

# Miért különleges a beszéd?

## A beszéd mint kitüntetett hanginger a hatékony idegi kódolás elmélete szerint

*Gervain Judit*

Laboratoire Psychologie de la Perception, CNRS  
Université Paris Descartes

### 1. Bevezetés

Az elképzelést, miszerint „a beszéd különleges”, többen több megközelítésből is megfogalmazták már. A társas viselkedés kutatói szerint a beszélt nyelv azért különleges, mert az emberi faj kommunikációs rendszere, így kiemelt szerepe van fajunk társas viselkedésének szervezésében (Tomasello 2000). Mások szerint az teszi a beszédet a többi hangingerrel különbözővé, hogy saját magunk hozzuk azt létre, így az egyetlen olyan hang, amelyet nemcsak hallunk, hanem testünkkel érzünk is, hiszen érezzük, milyen helyzetben vannak a beszélőszerveink egy adott hang, szó vagy mondat kimondásakor. Így a mozgásérzékelés és a hallás egyaránt szerepet játszanak a beszéd észlelésében (Liberman et al. 1967). A jelen dolgozat célja, hogy olyan elméleti keretet mutasson be – a hatékony idegi kódolás elméletét –, amely új, a korábbiakat kiegészítő módon magyarázza a beszéd különleges természetét.

### 2. A hatékony idegi kódolás elmélete

A hatékony idegi kódolás elmélete abból indul ki, hogy az idegrendszer egyik alapvető feladata az információfeldolgozás és -tárolás. Különösen releváns ez az érzékszervek esetében, amikor is az idegrendszernek külső információt – a környezetből jövő ingereket – kell feldolgoznia, ezeknek a külső ingereknek ugyanis megvan a maguk információs szerkezete. Az ingerek, még azonos modalitáson belül is, eltérnek egymástól gyakoriságukban és szerkezetükben. Logikusnak tűnik hát feltételezni, hogy ezek a statisztikai és szerkezetbeli tulajdonságok valamilyen módon leképeződnek az érzékszervek anatómiai és funkcionális jellemzőibe.

Az információelmélet empirikusan vizsgálható keretet biztosít e hipotézis tesztelésére (Shannon 1948). E szerint az elméleti keret szerint – amely hatékony idegi kódolás néven vált ismertté (Attneave 1954; Barlow 1961) – az érzékszervek akkor tudják a lehető leghatékonyabban, azaz a legkisebb költség mellett a legtöbb információt tárolva, feldolgozni a külső ingereket, ha azok statisztikai szerkezetét kódolják, azaz a lehető legkisebbre csökkentik a tárolt információban a redundanciát. (Megjegyzendő, hogy a redundancia csökkentése matematikai elv; egy biológiai rendszerben, amilyen a nyelv is, a redundancia teljes kiküszöbölése nem lehetséges és nem is kívánatos, hiszen ez a rendszert sérülékennyé teszi, és csökkenti annak zajjal szembeni ellenállását.)

A fenti predikció teszteléséhez arra van szükség, hogy (i) (egy vagy több fizikai dimenzió mentén) meghatározzuk egy adott környezeti inger statisztikai tulajdonságait, majd (ii) ezeket összehasonlítsuk az inger idegi kódolásával. Mint ahogy a hatékony idegi kódolás elméletének kritikusai is felvetik, mindkét feladat aluldeterminált, azaz maga az elmélet semmit nem mond arról, melyek a releváns dimenziók, illetve statisztikai eloszlások, amelyeket az idegrendszer kódolhat, és arról sem szól, hogy az idegi feldolgozás mely szintjén kell a hatékony kódolásnak megnyilvánulnia (az egyes idegsejtek, az érzékszervek vagy a magasabb rendű feldolgozás, pl. az agykéreg szintjén).

Ennek ellenére az utóbbi fél évszázadban viszonylag sok eredmény született a látórendszer vizsgálata kapcsán, amelyek igazolni látszanak a hatékony kódolás elméletét (Simoncelli–Olshausen 2001). Kísérletek bizonyítják például, hogy a légy összetett szeme, valamint a macskák és majmok látókérge optimálisan képezi le a fényintenzitás azon értékeit, amelyek ezen állatok természetes környezetében tipikusan előfordulnak. Az emlősök retinájában található csapok pedig a lehető legkevesebb dimenzióval, mindössze hárommal lefedik a fény látható tartományát. Úgyszintén bizonyított, hogy a képek feldolgozásában nagyon fontos szerepet játszanak az élek, amelyeket az elsődleges, illetve a magasabb szintű agyi látómezőkben kifejezetten az élekre (és azok irányára/dőlésszögére) érzékeny idegsejtek kódolnak. A különféle természetes képek matematikai elemzése azt mutatja, hogy az élek ebből a szempontból is lényeges elemei a vizuális ingerek optimális, redundancia nélküli kódolásának.

A látórendszer kutatása arra enged tehát következtetni, hogy a hatékony idegi kódolás elmélete valóban hozzájárulhat az idegrendszer működésének jobb megértéséhez. Az elmúlt egy-másfél évtizedben néhány munka az elméletet a hallórendszerre is kiterjesztette, ami így relevánssá vált a beszélt nyelv kitüntetett szerepének vizsgálatában is.

### 3. A hangingerek statisztikai szerkezete

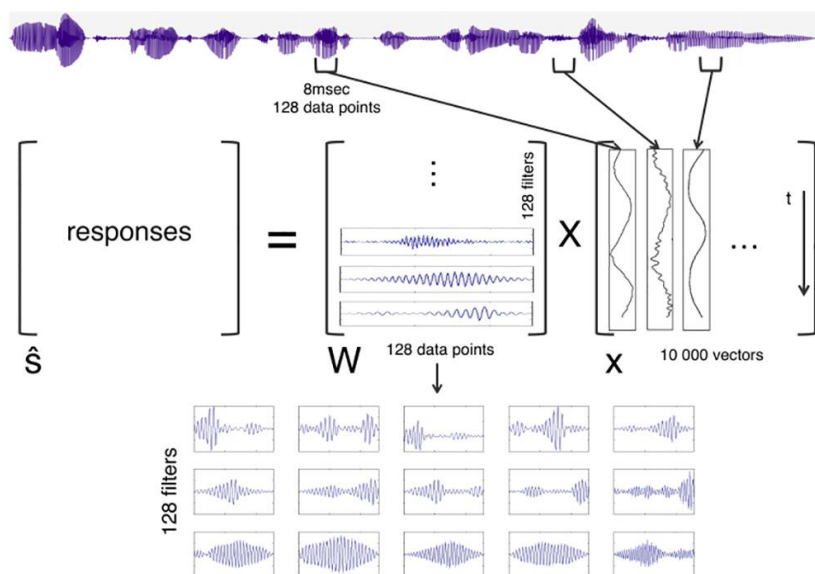
Ahhoz, hogy a hatékony idegi kódolás elméletét a hallórendszerrel kapcsolatban is értelmezni lehessen, szükségünk van olyan matematikai leírásokra, amelyek a hangok tulajdonságait az információ elmélet keretében vizsgálják. A hangingerek hagyományos leírásainak, pl. a Fourier-elemzésnek, nem célja, hogy optimális, a redundanciát a lehető legnagyobb mértékben lecsökkentő reprezentációkat adjanak. Az elmúlt néhány évtizedben azonban megjelentek olyan algoritmusok, matematikai módszerek, amelyeknek kifejezetten az a célja, hogy az elemzett ingert egymástól teljesen független, azaz redundanciamentes elemek együtteseként kódolják. Ezeket először a jelfeldolgozás területén, illetve a látás vizsgálatában használták, de az utóbbi 10–15 évben a hangokra is alkalmazni kezdték őket.

#### 3.1. Optimális szűrők a hallórendszerben

Az optimális kódolási algoritmusok egyik fajtája (pl. a független komponens analízis, *independent component analysis*, ICA vagy a főkomponens analízis, *principal component analysis*, PCA) arra törekszik, hogy a bemeneti adatokat egymástól teljesen független összetevőkre bontsa le (1. ábra). Hogy az algoritmus használatát hétköznapi példával illusztráljuk, képzeljünk el egy forgalmas várótermet, melynek különböző pontjain mikrofonok vannak elhelyezve – legalább annyi, ahány hangforrás található a váróteremben (emberek, hangosbemondó stb.). Fizikai elhelyezkedésénél fogva minden mikrofon kicsit másképp veszi fel a váróterem hangzavarát. A feladatunk az, hogy a mikrofonok felvételeit használva rekonstruáljuk minden egyes hangforrás eredeti megnyilvánulását. Az ICA algoritmus épp erre képes, mivel a hangforrások egymástól függetlenül szólalnak meg.

Az ICA algoritmust használva Lewicki (2002) három hangtípus statisztikai szerkezetét vizsgálta: a természeti környezetből jövő hangokét (ágak recsegése, rianás, kövek kopogása stb.), az állathangokét (nagyraest emlősök vokalizációit) és a beszédét (amerikai anyanyelvű férfi és női beszélők olvasott megnyilvánulásait). Az algoritmus a bemenetet képező hangokat (1. ábra felső sora, illetve a középső sorban az  $X$  mátrix) egy szűrőkészleten engedi át, amelyek azokat a hatékony kódolás elméletének értelmében egymástól független komponensekre bontják. A szűrőkészlet adja meg tehát azt a matematikai átalakítást, amelyet a idegrendszer (valamely része) végez a hangingereken, hogy azokból redundanciamentes idegi reprezentációt alkosson. Lewicki azt találta, hogy a természetbeli hangok kódolásakor kapott szűrőkészlet szűrői wavelet alakúak, azaz időben rövidek, lokalizáltak, és meghatározott frekvenciájuk

van, míg az állathangok szűrői inkább Fourier-transzformációhoz hasonlítanak, vagyis meghatározott frekvenciájuk van, de időbeli információt nem vagy csak keveset közvetítenek. Önmagában ezen szűrőkészletek egyike sem hasonlít a hallórendszerben mért idegi reakciókra. Viszont a két hangtípust közel azonos arányban keverő hangegyüttes elemzésekor kapott szűrőkészlet megegyezik az emlősök hallóidegében (a cochlea szintjén) mért elektrofiziológiai válaszokra. Ezen válaszok, illetve szűrők formája átmenet a Fourier- és a waveletszerű szűrők között. Pontosan kódolják a frekvenciát, és időbeli információt is hordoznak, sőt mi több, a kettő összefüggésben is van egymással, mert a szűrők hossza fordítottan arányos a frekvenciájukkal. Ez azt jelenti, hogy a szűrők alakja léptékfüggetlen, mert különböző frekvenciák mellett is változatlan marad. A léptékfüggetlenség (vagy önazonosság vagy fraktálszerkezet) számos természeti jelenség és inger sajátja (Simoncelli–Olshausen 2001), így például a hópelek, a pagodakarfiol, a tengeri csigák háza vagy a kontinensek partvonala is léptékfüggetlen.



1. ábra

A független komponens analízis sematikus bemutatása (Guevara Erra–Gervain 2016 alapján). A beszélt nyelvi hangingereket a felső sor illusztrálja. Ezekből az algoritmus véletlenszerűen kiválaszt 8msec hosszú részeket, amelyek a bemeneti mátrixot ( $X$  mátrix a középső sorban) képezik. Az algoritmus kimenete az  $S$  mátrix jelölte független idegi válaszok együttese. Az idegi kódolás matematikai modelljét ennek megfelelően a  $W$  keverőmátrix adja meg. A kapott szűrők közül néhányat a legalsó sorok illusztrálnak.

A fenti eredmények igazolják tehát a feltételezést, miszerint a hallórendszer hatékonyan kódolja a hangingereket, legalábbis ha elfogadjuk, hogy a környezeti hangok és az állati vokalizációk valamiféle keveréke valóban reprezentatív mintája annak, amihez az evolúció során a hallás adaptálódhatott. Elméleti kérdésünk szempontjából még fontosabb, hogy a beszédre kapott szűrőkészlet egybe esett az előbb említett hangkeverék szűrőkészletével, és így a kísérletesen mért idegi válaszokkal is. Ez arra enged következtetni, hogy a beszéd az evolúció során úgy alakult ki, hogy a már meglévő idegi kódot a lehető legoptimálisabban használja ki.

A beszéd tehát azért (is) különleges inger az idegrendszer számára, mert matematikai szempontból hatékonyan kódolható. Lewicki (2002) feltételezése szerint azért, mert a nyelv bizonyos fonémái az állati vokalizációkra, mások a természeti környezet hangjaira hasonlítanak akusztikai tulajdonságaikban. A magánhangzók például jellegzetes formánsaik vannak, tehát meghatározott frekvenciák hordozzák őket, és időben hosszúak, míg a zárhangok rövidek, és az akusztikus energia gyors, robbanásszerű emelkedése, majd valamivel lassabb csökkenése, és a hirtelen átmenetek jellemzik őket.

Erre az ötletre építve Stilp és Lewicki (2013) a fonémák különböző osztályainak hatékony kódolását vizsgálták. Lewicki (2002)-höz hasonlóan az ICA algoritmust használták, de nem egész mondatokat, hanem egyes fonémákat elemeztek. Eredményeik csak részben igazolták az eredeti elképzelést. Azt találták ugyanis, hogy a magánhangzók szűrői nem az állathangok szűrőihöz, hanem általában a beszéd (illetve az állathang és környezeti hang keverék) szűrőihöz hasonlítanak. Ez a magánhangzók minden csoportjára (elől, középen, hátul képzett magánhangzók, kettőshangzók stb.) igaz volt. A mássalhangzók kategóriái azonban nagyon eltérő szűrőket produkáltak. A zárhangok a vártak megfelelően a környezeti hangterek szűrőihöz hasonló rövid, időben meghatározott, a frekvenciával fordítottan arányos hosszúságú szűrőket adtak. A zár-rés hangok a beszéd általános szűrőihöz (és így a magánhangzókéihoz) voltak hasonlóak, míg a réshangok és a nazálisok az állati vokalizációk Fourier-szerű, időben meg nem határozott, de a frekvenciát pontosan kódoló szűrőihöz hasonló filtereket eredményeztek. Úgy tűnik tehát, hogy nem a magán- és mássalhangzó státusz, hanem inkább a hangok akusztikai jellemzői határozzák meg a matematikailag optimális kódolást. Azok a hangok kódolódnak leginkább waveletszerű, időben jól meghatározott szűrőkkel, amelyek rövidek és hirtelen átmeneteket tartalmaznak, míg a hosszabban kitarított hangok (akár harmonikusak, mint a magánhangzók, akár nem, mint a réshangok) inkább a frekvenciatartományban kódolódnak pontosan.

Mivel a nyelvekben a különböző fonématisípusok nem egyforma arányban találhatóak meg, és mivel a fonémarepertoár a nyelvek más akusztikai-fonológiai tulajdonságaival, így például ritmusukkal, szótagszerkezetükkel stb. is korrelál, felmerül a kérdés, hogy minden nyelvnek hasonló-e az optimális kódja, vagy ez a kód nyelvenként eltérő-e. Guevara Erra és Gervain (2016) Lewicki módszerét hét tipológiailag és történetileg különböző nyelvre (holland, angol, maráti, japán, lengyel, spanyol és török) alkalmazva azt találta, hogy az optimális szűrők nyelvenként némileg eltérők. Bár minden nyelv szűrőkészlete nagyjából a wavelet és a Fourier-szűrők közötti átmenetet mutatja (1. ábra, alsó rész), az angol, a spanyol, a lengyel és a holland szűrői valamivel közelebb állnak a Fourier-szerű kódhoz, míg a török, a japán és a maráti szűrői kicsit inkább waveletszerűek. Ennek az az oka, hogy az utóbbi nyelvekben gyakoribbak a gyors akusztikai átmenetet tartalmazó fonémák (pl. a zár mássalhangzók), míg az előbbieken a kitarított hangok gyakoribbak.

A fonológiailag különböző nyelvek optimális kódjai eltérnek tehát egymástól, így felmerül a kérdés, hogy az idegrendszer az egyedfejlődés során alkalmazkodik-e ehhez, és ha igen, hogyan. Ez annál is inkább releváns, hiszen jól dokumentált, hogy a felnőtt beszélők beszédészlelési és -értési képességeit az anyanyelv tulajdonságai erősen befolyásolják. Idegen nyelvek hallgatásakor az anyanyelvünkhöz hasonló prozódiajű, ritmusú, fonémakészletű nyelveket könnyebben értjük, mint az anyanyelvünktől távol esőeket (Cutler 1994; Cutler-Demuth-McQueen 2002; Werker-Hensch 2015). Az anyanyelv ezen meghatározó szerepe az első néhány életévben, a nyelvtanulás kritikus periódusában alakul ki. Lehetséges, hogy az anyanyelvre való ráhangolódás egyik mechanizmusa az idegrendszeri kód finomra hangolása, az anyanyelvre optimalizált szűrők beállítása. E feltételezést kísérletesen egyelőre még nem vizsgálták, de a jövőben érdekes lehet olyan csecsemőkísérleteket végezni, amelyek tesztelik ezt a lehetőséget.

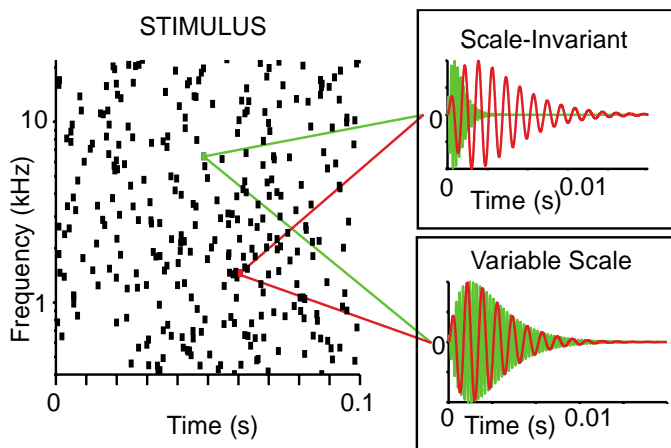
Ezzel kapcsolatban releváns megjegyezni, hogy mivel a nyelvek fonológiai tulajdonságai gyakran korrelálnak morfoszintaktikai jegyeikkel, a nyelvek korai akusztikai elkülönítése egyben a későbbi nyelvtanulást is előkészítheti. A fenti tanulmányban vizsgált nyelvek a nekik megfelelő optimális szűrők tulajdonságai alapján két nagyobb csoportra oszlottak. Ez a két nagyobb csoport egyben megfelel e nyelvek morfoszintaktikai tipológiájának is, hiszen a spanyol, a lengyel, a holland és az angol ige-tárgy szórendű flektáló nyelvek, míg a maráti, a japán és a török tárgy-ige szórendű agglutináló nyelvek. Nem egyértelmű, hogy a csecsemők hogyan tudhatnak a fonológia-akusztika és a morfoszintaxis ezen korrelációiról, de ha feltételezzük, hogy ezek a korrelációk valamilyen módon hozzáférhetőek számukra (pl. mert részei a veleszületett nyelvtanulási képességnek, vagy mert még ismeretlen, de létező biológiai, nyelvi, matematikai vagy más elvekből levezethetőek), akkor az optimális kód korai beállítása fontos „csizmahúzó” (bootstrapping) mechanizmus lehet a nyelvvelsajátítás során (Gervain–Werker 2013; Morgan–Demuth 1996; Nespor et al. 2008).

### 3.2. Ritka kódolás („sparse coding”) a hallórendszerben

Az optimális kódolási algoritmusok másik fajtája a ritka kódolás (angolul „sparse coding”). A ritka kódolási módszerek a bemeneti jeleket sok reprezentációs egység (pl. idegsejt vagy bázisvektor) segítségével bontják le, de úgy, hogy ezek közül az egységek közül egy-egy bemeneti inger kódolásakor egyszerre csak a lehető legkevesebb aktív.

Smith és Lewicki (2006) adott frekvenciájú, hosszúságú és meghatározott időpontban kezdődő hullámocskák készletével kódolt a Lewicki (2002) tanulmányban használthoz hasonló hangfajtaikat. Az eredmények a korábbiakkal egyezően azt mutatták, hogy az állati vokalizációk és a különféle természeti hangok keveréke az emlősök hallóidegében mért válaszokhoz hasonló formájú kódot (hullámocskákat) produkált, míg a csak egyfajta hangtípusra, pl. csak állathangokra, optimalizált hullámkészletek kevésbé hasonlítottak a mért idegi válaszokhoz. A mi szempontunkból releváns, hogy a beszéd (a már korábban is használt angol nyelvi beszéd-korpusz) szintén a mért idegi válaszokhoz hasonló hullámkészletet eredményezett. Különbség viszont a független komponens elemzés (Lewicki 2002) és a ritka kódolás (Smith–Lewicki 2006) között, hogy az ICA algoritmus nem tud időben aszimmetrikus kódot létrehozni, míg a ritka kódolás nem korlátozza a hullámok lehetséges formáját. Ennek megfelelően a kapott hatékony reprezentációk között lényeges eltérés volt, hogy míg az első módszer eredményezett szűrők *per definitionem* időben szimmetrikusak voltak, addig a ritka kódolással elért optimális hullámocskák hirtelen, nagy amplitúdóval kezdődtek, de fokozatosan csökkenő energiával értek véget. Ez a fajta aszimmetria jól megfigyelhető mind az elektrofiziológiailag mért válaszokban, mind a természetes hangok egy jó részében, pl. olyan környezeti hangokban, mint koppanás, reccsenés stb., vagy például a zár mássalhangzóiban is.

Geffen és munkatársainak (Geffen et al. 2011) tanulmánya a víz hangjainak ritka kódolását vizsgálta. Olyan generatív modellt dolgoztak ki (2. ábra), amelyben adott frekvenciájú, amplitúdójú és alakú gammatónusok készletével lehetett hangokat létrehozni. A gammatónusok vagy úgy voltak megválasztva, hogy léptékfüggetlenek legyenek, azaz az alakjuk legyen állandó (bennük a szinusz hullám mindig ugyanannyi ciklust futott be), vagy úgy, hogy változó léptékűek legyenek, azaz időtartamuk legyen mindig azonos. Az első esetben az alak állandóságából következően a tónusok frekvenciája fordított arányban állt hosszukkal, míg a második esetben az alakjuk a frekvencia függvényében változott. Ahogyan korábban említettük, a természetes ingerek, pl. természeti formák, bizonyos hangok és a beszéd is, léptékfüggetlenek (Simoncelli–Olshausen 2001; Voss–Clarke 1975). Ennek megfelelően Geffen és munkatársai azt várták, hogy a modell előállította léptékfüggetlen hangok természetes vízhangoknak tűnnek, míg a változó léptékű hangok nem.



2. ábra

A vízhangok ritka kódja időben és frekvenciában meghatározott gammatónusokkal Gervain és munkatársai (Geffen et al. 2011; Gervain–Werker–Geffen 2014; Gervain et al. 2016) munkájában. A gammatónusok időbeli és frekvenciabeli paraméterei kétféleképpen lehetnek meghatározva. A léptékfüggetlen (scale-invariant) kód esetében a gammatónusok formája minden frekvenciánál állandó volt, így a tónusok frekvenciája és időbeli hossza egymással fordított arányban állt. A változó léptékű kód esetében a tónusok időbeli hossza volt állandó, így formájuk a frekvencia függvényében változott.

### 3.3. A léptékfüggetlenség

E hipotézist kísérletesen is tesztelték mind felnőtteknél, mind csecsemőknél. Azt találták (Geffen et al. 2011), hogy a felnőttek valóban ugyanannyira természetesnek ítélték a léptékfüggetlen szintetikus vízhangokat, mint a valódi, természetben rögzített vízhangokat (patak zúgása), és ennek megfelelő leírásokat is adtak a hangokról (pl. eső, tenger morajlása, patak, hullámos, csepegés stb.). A változó léptékű szintetikus hangokat ezzel szemben nem ítélték természetesnek, és nem is tulajdonítottak nekik vízszerű jellemzőket. A csecsemőknél ugyan leírást és explicit természetességi ítéletet nem lehetett felvenni, de viselkedéses (Gervain–Werker–Geffen 2014), valamint agyi képpalkotó eljárást használó (Gervain et al. 2016) tesztekben a szerzők azt találták, hogy az 5 hónapos csecsemők és a 3 napos újszülöttek szintén meg tudják egymástól különböztetni a léptékfüggetlen és a változó léptékű vízhangokat, és ami még érdekesebb, csak az előbbiekből hoznak létre perceptuális kategóriát. Így az 5 hónapos csecsemők egy kategóriába tartozónak ítélték a különféle léptékfüggetlen hangokat (amelyeket a felnőttek más-más fajta vízként írtak le), míg a léptékfüggetlenségtől eltekintve hasonló paraméterekkel rendelkező változó léptékű hangokat nem. Ez azt jelenti, hogy a léptékfüggetlenség – a természetes ingerek egyik fontos jellemzője – valóban kitüntetett szerepet játszik a hallásban és a hangingerek kódolásában.

A léptékfüggetlenség a beszélt nyelvre is jellemző. Ezt mind a beszéd energiaspektrumának elemzése (Voss–Clarke 1975), mind azon pszicholingvisztikai eredmények is alátámasztják, melyek szerint a beszéd időben léptékfüggetlen, mert a beszéd bizonyos mértékű felgyorsítás és lelassítás ellenére érhető marad (Pallier et al. 1998; Sebastián-Gallés et al. 2000), s ez az adaptáció valószínűleg a hangalak, nem pedig az absztrakt nyelvi tudás szintjén történik, mert újszülöttek is képesek adaptálódni a felgyorsított beszédhez, legalábbis az anyanyelvükön

(Issard–Gervain 2016), a felnőttek pedig nem csak anyanyelvükön, hanem számukra ismeretlen, bár az anyanyelvükhöz ritmusában hasonló nyelveken is képesek erre (Pallier et al. 1998).

#### 4. Összegzés és jövőbeli kutatási irányok

A fentebb leírt komputációs és pszicholingvisztikai kísérletek alátámasztani látszanak azt a feltételezést, miszerint a hallórendszer – a látórendszerhez hasonlóan – olyan matematikai szabályszerűségeknek engedelmessé válik, amelyek lehetővé teszik számára, hogy a természetes hangokat, amelyekhez az evolúció során adaptálódott, hatékonyan és optimálisan képezze le. Az eredményekből az is kirajzolódik, hogy a beszélt nyelv – mint a hallás evolúciója során viszonylag későn megjelenő hanginger – olyan akusztikai tulajdonságokkal rendelkezik, amelyek a lehető legjobbban illeszkednek a hallórendszer által használt idegi kódhoz. Ilyen szempontból tehát a beszéd különleges, optimális hanginger.

Ezek az eredmények új perspektívába helyezik a beszédpercepció és beszédmegértés kutatását. A beszédpercepciót ezen új szempontból vizsgálva felmerül a kérdés, pontosan hol és milyen módon kódolódnak a beszéd különböző, statisztikailag releváns, nem redundáns tulajdonságai a hallórendszerben.

A hatékony kódolás modellje azt a kérdést is felveti, hogy a hatékony idegi kód változik-e, és ha igen, hogyan az egyed fejlődése során. Hatással van-e, és ha igen mennyire, a hallórendszer fejlődésére az anyanyelv vagy például a számottevő zenei tapasztalat.

#### Hivatkozások

- Atneave, Fred 1954. Some informational aspects of visual perception. *Psychological review* **61/3**:183–193.
- Barlow, H. B 1961. Possible principles underlying the transformation of sensory messages. In Walter A. Rosenblith (szerk.) *Sensory Communication*. Massachusetts: The MIT Press. 217–234.
- Cutler, Anne 1994. Segmentation problems, rhythmic solutions. *Lingua* **92**:81–104.
- Cutler, Anne – Katherine Demuth – James M. McQueen 2002. Universality versus language-specificity in listening to running speech. *Psychological Science* **13/3**:258–262.
- Geffen, Maria N. – Judit Gervain – Janet F. Werker – Marcelo O. Magnasco 2011. Auditory perception of self-similarity in water sounds. *Frontiers in Integrative Neuroscience* **5**:15.
- Gervain, Judit – Janet F. Werker 2013. Prosody cues word order in 7-month-old bilingual infants. *Nature Communications* **4**:1490.
- Gervain, Judit – Janet F. Werker – Alexis Black – Maria N. Geffen 2016. The neural correlates of processing scale-invariant environmental sounds at birth. *NeuroImage* **133**:144–150.
- Gervain, Judit – Janet F. Werker – Maria N. Geffen 2014. Category-specific processing of scale-invariant sounds in infancy. *PLoS ONE* **9/5**:e96278.
- Guevara Erra, Ramón – Judit Gervain 2016. The efficient coding of speech: Cross-linguistic differences. *PLoS ONE* **11/2**:e0148861.
- Issard, Cécile – Judit Gervain 2016. Adult-like processing of time-compressed speech by newborns: A NIRS study. *Developmental Cognitive Neuroscience* **megj. előtt**.
- Lewicki, Michael S. 2002. Efficient coding of natural sounds. *Nature Neuroscience* **5/4**:356–363.
- Liberman, A. M. – F. S. Cooper – D. P. Shankweiler – M. Studdert-Kennedy 1967. Perception of the speech code. *Psychological Review* **74/6**:431–461.
- Morgan, James L. – Katherine Demuth (szerk.) 1996. *Signal to Syntax. Bootstrapping from Speech to Grammar in Early Acquisition*. Hillsdale: Lawrence Erlbaum Associates.

- Nespor, Marina – Mohinish Shukla – Ruben van de Vijver – Cinzia Avesani – Hanna Schraudolf – Caterina Donati 2008. Different phrasal prominence realizations in VO and OV languages. *Lingue e linguaggio* **7/2**:139–168.
- Pallier, Christophe – Nuria Sebastian-Gallés – Emmanuel Dupoux – Anne Christophe – Jacques Mehler 1998. Perceptual adjustment to time-compressed speech: A cross-linguistic study. *Memory & Cognition* **26/4**:844–851.
- Sebastián-Gallés, Núria – Emmanuel Dupoux – Albert Costa – Jacques Mehler 2000. Adaptation to time-compressed speech: Phonological determinants. *Perception & Psychophysics* **62/4**:834–842.
- Shannon, Claude Elwood 1948. A mathematical theory of communication. *Bell System Technical Journal* **27/3–4**:379–423, 623–656.
- Simoncelli, Eero P. – Bruno A. Olshausen 2001. Natural image statistics and neural representation. *Annual Review of Neuroscience* **24/1**:1193–1216.
- Smith, Evan C. – Michael S. Lewicki 2006. Efficient auditory coding. *Nature* **439/7079**:978–982.
- Stilp, Christian E. – Michael S. Lewicki 2013. Statistical structure of speech sound classes is congruent with cochlear nucleus response properties. *Proceedings of Meetings on Acoustics* **20/1**:50001.
- Tomasello, Michael 2000. Do young children have adult syntactic competence? *Cognition* **74/3**:209–253.
- Voss, Richard F. – John Clarke 1975. „1/fnoise” in music and speech. *Nature* **258/5533**:317–318.
- Werker, Janet F. – Takao K. Hensch 2015. Critical periods in speech perception: New directions. *Annual Review of Psychology* **66/1**:173–196.