll rendelkezésünkre. Az elemzés során tehát változókat klasztereztünk. Először a Pearson korrelációs metrikát alkalmaztuk (abszolútérték kikötéssel; I. Vargha, 2019, 7. fejezet). Klaszterösszevonásra az átlagos távolság módszerét használtuk minden alább részletezett esetben. Pearson metrika mellett csak olyan változókat tudtunk bevonni, melyek eloszlása nem nagyon ferde, és nem nagyon alacsony a szórása. A redukált beszéd észlelési teljesítménye ezt mindkét ingertípusra teljesítette.⁶ A SZÓL-E alkálák esetében azonban többen a plafonhatás miatt erős negatív ferdeség volt megfigyelhető. Így itt az alkálák közül csak hat volt alkalmas erre az elemzésre, ezek a következők voltak: *álszavak hallási megkülönböztetése, gyors megnevezés ideje, szóemlékezet, nyelvtani morfémák produkciója, formaegyeztetés, hangtani tudatosság*. Ezek közül azonban a *formaegyeztetés* a vizuális észlelést mérő feladat, és mint várható, teljes mértékben független is mindkét fajta beszédingerünk észlelésétől (6. táblázat). Így ezt az alkálát az összes alábbi elemzésből kihagytuk, és a Pearson metrikán alapuló klaszteranalízisbe csak 5 SZÓL-E változót vontunk be. A 2. ábra ennek az elemzésnek a dendrogramjait mutatja be a két ingertípus szerinti alcsoportra.



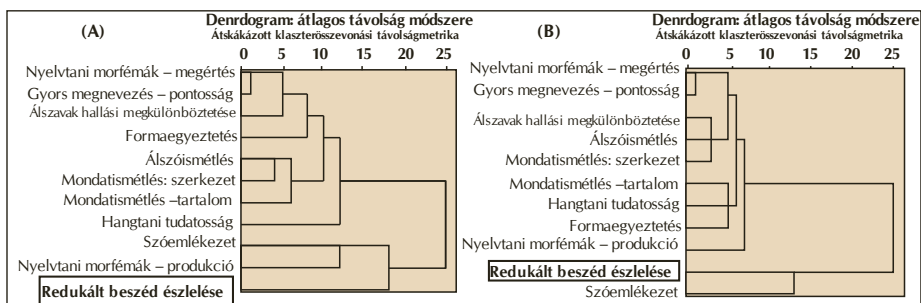
2. ábra. A beszédészlelési teljesítmény összefüggése a SZÓL-E teszt alkáláival. (A): szinuszhullámú beszéd, (B): amplitúdómodulált beszéd. Pearson abszolútérték metrika; átlagos távolság módszere.

A 2. ábrán számunkra most csak a redukált beszéd észlelési teljesítménye és a SZÓL-E alkálái közti viszony az érdekes. Ezzel kapcsolatban a következőket láthatjuk. A szinuszhullámú

⁶ Bár a szinuszhullámú alcsoportban eltért a teljesítmény eloszlása a normálistól, ennek oka bimodalitás volt; a ferdeség ebben az alcsoportban nem tért el szignifikánsan a nullától. Bár mint a 4.1.1. részben és az 1. táblázatban olvasható, a teljes óvodás és iskolás korcsoportban szignifikáns negatív ferdeséget találtunk erre a változóra, azonban a SZÓL-E tesztet az óvodás és iskolás korcsoport egy-egy részével vettük föl, és ebben az alcsoportban nem volt szignifikáns a negatív ferdeség.

beszéd észlelése adataink szerint a nyelvtani morfémák produkciójával és a hangtani tudatossággal áll a legszorosabb kapcsolatban. Az amplitúdómodulált beszéd észleléséhez a morfémaprodukción kívül az álszavak hallási megkülönböztetése áll a legközelebb, amely szintén egy döntően fonológiai feldolgozáson alapuló feladat.⁷

Részen a fentiek megerősítése, részben pedig az eddig kimaradt változók megvizsgálása céljából a klaszteranalízist négyzetes euklidészi távolságmérika segítségével is elvégeztük. E célból a SZÓL-E alskálák mindegyikét 1-re normáltuk (a személyek értékeit az elérhető maximális pontszámmal osztottuk) kivéve a gyors megnevezési időt, melynek nem volt maximuma. Így utóbbi változó (és a formaegyeztetés) kivételével valamennyi alskálát bevontuk a második elemzésbe. A 3. ábra mutatja a megfelelő dendrogramokat.



3. ábra. A beszédészlelési teljesítmény és a SZÓL-E skálák összefüggésének ábrázolása: dendrogram négyzetes euklidészi távolság metrika alapján. (A): szinuszhullámú beszéd, (B): amplitúdómodulált beszéd.

A 3(A) ábra alapján a szinuszhullámú beszéd észlelése a nyelvtani morfémák produkciójával és a szóemlékezettel áll a legszorosabb kapcsolatban; a hangtani tudatosság kissé messzebb került a kapcsolatrendszerben. Bár nem sokkal távolabb: a kilenc klaszterösszevonási lépésből a hetedikig a hangtani tudatosság önmagában állt, és csak ekkor olvadt össze a dendrogramon fölülte elhelyezkedő SZÓL-E változócsoporthal. A 3(B) ábra szerint hasonló figyelhető meg a nyelvi morfémák produkciója és az amplitúdómodulált beszéd észlelése esetében: a tíz összevonási lépésből az első hatban a morfémaprodukciónak egymaga alkotott egy klasztert, míg a hetedik lépésben csatlakozott az alskálák egy nagy csoportjához, melyből pont a szóemlékezet és a beszédészlelés maradt ki. Egészében a Pearson-féle metrika szorosabb kapcsolatot jelez a morfémaprodukción és a redukált beszéd észlelése között, mint az euklidészi, és utóbbi alapján a mi beszédészlelési feladatunk csak meglehetősen lazán kapcsolódik a SZÓL-E teszthez. Mindezek ellenére a morfémaprodukción a választott távolságmétrikától függetlenül váratlanul konzisztens módon függ össze mindkétféle redukált beszéd észleléssel. E következtetések megszilárdításához azonban még szükség lenne egy hasonló vizsgálat nagyobb mintán való elvégzésére.

⁷ A szinuszhullámú esetben a morfémaprodukción a második összevonási szinten, a hangtani tudatosság pedig a harmadikon olvadt össze a redukált beszéd észleléssel. Amplitúdómodulált ingereknél a morfémaprodukción az első összevonási szinten, az álszó megkülönböztetés pedig a harmadikon csatlakozott a beszédészlelési feladat teljesítményéhez.

A redukált beszéd észlelésének összefüggése további nyelvi tesztekkel

Ezekben a mérésekben egy 76 fős al minta tagjai vettek részt, óvodások és iskolások vegyesen. A résztvevők többsége életkorának megfelelően teljesített a vizsgálatokban, de voltak az elvárható szint alatti teljesítmények is. Ilyenek elsősorban az iskolás korosztályban a Magyar Mondatutánmondás (MAMUT-R) tesztben, a Számterjedelem tesztben és kisebb arányban a LAPP szókinccstesztben fordultak elő. Átlag alatti teljesítmény miatt nem zártunk ki személyeket ebből a vizsgálatból.

A redukált beszéd észlelése és a használt nyelvi képességtesztek közti korrelációkat a 7. táblázat összegzi.

Pearson r	LAPP T	MAMUT R	GMP17	GMP2	GMP3	GMP4	GMP5	Álszóism.	Számterj.	Ford. számt.
SWS	0.360*	0.563***	0.294	0.158	0.178	0.380*	0.334*	0.283	0.242	0.388*
VOC	0.449**	0.305	0.533***	0.448**	0.513**	0.046	0.381*	0.124	0.402*	0.221
Kendall tau										
SWS	0.214	0.321*	0.254*	0.138	0.137	0.216	0.239*	0.320*	0.163	0.252
VOC	0.390**	0.198	0.331**	0.332**	0.398**	0.023	0.293*	0.117	0.317*	0.161

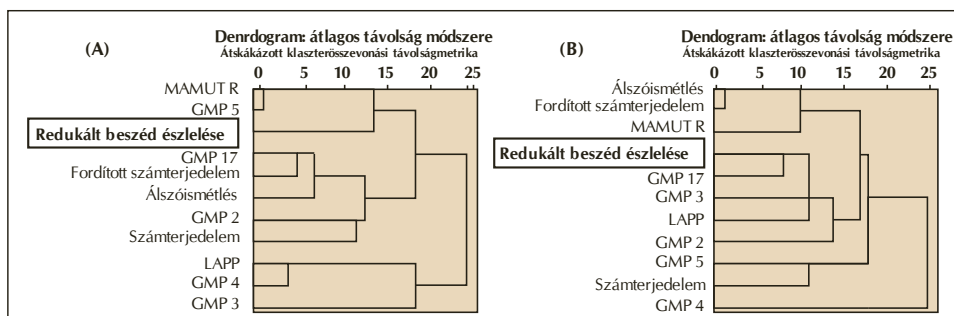
* p<0.05, **p<0.01, ***p<0.001

7. táblázat. Nyelvi képességtesztek és a redukált beszéd észlelése közötti korrelációk a redukáltbeszéd-észlelési vizsgálat ingertípusa szerint csoportosítva.

Ahogy a SZÓL-E teszt esetében, úgy itt sem kaptunk kiugróan magas korrelációt egyetlen nyelvi képességteszttel sem, annyi azonban látható, hogy van bizonyos összefüggés a nyelvi részképességek és a redukált beszédformák észlelése között. Sőt, némi különbség is látható a két ingertípusnak a nyelvi képességek rendszerébe való beilleszkedése tekintetében. Például a mondat- és szóazonosítás zajban (GMP2 és GMP3) feladatok szorosabban látszanak kapcsolódni a szélsusugásra, suttogásra emlékeztető amplitúdómodulált beszéd észlelésével, mint a szinusz hullámúéval; ehhez a Megvitatás részben még visszatérünk. Hasonló korrelációs kontraszt figyelhető meg a GMP17 esetén is (Beszédhang-differenciálás), és ennek értelmezése nem is tűnik egyszerű kérdésnek. A szókinccs (LAPP) és a gyorsított mondatok észlelése (GMP5) ellenben nagyjából egyformán függ össze az általunk használt két ingertípus észlelésével.

Itt is fontosnak tűnik tehát a továbblépés az összetettebb elemzések irányába. A hierarchikus klaszteranalízis mellett itt már többszörös lineáris regressziós (TLR) elemzést is végeztünk, mivel a mintaelemszám minimálisan elégségesnek tűnt háromnál nem több prediktort alkalmazó TLR modellek vizsgálatára.⁸ A hierarchikus klaszteranalízist most csak Pearson abszolútérték metrikával végeztük el, ennek eredményét mutatja a 4. ábra.

⁸ A Green (1991) javasolta kritérium szerint $n=50+8k$ elemszám már elegendő lehet (ahol k a független változók száma), regressziós modellek illeszkedésének vizsgálatára. A mi 76-os elemszámunk esetében ez éppen teljesül $k=3$ -ra.



4. ábra. A redukált beszédformák észlelése és nyelvi képességetesztek közti összefüggés hierarchikus klaszteranalízis (Pearson abszolútérték metrika) alapján. (A): szinusz hullámú beszéd, (B): amplitúdómodulált beszéd.

A 4 (A) ábra szerint a szinusz hullámú beszéd észlelése a gyorsított mondatok észleléséhez és a MAMUT-R mondatismétlési feladatahoz áll a legközelebb. Ez nem túl meglepő, hiszen három mondatismétlési feladatról van szó, különböző feltételek mellett (A GMP5-ben is a hallott gyorsított mondatok elismétlése a feladat). A többi nyelvi teszt egymás közötti viszonyait most nem elemezzük, de a köztük lévő korrelációs mátrixot közöljük a (B) függelékben. Ott megfigyelhető, hogy a MAMUT-R teljesítménye elég határozottan korrelál lényegében az összes többi nyelvi képességeteszttel.

A 4 (B) ábrán látható, hogy az amplitúdómodulált beszéd észlelése a GMP17 szubteszttel (beszédhang-differenciálás) függ össze leginkább – itt visszaköszön a 7. táblázat 3. oszlopában látható határozott VOC-GMP17 korreláció. Emellett a GMP3 és a LAPP kötődnek szorosan az AM beszéd észleléséhez (vö. a 7. táblázatban található korrelációkkal). A jó szókinca a lexikai restaurációs folyamatokat támogathatja hatékonyan; a zajban való szóazonosítás szerepéhez a Megvitatás részben még visszatérünk.

A többszörös regressziós elemzések során több egyszerű, három független változót tartalmazó modellt is kipróbáltunk. A függő változó minden esetben a redukált beszéd észlelése volt; független változóink közül egy az ingertípus (SWS vagy VOC), kettő pedig a nyelvi képességetesztek változói közül került ki. A nyelvi képességváltozókat az egyes – alább kipróbált – modellekben a 7. táblázat korrelációi alapján válogattuk ki. Három ilyen egyszerű modellt mutatunk itt be, mivel ezek elég jól lefedik a rendszerünkben lévő legfontosabb információkat. A 8. táblázat foglalja össze az első modellt.

1. modell	Együtthatók						Korrigált R ²	elsőfajú hiba
	B	sztdev hiba	CI95 alsó	CI95 felső	béta	elsőfajú hiba		
Ingertípus	0.362	0.048	0.266	0.458	0.645	p<0.001	0.359	p<0.001
MAMUT R	0.006	0.002	0.002	0.010	0.304	p<0.005	0.128	p<0.001
GMP5	0.002	0.002	-0.001	0.005	0.129	NS	0.005	NS

8. táblázat. A redukált beszéd észlelésének első kipróbált regressziós modellje. Ingertípus: SWS, illetve VOC. Az Együtthatók részben található elsőfajú hiba értékek a B együtthatók meredekségének nullától való eltéréseire vonatkoznak, a jobb szélső oszlopban lévő értékek pedig a korrigált R négyzet növekményére, tehát arra, hogy az újabb változó bevonása szignifikáns növekményt eredményez-e a függő változó megmagyarázott varianciájában. Ennek megfelelően a Korrigált R² oszlop nem a kumulatív értékeket, hanem az egyes változók hozzáadása hatására előálló növekményeket tartalmazza.

A 8. táblázatban az látható, hogy az ingertípus mint bináris függő változó után a MAMUT-R teljesítményt bevonva szignifikánsan és csaknem 13 százalékkal nő a függő változó variációjának megmagyarázott része. A GMP5 (gyorsított mondatok megértése) bevonása azonban ehhez már nem tesz hozzá semmit – úgy tűnik, hogy a GMP5 változó variációja a MAMUT-R-hez képest már redundáns. Következő próbálkozásként a MAMUT-R-t lecseréltük a LAPP-szókincesvizsgálatra, a GMP5-öt pedig a GMP17-re (beszédhangok differenciálása) (lásd 9. táblázat).

Független változó	Együtthatók							Korrigált R ²	elsőfajú hiba
	B	sztđ hiba	CI95 alsó	CI95 felső	béta	elsőfajú hiba			
Ingertípus	0.352	0.047	0.258	0.446	0.627	p<0.001	0.359	p<0.001	
LAPP	0.007	0.003	0.001	0.014	0.216	p<0.05	0.083	p<0.002	
GMP17	0.004	0.002	0.000	0.008	0.204	p<0.05	0.028	p<0.05	

9. táblázat. A második regressziós modell, mely a szókinces és a beszédhang-differenciálás hatásait vizsgálja. Az elsőfajú hibák és a korrigált R² értékek értelmezése ugyanaz, mint a 8. táblázatban.

A második modell valamivel érdekesebb az elsőnél, mivel itt a LAPP szókinceseszt bevonása első lépésben 8.3%-os növekedést eredményez a megmagyarázott variációban, majd ehhez adódik hozzá egy szerény 2.8%-os növekmény a GMP17 hatására. E modell értelmezéséhez alább még visszatérünk. Végül harmadik modellünket mutatja a 10. táblázat.

Független változó	Együtthatók							Korrigált R ²	elsőfajú hiba
	B	sztđ hiba	CI95 alsó	CI95 felső	béta	elsőfajú hiba			
Ingertípus	0.382	0.049	0.285	0.478	0.680	p<0.001	0.359	p<0.001	
LAPP	0.008	0.003	0.002	0.014	0.242	p<0.01	0.083	p<0.002	
GMP5	0.003	0.001	0.001	0.006	0.214	p<0.05	0.034	p<0.05	

10. táblázat. A harmadik regressziós modell. Az elsőfajú hibák és a korrigált R² értékek értelmezése ugyanaz, mint a 8. táblázatban.

A 10. táblázatban a szókinces mellett visszatérünk a gyorsított mondatok észleléséhez (és elismétléséhez). Itt a LAPP teszt 8.3 százalékos hozzájárulása a megmagyarázott variációhoz kiegészül a GMP5 3.4 százalékával, ami továbbra is csekély. Az ingertípus mellett a MAMUT-R-t és a GMP17-et bevonva utóbbi már nem hozott szignifikáns növekményt a megmagyarázott variációban – hasonlóan a GMP5-höz az 1. modellben. Végül az ingertípus kihagyásával és a MAMUT-R, a LAPP T érték, és a GMP 17 bevonásával végzett elemzés szintén azt mutatta, hogy a MAMUT-R mellett a másik két feladat egyike sem volt érdemi hatással a magyarázó erőre. Mindenesetre az eddigiek alapján úgy tűnik, hogy fennáll a lehetőség a redukált beszédformák észlelése mögötti feldolgozási komponensek pontosabb modellezésére, egymástól való jobb elkülönítésére; ezen belül érdekes kérdés a két szintézistípus feldolgozása közötti különbségek további vizsgálata.

KÖVETKEZTETÉSEK

Az e tanulmányban bemutatott eredmények egyik része – az életkor, ingertípus, és rövid távú tanulás hatása – hasonló a korábbi vizsgálatunkban (Jakab és mtsai, 2018) kapott eredményeinkhez. A nyelvi képességekhez való viszonyt illetően igen kevés előzetes támpontunk volt; egészében tehát jelen tanulmány egy exploratív vizsgálatot ismertet. A kapott eredmények egy része némileg váratlan, és több kérdést is fölvet, melyek további vizsgálatot igényelnek.

Az egyértelmű életkori hatást korábbi vizsgálatunkhoz hasonlóan megkaptuk; a redukált beszédformák észlelési teljesítménye javul az életkorral. Szinuszhullámú ingerek esetén a felnőttek jobban teljesítettek az óvodásoknál és az iskolásoknál is. Amplitúdómodulált ingerekre már az óvodások és az iskolások teljesítménye is különbözik egymástól, az iskolásoké és a felnőtteké szintén. Az ingertípust illetően megfordult a viszony a korábbi vizsgálatunkhoz képest: most a szinuszhullámú beszéd észlelésében voltak jobbak a személyeink. Ennek azonban módszertani oka van, a jelen vizsgálatban a korábbi hatsávós helyett négyávós amplitúdómodulációt használtunk, hasonlóan Nittrouer és munkatársai vizsgálataihoz.⁹ Ez nehezebbnek bizonyult a szinuszhullámú beszédnél minden koresoportban. Hasonló volt a viszony a két szintézistípus között az idézett angol nyelvű vizsgálatokban is. Az életkor és ingertípus között méretében kicsi, de a trimmelt VA szerint szignifikáns interakciót kaptunk. Ennek lényege, hogy felnőtteknél a szinuszhullámú ingerek előnye az amplitúdómoduláltakkal szemben valamivel kisebb, mint a másik két csoportban. Ez az interakció irányában egybeesik egy korábbi jóslatunkkal (Jakab és mtsai, 2018, p. 204–205 és 6. ábra; Nittrouer & Lowenstein, 2010), mely szerint a szinuszhullámú és amplitúdómodulált ingerek közti teljesítménykülönbség csökken az életkorral – és a bevezetőben írottak alapján ez nem csak plafonhatás következménye lehet. Ugyanakkor a mi adatainkban ez az interakció olyan kicsi volt, hogy abba – a bevezető részben említett – életkorral megjelenő eltolódást a statikus koartikuláció növekvő szerepe irányába aligha lehet beleértelmezni. Ez nem véletlen, hiszen ilyen finom hatásokat elsősorban szótagészlelési vizsgálatban lehet kimutatni, ahogyan Nittrouer és munkatársai (2014) is tették; a mondatészlelésben megjelenő fogalomhajtott folyamatok ezt nagyjából elmoshatják.

A próbasorszám és teljesítmény közti korreláció véleményünk szerint értelmezhető egyfajta gyors perceptuális tanulásként; a javulás eszerint a szokatlan hangzású ingerek mondatokként való azonosításában következhet be. E hatást a hat vizsgált alcsoportból ötben sikerült kimutatni, és a hatodikban (a felnőtt szinuszhullámú mintán) vélhetőleg ennek csak a plafonhatás volt az akadálya. Korábbi vizsgálatunkban, ahol hosszabb, bonyolultabb mondatokat használtunk, és ahol emiatt a felnőtt szinuszos csoportban nem kaptunk plafonhatást, épp ez a csoport volt az egyetlen, ahol a kísérleti ülés során kimutatható volt a teljesítmény javulása. Ennek alapján úgy gondoljuk, hogy a gyors perceptuális tanulás jelen van a szinuszhullámú és amplitúdómodulált beszédingerek észlelésében is, gyerekeknél és felnőtteknél egyaránt. Sőt, ez a hatás elég jelentős: adataink szerint egy viszonylag rövid kísérleti

⁹ Másik motívációnk a négyávós szintézisre való áttérésre pont az volt, hogy a hatsávós amplitúdómoduláció már gyerekeknek is elég könnyű volt, felnőtteknél pedig egyértelmű plafonhatást eredményezett. A négyávós szintézis minden koresoportban jól differenciált. Plafonhatást most a szinuszhullámú felnőtt csoportban kaptunk, ennek oka azonban az lehetett, hogy mostani mondataink lényegesen rövidebbek és egyszerűbbek voltak a korábbiaknál.

ülés során a csoport átlagteljesítményében megjelenő variancia 40-50 százalékát is magyarázhatja a gyakorlás következtében előálló javulás. Ezt a jelenséget ismereteink szerint mások még nem vizsgálták.

Ellenvetésként fölmerülhet, hogy a teljesítmény javulását a kísérleti ülés során más tényezőkkel – pl. a munkamemória, illetve a végrehajtó funkciók szintjén történő ráhangolódással – is lehet magyarázni, tehát nem biztos, hogy itt perceptuális tanulásról van szó. Válaszunk erre a következő: egyrészt a verbális munkamemória szintjén egy 15 perces ülés során kevésbé várhatunk határozott javulást, mint az észlelési feldolgozásban, ahol, mint a bevezetőben láttuk, van erre némi evidencia. Másrészt a jelen vizsgálatban az ingermondataink lényegesen rövidebbek voltak, mint a korábbiakban, és ezzel összhangban az általános teljesítményszint is jobb volt. Ez jól magyarázható a csökkent munkamemória-terheléssel, főleg gyerekeknél. Ugyanakkor a rövidebb ingermondatokkal (tehát a munkamemória tényező jelentőségének csökkentésével) és több próbával a korábbinál határozottabb teljesítményjavulást sikerült elérnünk. Természetesen a perceptuális tanulás kérdése a jelenlegihez hasonló kísérleti helyzetekben ezzel még nem tekinthető lezártnak.

A SZÓL-E teszt és a két redukált beszédforma észlelése közötti összefüggés az a terület, ahol eredményeinket némi bizonytalanság terhelheti a kis mintac elemszám miatt. A hangtani tudatosság és az álszó-megkülönböztetés kapcsolata a beszédészlelési feladatunkkal jól értelmezhető, hiszen mindkét nehezített beszédészlelési helyzetben a fonológiai feldolgozásra jelentős teher hárul. A morfémaprodukció szerepe már elgondolkoztatóbb; a kétféle beszédingerre, két különböző számítási eljárással elvégzett hierarchikus klaszteranalízis lényegében mind a négy esetben jelzett bizonyos kapcsolatot a morfémaprodukciós képesség és a redukált beszéd észlelése között. Ez a kapcsolat talán a nyelvtani feldolgozás támogató, restaurációs hatását jelezheti, amely a mondat elismétlése, „összerakása” során fejt ki hatását. A *nyelvtani morféma megértése* változó igen erős plafonhatást mutatván nem volt használható ennek az értelmezésnek a kiegészítésére; itt tehát nyitva maradt egy fontos kérdés, melyet későbbi vizsgálatokban kell tisztázni.

A többi nyelvi tesztnek a beszédészleléssel mutatott kapcsolatai szintén jól értelmezhetők. Érdekes megfigyelés a 7. táblázat alapján, hogy a mondat- és szóazonosítás zajban (GMP2 és GMP3) szorosabban látszanak kapcsolódni a szélsusugásra, suttogásra emlékeztető amplitúdómodulált beszéd észlelésével, mint a szinuszhullámúval. A bevezetőben említettük, hogy a szinuszhullámú beszéd statikus és dinamikus koartikulációs információt is megőriz, míg az amplitúdómodulált inkább csak statikust (Nittrouer & Lowenstein, 2014). Lehetséges tehát, hogy a zajban való beszédészlelés és az amplitúdómodulált ingerek dekódolása részben átfedő mechanizmusokon alapul. Ennek egy lehetséges módja, hogy a zajban való beszédészlelés, az amplitúdómodulált beszédhez hasonlóan, nagyobb mértékben támaszkodik statikus koartikulációs jegyekre, mint dinamikusakra. Hogy ez így van-e, arra vonatkozóan eddig nem találtunk szakirodalmi adatot. Egyes vizsgálatok szerint a beszédészleléshez szükséges jel-zaj arány észrevehetően csökken iskolás korban (Bradley & Sato, 2008), ám nem teljesen világos, hogy van-e kapcsolat a hangtani tudatosság és a zajban való beszédészlelés képessége között (Lewis és mts., 2010). Ami a további eredményeket illeti e csoportban, adataink szerint a gyorsított mondatok észlelése (és visszamondása; GMP5) és a LAPP szókinceszt is valamivel specifikusabb kapcsolatban van a redukált beszéd észlelési teljesítményével, mint a MAMUT-R teszt eredménye, ami a többi nyelvi képességeteszt közül szinte mindegyikkel elég szorosan együtt mozog (lásd a (B) függelékét). A beszédészleléshez szorosabban kapcsolódó szubtesztek

mellett az aktív szókincs mértéke bír észrevehető magyarázó erővel a szinuszhullámú és amplitúdómodulált mondatok megértésében.

Végül még egy megjegyzést teszünk regressziós modelljeinkkel kapcsolatban. Érzésünk szerint e modellekkel még nem sikerült kielégítően megragadnunk a fonológiai feldolgozás jelentőségét a redukált beszéd észlelésében. Ennek egyik oka valószínűleg az, hogy a választott tesztfeladatok, vagy azok egy része, nem elég hatékonyan szeparálja egymástól a nyelvi részképességeket (erről lásd pl. Bertalan, 2005). A kiutat ebből a helyzetből a prediktorváltozóink újraválogatása, illetve egyes komponens feladatok továbbfejlesztése jelentheti. A több képességetest esetén is nehezen elkerülhető, kisebb-nagyobb plafonhatás szintén nem tehetett jót a megmagyarázott varianciaarányoknak.

Köszönetnyilvánítás

A kézirat szerzői köszönetüket fejezik ki az ELTE Bárczi Gusztáv Gyógypedagógiai Kar munkatársainak, elsősorban Keresztessy Évának, a kar audioeszközeinek és hallásvizsgáló laborjának rendelkezésre bocsátásáért, Vargha Andrásnak és Kas Bencének a kézirat első változatával kapcsolatos javaslataikért, valamint az adatok felvételében közreműködő gyógypedagógus hallgatóknak: Szanyi Andreának, Süveggyártó Tibornének, Táló Lillának, Horváth Lillának és Szűcs Eszternek.

IRODALOMJEGYZÉK

- Bertalan R. (2005). A GMP diagnosztikai eljárás bemutatása és kritikus próbája. *Beszédgyógyítás*, 16(2), 7–26.
- Blumstein, S. E., & Stevens, K. N. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America*, 66(4), 1001–1017.
- Bradley, J. S., & Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *Journal of the Acoustical Society of America*, 123(4), 2078–2086.
- Bregman, A. S. (1990). *Auditory Scene Analysis*. Cambridge Mass: The MIT Press.
- Gósy M. (1995). *GMP-diagnosztika. A beszédészlelés és beszédmegértés folyamatának vizsgálata*. Budapest: Nikol GMK.
- Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research*, 26(3), 499–510.
- Hudák M., Jakab Z., & Kovács I. (2013). Phenomenal qualities and the development of perceptual integration. In Albertazzi, L. (Ed.), *The Wiley-Blackwell Handbook of Experimental Phenomenology: Visual Perception of Shape, Space and Appearance*, Malden, MA, Wiley & Blackwell (pp. 145–161).
- Jakab Z., Nagyné Ringer G., Víg J., & Szabó P. T. (2018). Szinuszhullámú és amplitúdómodulált beszéd észlelésének vizsgálata magyar mondatok segítségével. *Magyar Pszichológiai Szemle*, 73(2), 183–211.
- Jenkins, J. J., Strange, W., & Edman, T. R. (1983). Identification of vowels in “vowelless” syllables. *Perception & Psychophysics*, 34(5), 441–450.

- Kas B., Lőrík J. M., Bogáth R. Sz., & Mályi N. (2012). *SZÓL-E? Szűrőeljárás az óvodáskori logopédiai ellátáshoz*, Budapest: Logotech Kiadó.
- Kas B., & Lukács Á. (előkészületben). *Magyar Mondat-utánmondási Teszt (MAMUT)*.
- Kovács I. (2000). Human development of perceptual organization. *Vision Research*, 40(10-12), 1301–1310.
- Kovács I., & Eisenberg M. (2005) Human development of binocular rivalry. In Alais, D., & Blake, R. (Eds.), *Binocular Rivalry* (pp. 101–116). Cambridge Mass: The MIT Press.
- Lewis, D., Hoover, B., & Choi, S. (2010). The relationship between speech perception in noise and phonological awareness skills for children with normal hearing. *Ear and Hearing*, 31(6), 761–768.
- Lőrík J., Ajtony P., Palotás G., & Pléh Cs. (2015). *Aktívizókincs-vizsgálat*. Budapest: Educatio Társadalmi Szolgáltató Nonprofit Kft.
- Maurer, D., Lewis, T. L., & Mondloch, C. J. (2005). Missing sights: consequences for visual cognitive development. *Trends in Cognitive Sciences*, 9(3) 144–151.
- Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics*, 18(5), 331–340.
- Mooney, C. M. (1957). Age in the development of closure ability in children. *Canadian Journal of Psychology*, 2(4), 219–228.
- Newman, R. S., Chatterjee, M., Morini, G., & Nasuta M. (2013). 'Toddlers' comprehension of noise-coded speech and sine-wave analogs to speech. *Proceedings of Meetings on Acoustics, Vol., 19* (ICA 2013, Montreal), <http://acousticalsociety.org/>.
- Nilsson, M., Soli, S. D., & Gelnett, D. J. (1996). *Development and norming of a hearing in noise test for children* (House Ear Institute, Los Angeles, CA).
- Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and noise. *Journal of the Acoustical Society of America*, 95(2), 1085–1099.
- Nittrouer, S., Lowenstein, J. H., & Packer, R. (2009). Children discover the spectral skeletons in their native language before the amplitude envelopes. *Journal of Experimental Psychology: Human Perception and Performance*. 35(4), 1245–1253.
- Nittrouer, S., & Lowenstein, J. H. (2010). Learning to perceptually organize speech signals in native fashion. *Journal of the Acoustical Society of America*, 127(3), 1624–1635.
- Nittrouer, S., & Lowenstein, J. H. (2013). Perceptual organization of speech signals by children with and without dyslexia. *Research in Developmental Disabilities*, 34(8), 2304–2325.
- Nittrouer, S., & Lowenstein, J. H. (2014). Dynamic spectral structure specifies vowels for adults and children. *Language and Speech*, 57(0 4), 487–512.
- Racsmány M., Lukács Á., Németh D., & Pléh Cs. (2005). A verbális munkamemória magyar nyelvű vizsgálóeljárásai. *Magyar Pszichológiai Szemle*, IX(4), 479–505.
- Remez, R. (2005). Perceptual Organization of Speech. In Pisoni, D. B., & Remez, R. E. (Eds.), *Handbook of Speech Perception*. Oxford: Blackwell.
- Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science*, 212(4497), 947–950.
- Remez, R. E., Rubin, P. E., Nygaard, L. C., & Howell, W. A. (1987). Perceptual normalization of vowels produced by sinusoidal voices. *Journal of Experimental Psychology, Human Perception & Performance*, 13(1) 40–61.

- Remez, R. E., & Rubin, P. E. (1990). On the perception of speech from time-varying acoustic information: Contributions of amplitude variation. *Perception & Psychophysics*, 48(4), 313–325.
- Remez, R. E., Rubin, P. E., Berns, S. M., Pardo, J. S., & Lang, J. M. (1994). On the perceptual organization of speech. *Psychological Review*, 101(1), 129–156.
- Rosner, B. S., Talcott, J. B., Witton, C., Hogg, J. D., Richardson, A. J., Hanson, P. C., & Stein, J. F. (2003). The perception of "sine-wave speech" by adults with developmental dyslexia. *Journal of Speech, Language and Hearing Research*, 46(1), 68–79.
- Rueff, M., Furmann, A., & Muñoz-Sandoval, A. (szerk., 2003). *Woodcock–Johnson nemzetközi kiadás. Vizsgálatvezetői kézikönyv – magyar kiadás.* (Fordította: Katona N.). The Woodcock Muñoz Foundation, Nashville.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270(5234), 303–304.
- Strange, W., Verbrugge, R. R., Shankweiler, D. P., & Edman, T. R. (1976). Consonant environment specifies vowel identity. *Journal of the Acoustical Society of America*, 60(1), 213–224.
- Strange, W., Jenkins, J. J., & Johnson, T. L. (1983). Dynamic specification of coarticulated vowels. *Journal of the Acoustical Society of America*, 74(3), 695–705.
- Vargha A. (2000). *Matematikai statisztika pszichológiai, nyelvészeti és biológiai alkalmazásokkal*, Budapest: Pólya Kiadó.
- Vargha A. (2016). Szignifikanciatesztek – negyven éve hibás elemzéseket végzek, és téveszmét tanítok? *Statisztikai Szemle*, 94(4), 445–451.
- Vargha A. (2019). *Többváltozós statisztika dióhéjban: Változó-orientált módszerek*. Budapest: Pólya Kiadó.
- Vargha A. (2020). *Normális vagy? És ha nem? Statisztikai módszerek nem normális eloszlású változókkal, pszichológiai kutatásokban*. Budapest: Pólya Kiadó.

Függelék

(A) A beszédészlelési vizsgálatban felhasznált magyar mondatok és angol megfelelőik a HINT-C tesztben.

HINT-C eredeti	Magyar változat
Bemelegítő próbák:	
The mailman shut the gate.	A postás levelet hozott.
The book tells the story	Mesélek neked a könyvből.
The orange is very sweet.	A citrom nagyon savanyú.
Éles próbák:	
The yellow pears taste good	Fejemre esett egy alma.
The pond water was dirty	A pocsolya vize sáros
The little girl was shouting	A kislány nagyot kiáltott
Flowers grow in the garden	Virágok nyílnak a kertben.
She looked in her mirror.	A lány tükrébe pillantott.
He was sucking his thumb.	Ebéd után alszom egyet.
He grew lots of vegetables.	A bácsi zöldséget árul.
He hung up his raincoat.	Itt van az esőkabátod.
The police helped the driver.	A rendőr segített nekünk.
He wore his yellow shirt.	Szeretem a sárga ingem.
They had a wonderful day.	Egész nap együtt játszottunk.
The bananas were too ripe.	A banán héja megbarnult.
The kitchen window was clean.	A konyha ablaka párás.
The ball bounced very high.	A labda magasra pattant.
Someone is crossing the road.	Valaki átment a hídon.
The sun melted the snow.	A hó estére elolvadt.
The chicken laid some eggs.	Tyúkanyó három tojást tojt.
The ball broke the window.	A szoba ablaka betört.
The baby wants his bottle.	A baba elkezdett sírni.
The oven door was open.	A sütő ajtaja kinyílt.

(B) A SZÓL-E teszten kívül használt nyelvi képességsztek egymás közti korrelációi a 76 fős óvodás-iskolás al minta alapján.

	MAMUT R	GMP17	GMP2	GMP3	GMP4	GMP5	Álszóism.	Számterj.	Ford. számt.
LAPP T	0.401**	0.409**	0.303**	0.456**	0.336**	0.289*	0.374**	0.324**	0.289*
MAMUT R		0.563**	0.377**	0.294**	0.387**	0.480**	0.413**	0.417**	0.550**
GMP17			0.461**	0.278*	0.066	0.375**	0.500**	0.493**	0.547**
GMP2				0.338**	0.068	0.272*	0.451**	0.346**	0.407**
GMP3					0.249*	0.312**	0.391**	0.243*	0.322**
GMP4						0.262*	0.274*	0.288*	0.261*
GMP5							0.370**	0.412**	0.427**
Álszóism.								0.431**	0.595**
Számterj.									0.371**

Szignifikanciaszintek (elsőfajú hiba korrekció nélkül): * $p < 0.05$; ** $p < 0.01$

SINE-WAVE AND AMPLITUDE-MODULATED SPEECH: PERCEPTION, PERCEPTUAL LEARNING, AND THEIR RELATION TO TESTS OF LANGUAGE DEVELOPMENT

ZOLTÁN JAKAB – KATALIN MOHAI – JULIANNA VIG

jakab.zoltan@barczy.elte.hu

mohai.katalin@barczy.elte.hu

vig.julianna@barczy.elte.hu

ABSTRACT

Background and aims: Studies of the perception of sine-wave speech (SWS) and amplitude-modulated (AM or vocoded) speech initially addressed fundamental questions of language processing; subsequently researchers also became interested in what these phenomena tell us about the development of speech perception. In the present study, we pursue the developmental line, and address two questions related to the perception of SWS and AM speech. First, is it possible to demonstrate fast perceptual learning in the perception of these types of speech stimuli, in children and adults? Second, how do different perceptual and cognitive abilities subserving language acquisition relate to the perception of these forms of speech?

Methods: We examined three age groups (kindergarteners, grade schoolers, and adults; 134 participants in total). Our speech perception task used approximate Hungarian translations of sentences from the Hearing in Noise Test for Children (HINT-C) for SWS and AM synthesis. Twenty four words and eight syllables were used for three-formant SWS and four-band AM synthesis; the synthesized versions were used in the experimental trials. To assess language skills, a variety of language tests developed in Hungary were used in addition to a reverse digit span procedure from the Woodcock-Johnson test of cognitive abilities. To answer our questions, correlation- and regression-based analyses were conducted on the data.

Results: We found a substantial improvement in speech perception performance within an experimental session comprising 20 synthesized sentences (either SWS or AMS for a given subject) which we interpreted as the result of fast perceptual learning. Regarding the language tests, perception of SWS and AM speech was most strongly associated with subtests measuring phonological processing (phoneme differentiation; the understanding of accelerated speech), sentence repetition, and productive vocabulary. One subtest of morpheme production also appeared to be closely related to SW and AM speech perception.

Conclusions: The observed relationship between the perception of SWS and AMS on the one hand, and language subtests on the other is likely underlain by phonological processing as a common core; in addition, the association with morpheme production likely indicates a top-down, restoration effect due to grammatical processing. Based on our results we see a chance that the tasks of SWS and AMS perception can be used in developing phonological diagnostic and training procedures.

Keywords: development of speech perception, sine-wave speech, vocoded speech, perceptual learning, tests of linguistic ability
