

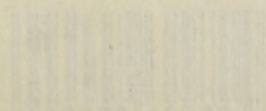
LINGUISTICA
SERIES A
STUDIA ET DISSERTATIONES, 2.

BESZÉDÉSZLELÉS

GÓSY MÁRIA

BESZÉDÉSZLELÉS

MTAK



9 80102 22970 4

189863

LINGUISTICA

SERIES A

STUDIA ET DISSERTATIONES, 2.

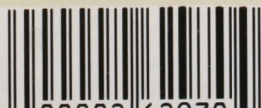
GÓSY MÁRIA:

BESZÉDÉSZLELÉS

A MAGYAR TUDOMÁNYOS AKADÉMIA NYELVTUDOMÁNYI INTÉZETE
INSTITUTUM LINGUISTICUM ACADEMIAE SCIENTIARUM HUNGARICAE

1989

MTAK



0 00002 62970 4

670021

Lektorok:

Götze Árpád

Pléh Csaba

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

ISBN 963-8461-35-7

ISSN 0238-8642

Copyright az MTA Nyelvtudományi Intézete, Budapest, 1989

Felelős kiadó: Herman József
Hozott anyagról sokszorosítva.

8918701 MTA Sokszorosító, Budapest. F. v.: dr. Héczey Lászlóné

M. TUD. AKADÉMIA KÖNYVTÁRA
Könyvtár 6327/10 89. sz.

TARTALOM

BEVEZETÉS	7
AZ ÉSZLELÉSI--MEGÉRTÉSI FOLYAMAT <i>B. felad</i>	14
A percepciós bázis <i>5.</i>	16
Beszédmegértési modellek <i>6.</i>	19
A féltreke-dominancia <i>2.</i>	28
A vizuális észlelés hatása a beszédpercepcióra <i>5.</i>	30
Az elemi percepciós egység <i>5.</i>	32
Az invariancia jelensége <i>5.</i>	36
A beszédmegértési folyamat alsóbb szintjei	47
A hallásról <i>4.</i>	49
Az emberi hallás vizsgálata <i>4.</i>	52
Akusztikai szint <i>4.</i>	57
Fonetikai szint <i>4.</i>	60
Fonológiai szint <i>4.</i>	63
Folyamatos szövegből szegmentált magánhangzók észlelése	68
MAGYAR BESZÉDHANGOK AZONOSÍTÁSA	74
Felismerési kulcsok	74
A magánhangzók percepciós vizsgálata	77
A frekvenciaszerkezet szerepe a felismerésben	78
A frekvencia- és időtartam-változás összefüggése a magánhangzók azonosításában	100
A mássalhangzók percepciós vizsgálata	107
A felpattanó zöngés zárhangok azonosítása	108
A nazális mássalhangzók felismerése	121
A frekvenciaszerkezet szerepe az azonosításban	122
A második formáns hatása a felismerésre	123
A [s, z, ʃ, ʒ] mássalhangzók felismerése	128
A felismerést biztosító frekvenciatartomány meghatározása	128
A zörejfrekvencia változásának hatása a felismerésre	132
A frekvencia- és időtartam-változás összefüggése a [s, z, ʃ, ʒ] mássalhangzók azonosításában	138

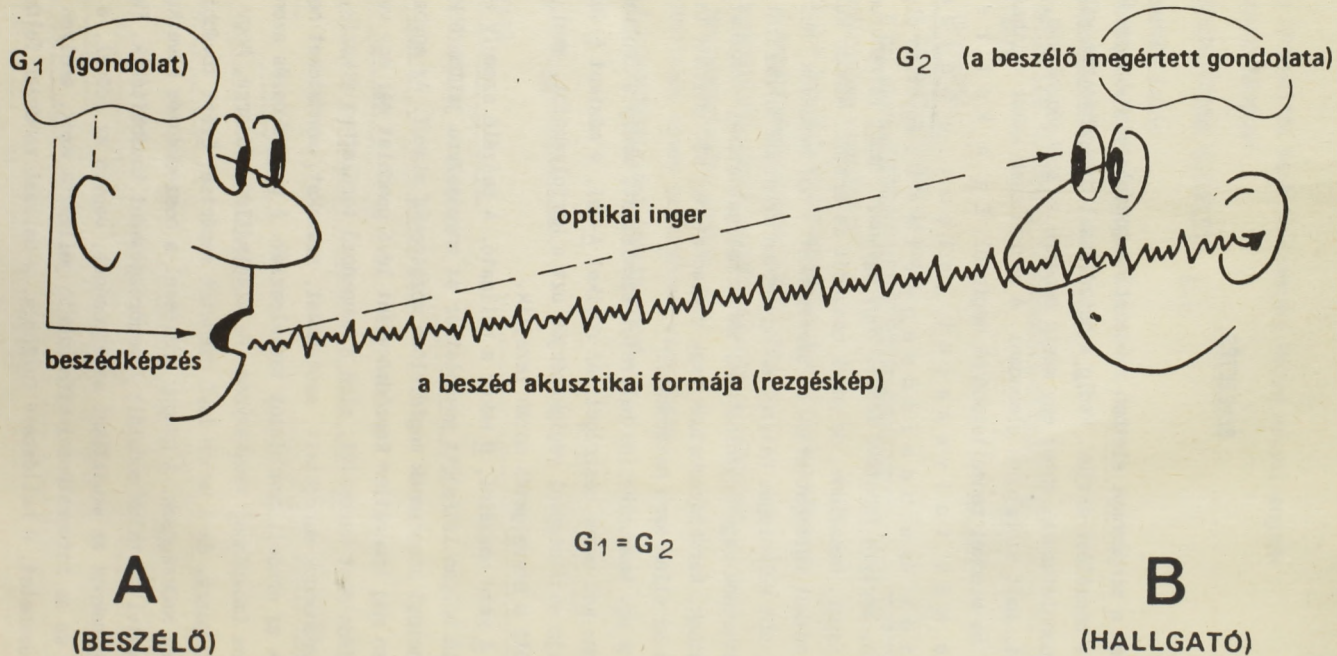
Az intenzitás mint felismerési kulcs mássalhangzók azonosításában	142
SZUPRASZEGMENTUMOK ÉSZLELÉSE 8.)	149
Tempóészlelés	149
A beszéddallam észlelése	164
Hangsúlyészlelés	174
10. A BESZÉDÉSZLELÉS FEJLŐDÉSE AZ ANYANYELV-ELSAJÁTÍTÁS FOLYAMATÁBAN	181
Bevezetés	181
A beszédészlelés kezdetei	183
A beszédészlelés fejlődése óvodáskortól	194
A beszédészlelés vizsgálata	201
A G-O-H hallást és beszédészlelést vizsgáló rendszer	203
A mesterséges szavak előállítás	203
A mesterséges szavak szűrőfunkciójának működése	205
Beszédészlelési adatok	220
11. A beszédészlelés egyes szintjeinek vizsgálata	223
A beszédészlelési tesztorozat	225
BEFEJEZÉS	233
IRODALOM	234
TÁRGYMUTATÓ	249
SPEECH PERCEPTION -- SUMMARY	255

BEVEZETÉS

1. Beszédén a mindennapi életben rendszerint gondolataink hangos közlését, vagyis a beszélést értjük. Pedig a beszéléssel önmagában nem valósulhat meg a gondolatcsere: ehhez egy másik emberre is szükségünk van, aki megérti mindazt, amit szeretnénk elmondani. A beszéd csak akkor töltheti be a szerepét, ha mindkét rendeltetésének megfelel. Ez a két rendeltetés két folyamatot jelent: a beszédprodukciót és a beszéd megértést. Egyetlen olyan képességünk van, amelyik lehetővé teszi, hogy ugyanazt, amit létrehozunk, fel is tudjuk fogni. Beszélünk, de mások beszédét is értjük. Minden különösebb nehézség nélkül váltogathatjuk a "szerepünket": hol beszélők, hol hallgatók vagyunk, sőt különleges helyzetben egy időben tehetjük mindkettőt. Beszéd közben köthetünk vagy rajzolhatunk, zenét hallgathatunk, szagokat, illatokat érezhetünk, kerékpározhatunk vagy szaladhatunk, sőt ehetünk is (ennek legfeljebb az illem szab határokat).

Az 1. ábra egy kommunikációs helyzetet szemléltet, amely a leggyakoribb az életben: két ember beszélget. Az egyiket A-val, a másikat B-vel jelöltük. A rajz segítségével végigkísérjük azt a körfolyamatot, amely a két ember között a beszélgetés során lezajlik.

Elsőként A kezd beszélni, B ekkor a hallgató. A beszélő személy agyában a megfelelő helyen létrejött gondolatsor az anyanyelvre jellemző képzési (artikulációs) programnak megfelelően mozgásokká alakul. Az agyban sok tekintetben még ismeretlen képződményként lévő gondolat már egy vagy több átalakuláson ment keresztül, mire a hangképző területhez érkezik. Itt a beszélő beszédhangok sorozatait, szótagokat, szavakat, mondatokat hoz létre, amelyek az eredeti gondolatot tartalmazzák. A beszédképzés azonban -- mint motoros (mozgásos) tevékenység -- alkalmatlan mód arra, hogy a hallgató feldolgozza, és a benne levő üzenetet megértse. Ezért további transzformációk szükségesek. A hangos közléssel a beszédképzés következményeként a nyelvi tartalom modulált levegőrezgésként továbbítódik. Ekkor ezeknek a rezgéseknek az akusztikai sajátosságai, vagyis az idővel, a frekvenciával és az intenzitással kifejezhető jellemzők azok, amelyek tartalmazzák az üzenetet. A hallószerv felfogja, a hallási rendszer feldolgozza



1. ábra

A beszéd körfolyamata: a kódolástól a dekódolásig

ezt a "kódolt" gondolatsort, végül a hallgató agyának megfelelő helyén megtörténik a megértés. Az egész folyamat emlékeztet arra, mint amikor egy üzenetet valamilyen okból rejtjeles formában kell továbbítanunk. A rejtjelezés, a közvetítés, a kódrendszer megfejtése, az eredeti szöveg azonosítása a megértési folyamathoz igen hasonló műveletek. Ha az A beszélőben és a B hallgatóban végbement működéssorozat tökéletes volt, akkor az eredeti és a megértett gondolat közé egyenlőségjelet tehetünk. Természetesen csak akkor, ha minden zavarmentesen folyt le, egyetlen átalakulás folyamán sem torzult el az eredeti üzenet. Zavaró körülmény pedig nagyon sok és sokféle adódhat:

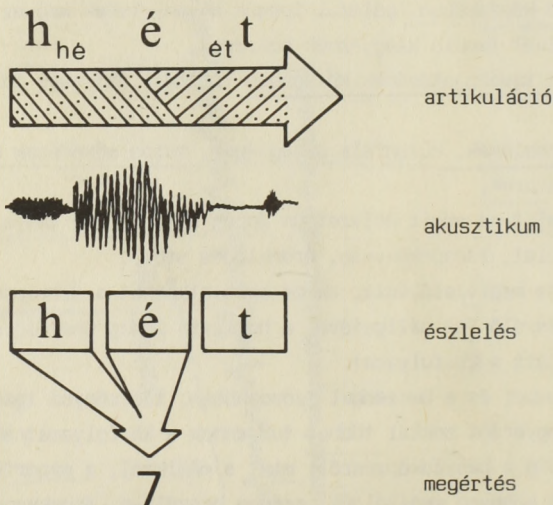
- (a) a beszéd képzésében (például idegen anyanyelvűek magyar kiejtése vagy az anyanyelvét tanuló kisgyermek beszéde),
- (b) a beszéd továbbjutásának közegében (például zaj vagy túl halk beszéd),
- (c) hallásproblémák, különféle betegségek, kóros működések következtében fellépő zavarok,
- (d) a hallgatót az adott helyzetben éppen befolyásoló tényezők (koncentráció, hangulat, fáradékonyság, érdeklődés stb.).

Ha a működés megfelelő volt, akkor bekövetkezhet a szerepcseré: a megelőzőekben beszélő fél hallgatóvá, a hallgató pedig beszélővé lesz. Ezzel létre is jött a körfolyamat.

A beszélő ember és a beszéddel szorosabban, tudományos igénnyel foglalkozó kutató egyaránt sokkal többet tud ennek a körfolyamatnak az egyik részéről, mégpedig a beszédképzésről, mint a másikról, a megértésről. A mások beszédét könnyen megítéljük: szépen beszélt-e, érdekesnek tűnt-e, könnyen érthető volt-e. Valamennyi ítéletünket annak alapján hozzuk meg, hogyan dolgozta fel a hallott közlést a beszédmegértési rendszerünk: de ezek mégis az elhangzottakra vonatkoznak. Valamivel nehezebb a helyzete az előadást tartónak, ha a hallgatósága megértési folyamatáról akar tájékozódni, mivel általában csak közvetett visszajelzések (arcmimika, gesztus stb.) alapján következtethet arra, hogy mondatait megértették-e, jól értették-e, sőt továbblépve: egyet értettek-e. A kutató azzal a nehézséggel találja szembe magát, hogy lényegesen könnyebb megfigyelni, vizsgálni azt, ami elhangzik, mint azt a folyamatot, amely ennek a hangzásnak a nyomán a hallgatóban végbemegy.

A beszédészlelés és beszédmegértés kutatása új terület, bár nem előzmény és hagyomány nélküli. Korábban a

pszichológia tudományterületén folytak a megértéssel kapcsolatos vizsgálatok, majd a nyelvészet, elsősorban a fonetika tűzte célul az észlelésnek, a beszédértésnek a kutatását. Ma mindenekelőtt a pszicholingvisztika foglalkozik a kérdéskörrel, mintegy egyesítve azokat a határterületeket, amelyeknek ismerete nélkül a beszédmegértés folyamata nem tanulmányozható. Az az igény, hogy a beszédet hármas egységben, a képzés, az akusztikum és az észlelés/megértés szempontjából vizsgáljuk, alig néhány évtizeddel ezelőtt merült föl (Fischer-Jørgensen 1957, vö. Hegedűs 1958, 205). A 2. ábra a hét szóval ezt az összefüggést szemlélteti.



2. ábra

Az artikuláció, az akusztikum, az észlelés és a megértés összefüggése

A magyar beszéd megértési sajátosságainak rendszeres alap kutatása éppenhogy megkezdődött, a korábbi szórványos adatok és eredmények azonban jó alapot adnak a további vizsgálatokhoz.

2. A múlt ismerete mindig hozzásegít a jelen értékeléséhez, a jövő megtervezéséhez. Amikor most mintegy 200 évnnyit visszalépünk

az időben, ezért tesszük. Akihez eljutunk, az Kempelen Farkas, akinek neve nem ismeretlen az utókor előtt, elsősorban alakjának regényes irodalmi feldolgozása (Szalatnai Rezső: Kempelen, a varázsló című könyve) és a híres sakkozó török, az "automata sakkjátékos" világhíre következtében. Kempelen volt az első, aki olyan mechanikus rendszerű beszélőgépet szerkesztett, amellyel magyar, francia, olasz és német nyelven szavakat és mondatokat lehetett kimondatni, a szerkezet részeinek megfelelő mozgatásával (1971). Német nyelven megjelent könyvéből idézzük: "A türelem, amivel a munkát folytattam, leírhatatlan. ...De a meggyőződésem, hogy a beszédnek utánozhatónak kell lennie, megerősített engem" (42). A több, mint egy évtizedes kitartó munka közben Kempelen számos, a megértéssel kapcsolatos megfigyelést tett. Ezeknek a legnagyobb része megfelel a valóságnak. Megállapítja például, hogy a hallás nagyon alkalmazkodó, ha a felismerést a jelentés is segíti. Ha az ember tudja, hogy mit akart a gép mondani, akkor könnyen megérti, és hajlandó azt hinni, hogy úgy is hallotta. A mássalhangzók közül a [p] már szólt, de a [d, t, g, k] előállítására gondot okozott. Ekkor írta, hogy a szerkezet működtetésében olyan kis különbségeket kell majd gyakorlással létrehozni, hogy "rászedje a fület, hogy a hallgató a hangot K-nak vagy T-nek ismerje fel, bár alapjaiban ott is P lesz található" (442). Megsejtett valamit a hangátmenetekről is. Leírja, hogy bármilyen kicsi szüneteket tartott az egyes hangok között, azt mindig szünetként lehetett észlelni; ha viszont nem voltak szünetek, akkor összefolytak a hangok, értelmetlenek lettek a szótagok. A hangátmenetek jelentőségéről, vagyis arról, hogy két beszédhang között rövid idejű átmeneti részek vannak, amelyek a két különböző beszédhangot illesztik egymáshoz, ugyanakkor csak meglehetősen rövid ideje tudunk: a huszadik századi mesterséges beszédelőállítások és megértési kísérletek eredményeként. Siketnémákkal történt találkozása ébreszti rá a szájról olvasás jelentőségére normál beszélynél is: "amit a fül elszalaszt [ti. a megértésből], azt a szem fogja befejezni" (23).

A múlt század elejéről származik a "Tapasztalásbeli emberlélekértékesség" című kéziratos lélektan, amelyben a szerző, Szűts István a hallással, a percepcióval is foglalkozik. Meghatározásában a hallás a levegő által továbbított rezgés érintése, felfogása szerint ugyanis végeredményképpen minden érzékelés a tapintásra megy vissza. Csírájában megtalálható itt a megértésnek egy sajátos értelmezése, amely szerint a hallási képzetek megőrzése és felújítása nem más, mint egy ingerre keletkezett mozgás és e

mozgás megújítása (Joó 1829/3-4, 90--100).

Csaknem 80 évvel később jelenik meg Stricker Salamon orvosprofesszor német nyelven írt könyve, amely valószínűleg elsőként tekinti át a beszéd-megértés mechanizmusát. Állítja, hogy nem minden egyes hangot fogunk fel, hanem a szó néhány jellegzetes része alapján következtetjük ki a teljes szót, hasonlóképpen néhány szó alapján a mondat értelmét! Kiemelésre méltó, hogy leírásában egészen modern felfogásban kapcsolja össze a megértést az artikulációval (1880).

Újabb jelentős állomás Sarbó Artúrnak, az idegkórtani diagnosztika magántanárának könyve (1906). A beszédmegértés az ő felfogásában is szoros kapcsolatot tart a beszédmozgató központtal, illetőleg a hangképzéssel. Az artikulációnak, a motoros komponensnek a részvétele a megértésben sokkal később, Liberman és munkatársainak kidolgozásában jelentkezett újra, s e formájában a modern elektroakusztikai vizsgálati eredmények következményeként teljesebben ki (1963).

Visszatérve a magyar szakirodalomra, két adat tudománytörténeti érdekességként kívánkozik rövid említésre. A XX. század elején nyelvészeti munkában még mindig nem olvashatni a beszédpercepcióról, ezért tartom jelentősnek, hogy két különböző folyóirat olyan műveket ismertet, amelyek e témát érintik, feldolgozzák. Az egyik a Nyelvtudományban jelent meg 1907-ben, a kitűnő Balassa József tollából, aki Scripture könyvét mutatja be, amelynek egy fejezete a hang felfogásával és megértésével foglalkozik. 1910-ben a Gyógypedagógia című folyóirat közli folytatásokban Wallis XVII. századi fonetikai munkáját, amelyik a beszédmegértés kialakulását, fejlődését, a fül szerkezetét tárgyalja.

A pszichológus Ranschburg Pál Az emberi elme című 1923-ban kiadott művében érinti az appercepció kérdését; mai kifejezéssel a megértés nyelvi meghatározottságáról értekezik.

1929-et fordulópontnak is nevezhetjük, mivel jelen ismereteink szerint Kempelen óta először történik meg, hogy fonetikai munka beszédpercepciót is tárgyal. Ez a könyv Horger Antal Általános fonetikája, amelynek egyik fejezete "A percipiáló szervek és működésük" címet viseli. Horger foglalkozik a fül működésével, a hallással, valamint az észlelésben szerepet játszó agyi és idegi folyamatokkal is. Felfogásában a levegőhullám a dobhártyán át a hallóideg közvetítésével jut a hallásközpontba, itt hangsorérzetté változik, majd a megértő központba jut, ahol a fogalom tudatosodik. Megemlíti, hogy a szemnek is van szerepe a beszédhangok per-

cepciójában! A motoros beidegzések, a hangképzés részvételéről nem tesz említést. (Ide kívánczik, hogy a 3 évvel későbbi, 1932-es első nemzetközi fonetikai kongresszus tárgyköreiben a beszédmegértés még nem szerepel.)

A Nobel-díjas Békésy György a hallás biofizikai problémáit kutatta, a 30-as években a beszéddel kapcsolatosan is végzett kísérleteket. Bizonyította, hogy ha a szótagok intenzitása túlnó a normál társalgási hangszinten, akkor érthetőségük csökken. Meghatározta a beszédpercepció szempontjából fontos "pszichológiai jelen" időtartamát (1960).

Érdeemes szó szerint idézni Hegedűs Lajos 1935-ből való sorait: "A hangok fizikai természetének megismerése után az is érdekel bennünket, hogyan jó létre a hang, mégpedig mind a beszélőben, mind a hallgatóban" (337). Mások kísérleteire hivatkozva állítja, hogy a magánhangzók felismerését két formáns dönti el. Részletesen foglalkozik a csiga felépítésével, a hangérzet keletkezésével.

Laziczius Gyula 1944-es Fonétikájában az artikuláció és az akusztikum mellett helyet kapnak a megértéshez tartozó témák is, mint a fül működése és a hallási folyamat. A kérdést azonban inkább fiziológiai, mint nyelvészeti-fonetikai problémaként mindössze 14 lapon a beszédhangok akusztikájának fejezetén belül tárgyalja. Az idegingerületek további sorsáról úgy nyilatkozik, hogy efelől a tudomány még mit sem tud (140).

Az utóbbi negyven évben a fonetikusokon, nyelvészeken kívül számos más tudomány képviselője foglalkozott a beszédpercepció különböző szintjeivel: orvosok, akusztikusok, pszichológusok, pszicholingvisták, gyógy-pedagógusok -- ez jelzi is egyben a téma érintkezési területeit, interdiszciplináris voltát. Örvendetes tény az is, hogy a nagyobb összefoglaló munkák, mint például Szende Tamás könyve, A beszédfolyamat alaptényezői (1976), Vértés O. András A magyar leíró hangtan története az újgrammatikusokig című műve (1980), az 1982-es Fejezetek a leíró magyar hangtanból című tanulmánykötet Bolla Kálmán szerkesztésében foglalkoznak a beszédmegértés kérdéseivel. Önálló kötetet írt e témáról Pauka Károly "Halláslélektan -- a beszédmegértés alaptényezői" címmel (1980). A megjelent számos tanulmány közül csupán néhányat említek. Hegedűs Lajos 1952-ben hallástani ismereteit szubjektív lehallgatásos eredményekkel egyezteteti, és így foglal állást a hangsúlyérzet kialakulásának kérdésében. Hat évvel később, 1958-ban Fónagy Iván a hangsúlyt elemző kismonográfijában döntő szerepet juttat az észlelésnek. "A hangsúlyélmény -- írja -- tehát nem a hallószervben, hanem a központi idegrendszerben ala-

kul ki a hangjelenség alapos, de gyors és öntudatlan elemzése nyomán" (21). Percepciós kísérlettel járul hozzá Hegedűs Lajos az affrikáták egyes- vagy kettőshang voltának eldöntéséhez. Magnetofonszalagra rögzített, zár-rés hangot tartalmazó szavakat visszafelé játszott le a hallgatóknak, akik az így keletkezett hangsorokat az elhangzást követően leírták (1958). Elsősorban orvosi vonatkozású eredmény: Götze Árpád 1960-ban összeállította a magyar beszédaudiometria szám- és szósorait, ezzel a beszédhallás-küszöb és a beszédérthetőség fokának klinikai vizsgálhatóságát teremtette meg. A magyar szakirodalomban addig nem alkalmazott új módszerrel, szintetizált magánhangzókkal vizsgálta Magdics Klára azok felismerési sajátosságait 1970-ben. Megállapítja, hogy legnagyobb jelentősége a formánsfrekvenciának van, az amplitúdó és az alaphangmagasság másodlagosak. Tarnóczy Tamás és Vicsi Klára -- más kísérleteik mellett -- a [p,t,k] mássalhangzók mikroszegmenseinek, ti. a magánhangzó-átmeneteknek, a szünetnek és a felpattanási zörejnek, illetőleg a zörej időtartamának a felismerésében játszott szerepét vizsgálták. Természetes és mesterséges hangkapcsolatokkal végzett kísérletek egyaránt igazolták, hogy a résmássalhangzók időtartamának csökkenése a hangminőség változását eredményezi (1979).

Grammatikai, szintaktikai sajátságoknak a megértésben játszott szerepével, valamint a jobb-fül dominanciával foglalkozik Pléh Csaba több dolgozata (1981a; 1981b; 1981c; 1984).

A Kempelen utáni csaknem kétszáz esztendő vázlatos áttekintése azt mutatja, hogy még a beszédészlelés területén is milyen sokat meríthet a mai kutatás a régi s az újabb kutatók eredményeiből; a múlt meglepően időtálló ismereteket és megsejtéseket őrző hagyatéka ösztönzést, útravalót és alapot ad a további munkához.

AZ ÉSZLELÉSI--MEGÉRTÉSI FOLYAMAT

A beszédmegértés folyamata két nagy szakaszból áll: 1. a nyelvi kódok (jelek) rendszerének megfelelő hangjelenségek észlelése és 2. ennek a kódrendszernek az értelmezése (Lurija 1979, 291--292). Mindkettő több fokozatból (szintből) épül fel, amelyek törvényszerű együttműködésben

biztosítják a hallott beszédjelenségek megértését. A beszéd megértése aktív folyamat, amelynek során a hallgató a fülével érzékelt hangjelenségeket (=beszédjelenségeket) értelmezi. (Érdekes, hogy Saussure és mások is a megértést passzív folyamatnak vélte a produkció aktív folyamatához képest: 1967, 29--30.) Az első nagy szaka a beszédhangok, a hangkapcsolatok felismerését jelenti, azt tehát, hogy azonosítani tudjuk a [b, p, v, u, e] hangokat, az [om, on, ong] vagy [sɛ:] hangkapcsolatokat, vagy hogy a magyar anyanyelvűek az [o] és az [a:] között egy [ɔ]-t, illetve a [d] és a [g] között egy [ʒ]-t is meg tudnak különböztetni. Az észlelési rendszer azokat a különbségeket ismeri fel, amelyeket a nyelv használ. A beszédmegértés másik nagy szaka a "kódrendszer értelmezése", vagyis a szavak és a mondatok megértése. Ez a szakasza a folyamatnak már a jelentés felismerését is tartalmazza, ezért már nem észlelésnek (azonosításnak), hanem megértésnek nevezzük. Például, amíg egy elhangzó zó hangkapcsolatról annyit azonosítunk, hogy a [z] mássalhangzó és az [ø:] magánhangzó kapcsolata, addig az ugyanezen beszédhangokból álló öz szótagról mindennek-előtt azt állapítjuk meg, hogy azonos az erdőben élő, "a szarvassal rokon, de kisebb, kecses termetű kérődző állat"-tal; vagyis megértjük a jelentését. (Külön kérdésre persze be tudunk számolni arról, hogy ez a felismerés milyen hangélmény alapján jött bennünk létre.)

A beszédmegértés komplex folyamat, amelyben -- normális körülmények között -- nem válik szét a beszéd értelmes és értelem nélküli szegmensekre vagy szegmentális és szupraszegmentális részre. A folyamat működésének jelentős részéről gyakorlatilag nincsenek vagy alig vannak egzakt ismereteink: sejtjük csak, milyen működések történnek az idegpályákon vagy az agyban. A meglévő kísérleti eredmények alapján jobbára csak következtetünk ezekre.

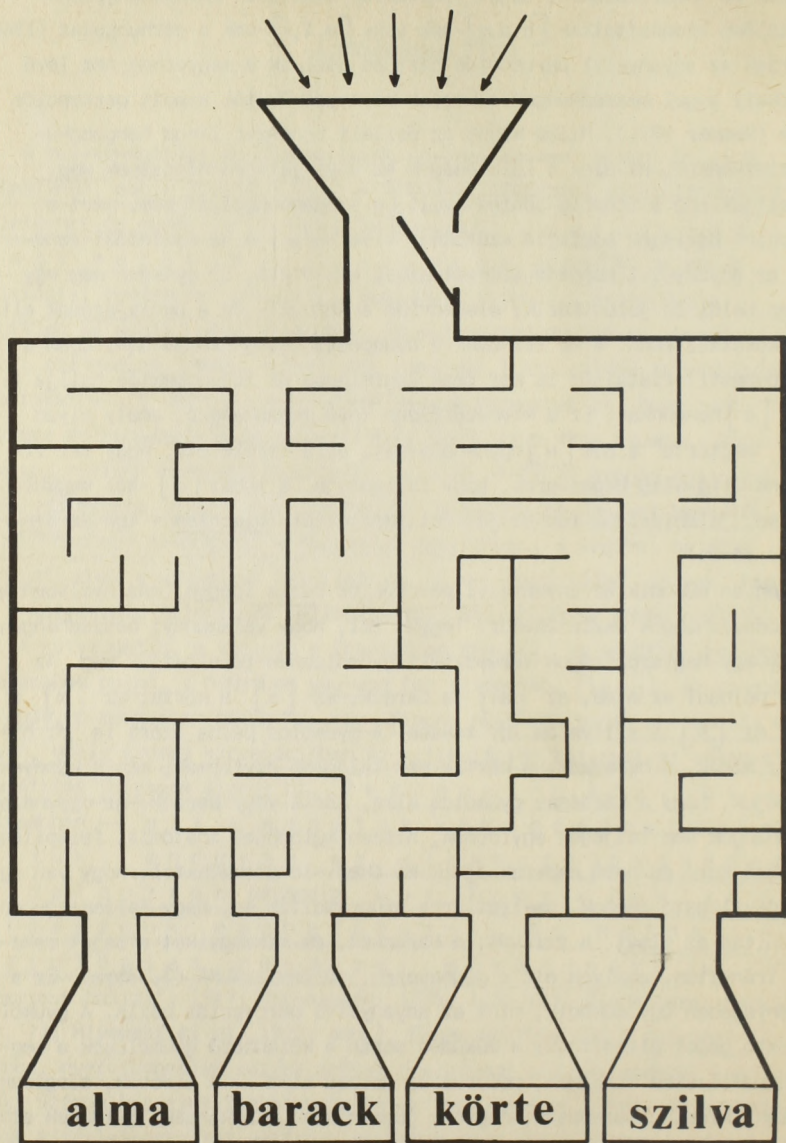
A percepciókutatásban a beszédpercepció sajátos helyet foglal el, mivel -- mint a bevezetésben említettük -- csak a beszéd esetében képes a "fogadó" (hallgató) ugyanazt a jelet kiadni, amely hozzá érkezett (Hörmann 1971, 68). A beszédmegértés a teljes kommunikációs lánc része.

A percepciós bázis

A beszédmegértés fejlődése során (gyermekkorban) alakul ki az a mechanizmus, amely lehetővé teszi a folyamat működését az adott nyelv észlelésére és megértésére. Ezt a mechanizmust percepciós (vagy megértési) bázisnak nevezzük. Az ember fiziológiai rendszere, amellyel a beszédet létrehozza és felfogja, független az adott nyelv típusától. Ugyanazon beszédszervek (hangszalagok, szájpadlás, nyelv, ajkak, fül, agy stb.) szolgálnak az eszkimónak, az afrikainak, a csehnek, a kínainak vagy a magyarnak is, hogy hangképzésre és hangészlelésre felhasználja őket. Az a tény azonban, hogy az egyes nyelvek hangzása különböző, mutatja, hogy az azonos anatómiai felépítés mégsem jelent azonos működést. A különböző nyelvet beszélők különbözőképpen működtetik a hangképző apparátust, s ezáltal sajátos jelleget adnak anyanyelvük hangzásának. Az ún. hangutánzó szavak (tehát a természeti jelenségek felfogása is?) különböző az egyes nyelvekben. A kutyaugatás, a vau-vau a franciában oua-oua, az angolban bow-bow, az olaszban bu-bu, a németben wau-wau (az ejtése kb. [va:u]), az oroszban gáv-gáv, a mongolban pedig háu-háu. Minden nyelvnek saját percepciós bázisa van, amely szoros korrelációban az artikulációs bázissal az anyanyelv-elsajátítás folyamán alakul ki. Percepciós bázison -- pontosabban -- azt a sajátos működéssorozatot értjük, ahogyan az elhangzó beszédet feldolgozzuk; mégpedig úgy, hogy a nyelvi sajátosságok meghatározók, és hatnak a fiziológiai rendszer működésére (vö. Lurija: "fonemikus hallás": 1947; Maruszewski 1975, 125--127; Glucksberg--Danks 1975, 27; Miyawaki et al. 1975; stb.). Ennek megfelelően a folyamat általános, nyelvfüggetlen vázára erősen nyelvfüggő szabályrendszer kapcsolódik. Számos kísérlet foglalkozik ennek igazolásával. Lotz és munkatársainak kísérletsorozatát említjük elsőként ezek közül, mivel a kísérleti személyek között magyar anyanyelvűek is szerepeltek. A mássalhangzó-felismerés vizsgálata volt a cél; a mássalhangzók zöngés, illetve zöngétlen zárhangként történő azonosíttatásával CV típusú hangkapcsolatokban. Az eredmények szerint az eltérő anyanyelvű kísérleti személyek (angolok, spanyolok,

magyarok és thaiföldiek) a saját nyelvükben realizált mássalhangzóknak megfelelően azonosították [b,d,g]-nek vagy [p,t,k]-nak a zárhangokat (1960). Ugyancsak az anyanyelvi percepciós bázisra utalnak a magyarban nem lévő (amerikai) angol beszédhangokra magyar anyanyelvűektől kapott percepciós adatok (Nemser 1971). Hiába képes az észlelő rendszer finom hangmégkülönbötésekre, ha ezek a különbségek az anyanyelvben nincsenek meg. Nem halljuk meg a francia beszéd nyílt [ɔ] magánhangzóját sem, mert a percepciós bázisunk egyfajta szűrőként viselkedik, s az észlelési rendszer az anyanyelvi hangzás irányításával működteti. Ez gyakran még egy nyelven belül is jelentkezik; elsősorban a köznyelv és a nyelvjárások eltérései következtében. A ma felnövekvő budapesti (és jelentős hányadban a nem budapesti) fiatalság is már csak legfeljebb az idősebbektől hallja (?) a zárt [e] hangunkat. Ez a kiveszőfélben lévő magánhangzó, amely olyan szépen "megtörte" a sok [e] egyhangúságát, arra íteltetett, hogy mai ifjúságunk alig-alig képes arra, hogy felismerje, a többi [e] -től megkülönböztesse, jelentéstani funkcióját felismerje (pl. szerelem - ige és sze-relem - főnév).

Hogyan működik az anyanyelvi percepciós bázis idegen (más nyelvben meg-lévő) beszédhangok észlelésekor? Tegyük fel, hogy valamennyi beszédhangunkat egy-egy Magyarországon termesztett gyümölcsnek feleltetjük meg. Az legyen például az alma, az [a:] a barack, az [o] a körte, az [m] a málna, az [s] a szilva és így tovább. A gyümölcs példa azért is jó, hiszen az almák, a barackok, a körték sem teljesen egyformák, mégis könnyen megmondjuk, hogy a kérdéses gyümölcs alma, körte vagy barack. Az ugyanazon beszédhangok sem teljesen egyformák, hiszen különböző anatómiai felépítésű, különböző korú és nemű emberek ejtik ki őket. Tételezzük fel, hogy van egy gyümölcsválogató gépünk, amelyet arra programoztak be, hogy felismerje és elkülönítse az almát, a szilvát, a barackot, és mindegyiket a saját rekeszébe irányítsa, amelyen ott a megnevezés, például: ALMA (3. ábra). Ez a gép lényegében úgy működik, mint az anyanyelvi percepciós bázis. A gyümölcs-felismerő gépet elindítjuk: a működés során a különböző gyümölcsök a megfelelő rekeszekbe kerülnek, például az összes alma -- a pirosak, sárgások, zöldek, kicsik, közepesek, nagyok, jonathánok, starkingok, goldenek stb. -- az ALMA felirat alá. A programozás tökéletes, a rendszer hibátlanul dolgozik. És ekkor váratlanul egy narancs jelenik meg a gyümölcsök között! Mi történik ekkor? A leggyakoribb az, hogy a gép megpróbálja -- a programnak megfelelően -- elemezni az ismeretlen gyümölcs jellemzőit. Ennek alap-



3. ábra

A gyümölcsválogató "gép" mint az anyanyelvi
percepció bázis analógiája

ján pedig abba a rekeszbe irányítja, amelynek sajátságai a legnagyobb hasonlóságot találta. Azaz a narancsot vagy az ALMA, vagy a BARACK felirathoz osztályozza. Pontosan az történik, mint az emberi percepció rendszerben; az ismeretlen hangot a hallgató valamelyik anyanyelvi hanggal azonosítja. (Például a román nyílt [ø] hangot a magyarok saját, zártabb [ø̄]-jeként azonosítják, így lesz Balcescu nevének értelése Bölcseszku, a spanyol flamenco-ból flamengó, az orosz felső nyelvállású, mediális, labiális magánhangzó, a jerü hol [y], hol [i] (ennek megfelelően ingadozik a nagy orosz meseíró nevének ejtése is: Krilov vagy Krülöv) vagy a francia pénz nevéből a franc-ból frank, az eredeti szóvégi orrhangú [ʒ] magánhangzó helyett.) Nem ritka azonban -- visszalépve a gyümölcsosztályozó gép példájához --, hogy létezik egy NARANCS feliratú rekesz is. Ekkor a rendszer minden nehézség nélkül különíti el a déligyümölcsöt az almáktól vagy a barackoktól. Ha az anyanyelvünkön kívül más nyelv(ek)et is ismerünk, a "fülünk" rugalmas marad (vagy rugalmassá válik) az újabb, a már ismerttől akár csak kismértékben különböző hangjelenségek befogadására és feldolgozására is. Az az ember, aki több nyelven beszél, annak több percepció bázisa van. Úgy is mondhatjuk, hogy észlelési-megértési rendszere bonyolultabban programozott.

Beszédmegértési modellek

A megértési folyamatot ún. modellekben próbálják ábrázolni, amelyekben lényegében egyszerűsítve reprezentálható a működéssorozat. A cél: annak bemutatása, hogyan milyen szinteken milyen jellegű működések történnek a megértés mechanizmusában. Pisoni szerint az alapkérdés a következő: hogyan konvertálja a hallgató a folyamatosan változó akusztikai jelet, amelyet a beszélő kiad, diszkrét nyelvi egységek sorozatává, és hogyan fogja fel a benne lévő üzenetet. Megjegyzi még, hogy mindez természetesen rengeteg alproblémát vet fel, amelyek megoldásra várnak (1981, 249). Ha feltételezzük, hogy a beszédet a megértés során bármiféle leírható mechanizmus dolgozza fel, akkor mindenekelőtt azt kell meghatározni, de legalábbis körvonaloznunk, hogy ez a mechanizmus milyen paraméterekkel jellemezhető és miként működik.

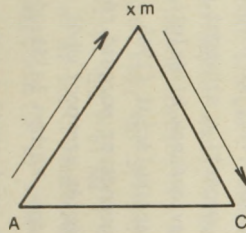
A beszédmegértés folyamatának modellálása nem újkeletű; az ezzel foglalkozó művek már egy évszázaddal korábban is próbálták valamilyen sémába

foglaltan értelmezni azt, amit a folyamatról tudtak, illetve elképzelték. A tudomány akkori érdeklődésének és állásának megfelelően az észlelést tekintették ugyan a kiindulásnak, sajátságait azonban kevésbé tárgyalták. Részletesebben a magasabb rendű (agyi) folyamatokkal foglalkoztak. Fontosnak tartjuk ismét megemlíteni Stricker Salamon nevét, aki több, mint fél évszázaddal korábban, mint hogy az elmélet elterjedt, hirdette a -- mai szakki-fejezéssel -- motoros beszédmegértési hipotézist (1880). Később Sarbó Artúr is ugyanezt a felfogást képviseli (1906). Az elmélet kidolgozása és kísérleti igazolása Liberman és munkatársai nevéhez fűződik 1957 után (valószínűleg teljesen függetlenül a múlt századi előzményektől. A 4. ábrán ezeket a feltételezett modelleket szemléltetjük.

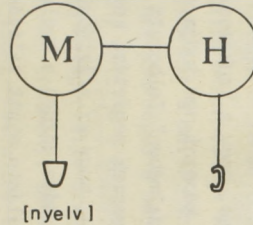
A ma ismert percepció modellek a teljes beszédmegértési folyamatot igyekeznek ábrázolni, továbbá kisebb-nagyobb mértékben foglalkoznak a folyamat működésmechanizmusával. Meglehetősen sok modellről tudunk (Flanagan 1965; Fant 1968; Schüller--Dévai--Kádár 1967; Glucksberg--Danks 1975; Lindner 1969; 1977; Pauka 1982; stb.), egyelőre nincs azonban olyan, amelyik maradéktalanul megoldaná az összes problémát, s igazítható lenne a különböző típusú nyelvekhez. Fant hat beszédészlelési modellt különít el: kibernetikai modell, neurológiai modell, nyelvészeti modell, analízis--szintézis (más megnevezéssel: aktív--passzív) modell, artikulációs modell és motoros modell (i.m.) Az egyes modellek közötti különbség mindenképp az elméleti felfogásban, a funkcionális szemléletben és a modell végcélja tekintetében jelentkezik. A ma számottevő modellek közös jellemzője, hogy nagymértékben hangsúlyozzák a nyelvi meghatározottságot -- ennek a felismerésnek azért óriási a jelentősége, mivel az ún. "általános percepció modell" csak elméleti úton hozható létre, s az egyes nyelvek megértési folyamatának reprezentálására alkalmatlan. Többnyire kikerülnek az a zsákutcát is, hogy a folyamat működésének egy vagy több részletét túlzottan hangsúlyozzák az összes többi rovására. Ilyen volt például a hallás: a hallásmechanizmus működésének szabályai nem alkalmazhatók minden esetben és módosítás nélkül a beszédészlelési folyamatra (ld. az elfedési jelenségek problémáját: Halm 1963; Pauka 1982; Denes--Pinson 1973; Tarnóczy 1984; stb.). A továbbiakban a legfontosabb modellkísérleteket ismertetjük, majd ezek és saját kísérleteink alapján megpróbálunk egy magyar nyelvi sajátosságokat tartalmazó hipotézist megfogalmazni.

1. A motoros teória. -- Az elmélet kidolgozása -- mint említettük -- az amerikai Liberman és munkatársai nevéhez fűződik (1957a; 1957b) a mel-

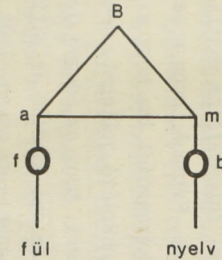
STICKER 1880



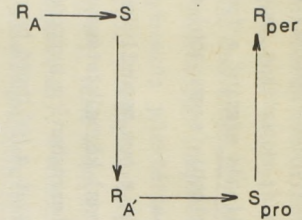
SARBÓ 1906



SIMONYI 1942



LIBERMAN 1957



4. ábra

A motoros komponens részvétele a megértésben különböző szerzők szerint
 (x=beszédhang-feldolgozás, m=motoros komponens, A=hangsor; C=jelentés;
 M=mozgató központ; H=hallómező; B=jelentés, a=hallásos kép, m=mozgásos kép,
 f=hallásos ingerfelfogás központja, b=mozgatásos beidegzési központ;
 R_A =artikuláció, $R_{\bar{A}}$ =immanens artikuláció, S=jel, per=percepció, pro=proprio-
 ceptíve létező jel)

lette felhozott érvek és ellenérvek csaknem napjainkig vita tárgyát képezik. Az elmélet lényege, hogy az agyban valahol a deskripció mindkét szintje (artikulációs és akusztikus) egyetlen mechanizmus formájában van kódolva, ebből következik, hogy a képzési és az akusztikus jegyeket a percepcióban is együttesen kell vizsgálni. A felfogás szerint a fonéma mint nyelvi entitás és az artikuláció között közelebbi (szorosabb) a kapcsolat, mint a fonéma és az akusztikai jel között.

Az akusztikai inger percepciója a nem megnyilvánuló, immanens artikulációs komponens (a motoros komponens) közbeiktatásával történik (vö. Hörmann 1971, 64--65). Más megfogalmazásban: a percepció különböző bonyolultságú szintjein az észlelő rendszer a hallható beszédet oly módon elemzi, hogy a szervezet saját beszédképzési folyamataira vezeti vissza. Az elmélet ellen számos érvet sorakoztattak fel. Például, hogy a beszédképzés hibája nem jelent feltétlenül percepció károsodást is; hogy az a képesség, hogy észleljük a beszédhullámok megkülönböztető jellegét, nem feltétlenül a beszédprodukciónk következménye. Ismert, hogy az idegen nyelvet általában könnyebb érteni, mint beszélni (s ez nem pusztán az aktív és a passzív szókincs különbségének a következménye); az anyanyelvét tanuló gyermek is előbb érti meg a hozzá intézett közléseket, minthogy ő maga képes lenne hasonlót létrehozni. Jakobson ellenvetése, hogy a kaukázusi nyelvek egy sor fonémáját meg tudja különböztetni, de képezni nem tudja azokat (Hörman 1971, 67); stb. A motoros elmélet "védelmében" felhozott bizonyítékok például, hogy az anyanyelvét tanuló gyermek megértési képessége más működésmechanizmuson alapszik; Lurija írja, hogy az afáziásnak lényegesen romlott a beszédértése, ha a beszéd hallgatása közben a nyelvét a fogaihoz kellett szorítania. Az immanens artikuláció miatt tiltják meg a hangszalaggyulladásal küszködő énekesnek, hogy zene- és énekszámokat hallgasson (!); avagy például az olvasni tanuló kisgyermek (vagy az idegen nyelven még bizonytalanul olvasó felnőtt is) jobban megérti a szöveget, ha hangosan vagy félhangosan kiejti a szavakat. Csisztovicséknak kísérletileg sikerült "létrehozni" és igazolni ezt a belső artikulációt (1965, 205). Ők úgy fogalmazznak, hogy a percepcióban az idegrendszer mozgató részei is szerepet kell, hogy játsszanak, mivel a beszédértés szempontjából aktivitásuk valamilyen módon fontos. A beszéd hallgatása közben lejátszódó imitációt Cherry írta le először (1953).

Min alapszik a libermani tétel? Megkülönböztetési kísérleteket végeztek felpattanó zárhangokkal ([b,d,g]) és magánhangzókkal (Lieberman et al.

1957a; 1957b). Az ABX és AX kísérletsorozat lényege az volt, hogy a szintetizált CV hangkapcsolatokban mindössze egyetlen paramétert változtattak, s így akusztikailag eltérő szekvenciákat hoztak létre. Ezek közül egyet vagy kettőt lejátszottak a kísérleti személyeknek, akiknek a második (vagy harmadik) hangkapcsolat elhangzása után el kellett döntenüik, hogy melyikek voltak hasonlóak (az ABX kísérletek esetében az első kettő közül melyik inkább a harmadikhoz, az AX esetében hasonlóak vagy azonosak voltak-e a jelek). Így kiküszöbölték a megnevezés jelentette nehézséget, és felhasználták egyszerűsre mind azt az észlelési törvényszerűséget, hogy mindig jóval több hangot tudunk megkülönböztetni, mint ahányat azonosítani, vagyis beszédjel esetében megnevezni (Flanagan 1965). A kísérletek eltérő eredménnyel zárultak a felpattanó zárhangok és a magánhangzók esetében. Az előbbieknél a fonémahatár közelében hirtelen átcsapást tapasztaltak, míg az utóbbiaknál az átmenet folyamatosnak mutatkozott (vö. például Stevens et al. 1969, 18). A magyarázat szerintük az adott beszédhangok képzési különbözőségében rejlik. A magánhangzók ugyanis folyamatosan alakíthatók egymásba az artikuláció során, míg a felpattanó zárhangok nem -- ez az eltérés jelentkezik a percepciójukban. A zárhangoknak a fonémahatár közelében tapasztalt hirtelen fonémaátcsapását nevezték kategoriális percepciónak, míg a "folyamatos" magánhangzókét nem kategoriálisnak (Lieberman et al. 1963). A mai álláspont az, hogy a kategoriális percepciónak valami sajátosan beszédjellegű sajátágként való felfogása már a múlté. Az újabb kísérletek azt mutatták, hogy se nem sajátos, se nem szükséges a beszédmegértéshez (Studdert-Kennedy 1980, 52). Számos tanulmány cáfolja a magánhangzók nem kategoriális felismerését, illetőleg bizonyítja kategoriális azonosításukat. A megfelelő időtartam megválasztásával azonos eredményt lehet felmutatni, mint a felpattanó zárhangok esetében (Fujisaki--Kawashima 1969; 1971; stb.). Fontos pontja a motoros elméletnek az átmeneti hangok kérdése. E tekintetben saját kísérletünk is bizonyítékkal szolgált (1982). A felpattanó zárhangok között is úgy léteznek beszédhangok, amelyek igazán sem egyik, sem másik (illetve akár mindkét) fonémaosztályba sorolhatók, mint a magánhangzók esetében. Ezeket a kísérleti személyek vagy úgy jelölik, hogy a mindkét fonémának megfelelő betűjelet leírják (bd vagy dg), vagy egy betűvel azonosítják ugyan, de az összesítésben az adott beszédhangra közel 50-50 % körüli felismerési arány adódik. Megpróbáltuk az artikuláció felől is ellenőrizni (ellen)feltevéssünket, hogy ti. a felpattanó zárhangokat -- természetesen a jellemző képzőmózzanatot megtartásával -- is lehet átmenetiként ejteni két fonéma-

realizáció között. Ezeket úgy hoztuk létre, hogy -- rövid begyakorlás után -- a nyelv és a szájpaddlás érintkezési pontját igen kis lépésekben változtattuk. Így több olyan hangot kaptunk, amelyek -- az azonosítás szempontjából -- a [b] és a [d] vagy a [d] és a [g] közé estek (vö. Gósy 1984e). Az erről az ejtéssorozatról készült hangszíneképek a második formáns finom mozgásáról--változásáról tanúskodnak (Gósy 1986, 12). Ez is azt támasztja alá, hogy a beszédhangok felismerése azonos módon történik, függetlenül a képzési sajátosságaitól. Egyúttal azonban nem tagadjuk az artikuláció részvételét a beszédmegértésben.

Az utóbbi évek kísérleti eredményei szerint a kategoriális percepció eredmények inkább a rövid idejű memória működésével kapcsolatosak, nem pedig a különböző beszédhangok eltérő felismerésének bizonyítékai (Pisoni--Sawusch 1975, 7). A motoros elmélet érvei tehát inkább logikaiak, mint empirikusan megalapozottak. A motoros teória modelljének legnagyobb hibája, hogy nem derül ki egyértelműen belőle, hol van a percepció elemzésnek az a szintje, ahol az artikulációs ismeret felhasználódik (Pisoni--Sawusch 1975, 8).

2. Az analízis szintézissel vagy aktív--passzív modell. -- Stevens és munkatársai javasolták a beszédmegértés modellálására (1960; Stevens--Halle 1967; Stevens--House 1972). Lényege, hogy a hallgató valamennyi beszédhangot saját maga számára generálja, "szintetizálja". Amikor hallja a beszédet, a rendelkezésére álló belső (artikulációs) rendszerrel megpróbálja létrehozni azt, amit hallott. A megértés akkor következik be, amikor a beszélő hangjelenségei és a hallgató belső, "szintetizált" jelei (a belülről generált képzetek, vö. Darwin 1976) megegyeznek. A folyamatban, ahol a fonetikai sajátosságok nem kontextusfüggőek, az elsődleges hallási elemzés során azonnal bizonyos összehasonlítások történnek a jel fonetikai ismérveivel. Ha nincs invariáns jegy, s az első valószínűsítési fonológiai mátrix hibás, akkor kiegészítő (újraanalizáló) folyamatok következnek. Ezek bizonyos szabályok szerint működnek, a cél ismét a bejövő jel és a "kontroll rendszer" megfeleltetése. E modell működésének igazolására Ladefoged és Broadbent kísérletsorozatát végeztek angol anyanyelvű kísérleti személyekkel. A kísérleti anyaghoz hatféleképpen szintetizálták a Kérem, mondja meg, melyik ez a szó! bevezető mondatot, majd a kísérleti személyeknek ugyancsak szintetizált CVC hangkapcsolatokat játszottak le.

Ezeket [bit], [bat], [bet] vagy [but]-ként kellett azonosítaniuk. A válaszok attól függtek, hogy a feladatot megjelölő mondat milyen volt. Ha relatíve magas hangon szólalt meg, akkor az első hangsort bit-nek érezték (87 %-ban). Ha mély hangon szólalt meg, akkor ugyanazt a hangsort bet-nek hallották (90 %-ban). A szerzők következtetése mindebből az, hogy a kísérleti alanyok a feladatmegjelölő mondatot voltaképpen a beszélővel azonosították, s ennek hatására ismerték fel különbözőnek az akusztikailag azonos hangsorokat.

Az "analízis szintézissel" elmélet alkalmas az izolált beszédhangok felismerési folyamatának szemléltetésére, és szolgálhat a globális beszédmegértés működési elvűl is. Állandó kapcsolatot tud tartani a megértés során az egyes beszédhangok és más nagyságrendű beszédjegyek között -- éppen a belső generálás következtében. Majdnem figyelmen kívül hagyja ugyanakkor az asszociációs folyamatokat; működésével néhány részletkérdés egyáltalán nem magyarázható meg (vö. Clark--Clark 1977, 205--207; Summary 1980). Az elmélet végül is egy óvatosan specializált változata a motoros teóriának, azzal a kivétellel, hogy az invariancia problémáját inkább neuroakusztikai szinten igyekszik megoldani, mint neuromotoros szinten. Meglehetősen absztrakt, ezért nehéz kísérletesen igazolni.

3. A globális beszédmegértés modellje. -- Ez a hipotézis a beszédmegértés során felvetődő számos problémát ugyan nem oldja meg, de a folyamatos beszéd felismerésére és értelmezésére minden más modellnél tökéletesebb. Eszerint a folyamatos beszéd megértése nem szavak egymásutániségának a felismerése és összerendezése a szemantikai jellemzők segítségével, hanem a globális egész alapján történik (Wingfield 1975, 146--160); mindebben fontos szerep jut az intonációs struktúráknak. A hosszabb szövegegységek felismerésével foglalkozó munkák azt mutatják, hogy gyakorlatilag nincs egyes szinteken lejátszódó működés, hanem az akusztikus inger szinte azonnal képes az asszociációs mezőket aktiválni, egyúttal pedig az alternatívákat (minőségileg és mennyiségileg is) korlátozni; jó összefoglalást ad erről Szirtes József tanulmánya (1983). A normál, a grammatikailag hibás és a grammatikailag helyes, de szemantikailag szabálytalan mondatok felismerése azt mutatja, hogy a grammatikai hibát kevésbé fogadjuk el, mint a szemantikailag rendhagyót -- azaz a grammatikailag hibás mondatok azonosítása a legnehezebb. Ez is utal(hat) arra, hogy a globális egészet érzékeljük, sokkal nagyobb ui. a választási lehetőség a szemantikai területen, mint a grammatikaiban. Utóbbiban jóval nagyobbak a megkötések, éppen ezért az ilyen

területen jelentkező eltérések könnyen bizonytalaná teszik a globális felismerést (vö. Miller és Isard grafikonban is szemléltetett eredményeit, 1963).

4. A Bondarko-féle elmélet (Bondarko et al. 1970). -- Tárgyalása azért is indokolt, mert jól látható, hogy a "szélsőséges" modellek részleges felhasználásával miképpen alakíthatók ki újabbak. A Bondarko-féle modell a hierarchikusan szervezett szintek egymásra épülésén alapszik, amely az akusztikai jeltől a valós üzenet felismeréséig terjed. Az egyes szintek a következők: hallási elemzés, fonetikai elemzés, morfológiai elemzés, szintaktikai elemzés, szemantikai elemzés. A fonetikai elemzés eredménye absztrakt (!); lehet akár akusztikai, akár artikulációs reprezentációja a beszédnek. Ez a reprezentáció pedig vagy a fonetikai szegmenseken, vagy a disztinktív jegyeken alapszik (vö. az artikulációs vagy akusztikus döntés megoldási kísérlete). Az invariancia problémáját úgy igyekeznek megoldani, hogy nem a felismerési kulcsokra alapoz, hanem feltételezi, hogy sokszoros döntés történik a hallási feldolgozás eredményeiből. Nem izolált hallási jegyekre összpontosít, tehát nem akusztikai kulcsokra, hanem a többszörös kulcsokat szimultán elemzi és hasonlítja össze. Bizonyos mértékig emlékeztet a "neurális spektrográfiának" nevezett felfogásra (Summary 1980; Szirtes 1983; stb.). (A neurális spektrográfia metaforikus kifejezés egy beszédhang/hangkapcsolat akusztikai szerkezetének, illetve az agyban kódolt megfelelőjének kapcsolatára utal.)

5. A hierarchikus megértési modell. -- A legtöbb modell -- több, kevesebb egyezéssel -- a fokozatos megértési hipotézis alapján építkezik. Lényege, hogy a bejövő akusztikus ingertől kezdődően egymásra épülő szintekkel jellemzi a megértési folyamatot, követve a beszéd nagyságrendi változásait egészen a jelentésig (Lindner 1969; Liberman 1970; Sawusch--Pisoni 1975; Clark--Clark 1977; stb.). Ennek a mechanizmusnak a modelljével és működésének részletes leírásával "A beszédmegértési folyamat alsóbb szintjei" című fejezetben foglalkozunk. A külön fejezetben történő tárgyalást az indokolja, hogy a beszédészlelési mechanizmusról kialakított elképzelésünk ehhez a modellhez közelít a legnagyobb mértékben.

Az áttekintéssel (amely természetesen korántsem teljes) érzékeltetni kívántuk, hogy a beszédmegértés folyamatának modellálásában általában csak a cél azonos, megoldási kísérlete közel annyiféle, ahányféle eltérő (módosított) felfogás létezik. Mindmostanáig nincs egy olyan modell, amelyik a különböző elképzeléseket egy empirikusan megalapozott leírásban egységesítené, és legalább a legfontosabb kérdéseket megnyugtatóan tisztázná. Utóbbiakhoz

tartozna elsősorban az elemi percepciós egység problémája, a szegmentálás, az invariancia (ezen belül a kontextus), a visszacsatolás, a késői kötés elvének kérdése, az artikuláció és a szupraszegmentumok elhelyezése és funkciója (a szupraszegmentumokra kiindulásul ld. Lieberman 1967). (E munkának nem célja valamennyinek a részletes ismertetése.)

A szegmentálás

A kérdés megfogalmazható úgy, hogy a szegmentálás felfogható-e egy speciális hallási folyamatnak, amely tartalmazza a fonetikai osztályozást; avagy a szegmentálás a fonetikai osztályozásnak magának az automatikus következménye. Példával megvilágítva, ez az első feltevés esetében azt jelentené, hogy a baba hangsor beszédhangjainak azonosításakor -- mint a megértés speciális hallási folyamatának eredménye -- azonnal képesek vagyunk az anyanyelvi fonémáknak megfelelő beszédhangrealizációk egymás után következésének felismerésére: b+a+b+a. A második feltevés esetében a baba akusztikai hullámformájának észlelését és feldolgozását követi egyfajta fonetikai osztályozás (bizonyos akusztikai kritériumok alapján), amelynek eredménye egyszersmind a szegmentálás, tehát fonémarealizációkra, szótagokra stb. bontás. A probléma azon alapszik, hogy nincs egy az egynek megfelelő az akusztikai jelben levő információ és a nyelvi analízisből kapott szegmentumok között (Pisoni--Sawusch 1975, 16). A kísérletek azt mutatják, hogy egy akusztikai jelrészlet tartalmazhat információt különböző fonetikai szegmensekről -- például egy bilaterális nazális és egy magánhangzó kapcsolatának átmeneti része a mássalhangzóról és a magánhangzóról is --, illetve igen különböző akusztikai jeleket gyakran azonos nyelvi információként észlelünk -- pl. a különböző magánhangzók előtt álló felpattanó zárhangot. Más megfogalmazásban: van néhány artikulációs konfiguráció, amely akusztikailag stabil abban az értelemben, hogy kis artikulációs változásnak kis akusztikai változás a következménye; nagy artikulációs változásnak pedig nagy akusztikai változás. A legtöbb artikulációs konfiguráció azonban instabil, mivel hasonlóan vagy ugyanolyan kis artikulációs változásoknak nagy akusztikai változás a következménye és megfordítva: nagy artikulációs változásnak kis akusztikai módosulás az eredménye (Stevens 1975). Vagyis az artikuláció és az akusztikum között nem lineáris a kapcsolat. Már Csisztovicsek leírják, hogy nehéz az akusztikai jelben meghatározni azt a részletet, amely az adott fonémát realizálja, mert néha az átmenetek, a szomszédos

hangok hordozzák az arra jellemző információt (1965). 15 évvel később úgy nyilatkozik, hogy a felismerésben néha másoknak vagy másoknak is szerepük van, például a formásátmenetek irányának, időszerkezeti állandóknak (1980, 67). Sok kutató úgy véli, hogy ha egyszer megértjük, hogy a fonémák hogyan alakulnak ki az akusztikai jelből, akkor a további nyelvi folyamatokat már könnyen meg tudjuk magyarázni (Summary 1980, 82). Közelebb jár a valósághoz az az óvatos megfogalmazás, amely szerint az előzetes hallási szegmentálás inkább szótag szintű, a szótagon belüli szegmentálás pedig valószínűleg szinonim az osztályozással magával (Studdert-Kennedy 1980, 49).

A visszacsatolás

Visszajelentés kifejezésben is használatos (vö. még feed-back). Tömör megfogalmazásban az a biológiai mechanizmus, amelyben egyik szerv kifejti hatását a másikra, ugyanakkor a másik szerv funkciója visszahat arra, amely funkcióját kiváltotta. A visszacsatolás fontossága a beszédmegértésben régóta ismert (vö. már Stricker 1880; pszichológiában korábban is); a mechanizmus gyakorlati felhasználása elsősorban a gyógyító pedagógiában jelentkezik.

A félteke-dominancia

Nem célunk, hogy az agykutatás eredményeivel foglalkozunk; néhány gondolatot azonban, amelyek szorosan hozzátartoznak a megértési folyamathoz, fontos elmondanunk. Agyunkat szívesen hasonlítják az elektronikus számítógépekhez; az összevetésben a gépek maradnak alul. Az agy tárolóképességét például 10^{14} bitre; fajlagos visszakeresési idejét 10^{-15} mp-re becsülik, vagyis gyorsabb, mint a másodperc egymilliomod részének az egymilliomod része (Fülöp 1984, 26).

A beszédmegértést a bal és a jobb agyfélteke koordinálja, mintegy a vezérlés helyei ebben a bonyolult rendszerben. Mind a mai napig vitatott az, hogy a két agyfélteke pontosan milyen szerepet vállal a beszédfeldolgozásban. Korábban úgy vélték, hogy a megértés csak a bal agyfélteke működési eredménye, a jobb agyfélteke egyéb észlelési sajátságok irányításáért felelős, (például a zenei hallás fejlettségéért). A kísérleti eredmények valószínűsítik, hogy a beszédmegértés szempontjából a bal agyfélteke a domináns, az uralkodó. Ha két beszédüzenetet juttatunk egy éphalló ember mindkét fülébe,

akkor a jobb fülbe érkezőket általában előbb és sokkal pontosabban fogja fel, mint a bal fülbe érkezőket (vö. dichotikus hallásvizsgálatok; a magyarra ld. Pléh 1984). A jobb fülnek közvetlen a kapcsolata a bal agyféltekével, a bal fülnek pedig a jobb agyféltekével. A bal agyfélteke uralkodóvá válása fejlődési eredmény; a fiziológiás fejlődés során alakul ki, bár a pontos életkort illetően nem alakult ki egységes vélemény. Az sem elképzelhetetlen, hogy a különböző nyelvek szerkezete kissé eltérően befolyásolja ezt a folyamatot. Magyar anyanyelvű óvodásokkal végzett kísérletek szerint 3 éves korban már kimutatható az ún. jobbfül-dominancia.

A bal agyfélteke tehát uralkodó a beszédmegértés vezérlésében, de a folyamat eredményes szervezésében jelentős mértékben részt vesz a jobb agyfélteke is (Cutting--Pisoni 1978). Ismeretesek olyan kísérleti eredmények, amelyek szerint a két agyfélteke eltérő feladatot lát el a beszédhangok felismerésében. A bal azonosítja az artikulációs helyet, míg a jobb reagál a zöngés és nem aspirált, illetve a zöngétlen, aspirált hangokra. A jelenség mind 2--5 hónapos csecsemőknél, mind (angol anyanyelvű) felnőtteknél kimutatható volt, újszülötteknél azonban nem (Eilers 1980, 30). Megint más kísérletek azt támasztják alá, hogy a jobb agyfélteke képes szavak és egyszerű mondatok felismerésére, de részletes nyelvi (például fonetikai) elemzésre alkalmatlan.

Újabb -- nem egy esetben a korábbiakkal ellentétes -- eredményeket hoztak azok a kutatások, amelyeket hasított agyú betegekkel végeztek (bizonyos okokból sebészeti úton kellett ezeknek az embereknek az agyában a két agyfélteke közötti kapcsolatot megszüntetni). A jobb szem látóterébe juttatott kép vagy a jobb kézbe adott tárgy megnevezése nem jelentett nehézséget a betegeknek (a bal agyfélteke működése); ugyanakkor képtelenek voltak a leg egyszerűbb válaszokra is, ha azok információját a jobb agyféltekére juttatták. Elektrosokk-kezelést követően vizsgálták a két agyfélteke funkcióját külön-külön. Ez azt jelentette, hogy egyik vagy másik félteke működését részlegesen kikapcsolták, felfüggesztették. Az eredmények azt mutatták, hogy a "bal féltekés" ember képes az elhangzó beszéd-részletek felismerésére, ismétlésére, de nem észleli azok prozódiai jellegét, például a kérdő vagy dühös hanglejtést. A "jobb féltekés" ember szóbeli kifejezőkészsége erősen csökkent, tárgyak neveit nehezen tudja megmondani. A beszédhangokat felismeri, mimikával, gesztusokkal pontosan jelzi az elhangzó beszéd prozódiai jellegét (Osmanné Sági--Erdélyi 1982).

A természetes kommunikációban a két agyfélteke bonyolult együttműködés-

ben és igen összetett funkciómegoszlásban látja el a beszédmegértés koordinálását.

A vizuális észlelés hatása a beszédpercepcióra

A beszédmegértés folyamatában más észlelési központok is részt vesznek (például az optikai, a tapintási központ). Szinkronizált film nézésekor a szereplők beszédét elsősorban az akusztikum, a hangzás alapján fogjuk fel. A látási élmény másodlagossá válik, ha az alapvetően nem áll ellentétben az akusztikus élménnyel. Ha azonban "rossz" a szinkron, s a szájmozgás, a látható beszédképzés nincs összhangban a hangzásélménnyel, akkor igen zavaró tényezővé válik, s meglehetősen nehéz a látási információt "kikapcsolni" ahhoz, hogy a beszédet könnyen értsük. (Előfordul olyan helyzet is, hogy a látási élmény mintegy uralkodik a hallásin. Moziban ülve a hangszóró mindig ugyanazon helyről (helyekről) közvetíti a közönségnek a beszédet. Hajlamosak vagyunk mégis különböző irányokba "tájéolni" a hangzást, pusztán annak a látási élménynek az alapján, hogy az adott szereplő a filmvászon melyik széléről beszél.)

A vizuális percepció segítségével a beszédfeldolgozásban a nagyothallók és a süketek használják a legintenzívebben; közismert megnevezése a szájról olvasás (vö. Uden 1976; Ling 1977; Csányi 1981). A beszédhangok egy része alkalmas arra, hogy az ajakartikuláció alapján többé-kevésbé pontosan felismerhető legyen. Ez nyújt lehetőséget a nagyothallónak, hogy a részlegesen kapott hallási információkat a látásiakkal kiegészítse. Nagyothallók 21 %-os beszédmegértése a szájról olvasás lehetőségével 64 %-osra emelkedett (Ling 1977, 50). Alig van kísérlet azonban, amely arról számolna be, hogy éphallók milyen mértékben használják a vizuális információt a megértésben (Sumbly--Pollack 1954).

Gyakorlati tapasztalat, hogy az idegen nyelvi szöveget könnyebben értjük meg, ha szemünkkel követhetjük a hangzásnak megfelelő leírt sorokat, vagy látjuk a beszélő száját. Angol nyelvi kísérletek kimutatták, hogy a vizuális információt jóval gyakrabban használja a beszélő/hallgató a beszédfeldolgozásban, mint feltételezték (Neisser 1984, 157). A hangzási és a látási élmény összefüggése "fordítottan" is jól ismert. A vizuális észlelés (pl. olvasás) révén kapott információk feldolgozását az "Üzenet" elhangzása nagymértékben segíti. Gondoljunk az olvasni tanuló gyermekre, aki

sokkal jobban megérti a leírtakat, ha hangosan olvassa (vö. Ladefoged 1967, 82).

Magyar anyaggal és magyar kísérleti személyekkel végzett kísérletek eredményei alapján a következőket állapíthatjuk meg. A hallási és a vizuális észlelés ép beszélők esetében is szoros összefüggést mutat; segítheti vagy nehezítheti a felismerést. A megegyező ajakartikuláció értelmetlen hangsorok szintjén átlagosan 10 %-kal, a bennük szereplő mássalhangzók felismerését több, mint 12 %-kal növeli. A zavaró ajakartikuláció hatására a logatomok azonosítása csaknem 20 %-kal csökken (a csak hallás alapján kapott helyes válaszokhoz képest). A különbségek értelmes (jelentéssel bíró) beszédanyag (szavak, mondatok) esetén még nagyobbak (részletesen vö. Gósy 1989).

Egészen más funkciója lesz mindezeknek a tényezőknek az érzékszerveiben károsodott ember beszédmegértése esetén. A vak-süket számára már csak a tapintás marad, amelynek segítségével valamiképpen mégis részt tud venni az emberi kommunikációban. Ez pedig igen kis esély az életben; mivel a tapasztalatok és kísérletek szerint ismereteink 83 %-át a látás, 11 %-át a hallás, 3,5 %-át a szaglás, 1,5 %-át a tapintás (!) és 1 %-át ízlelés útján szerezzük be (Fülöp 1984, 33). Különös jelentőségű tehát az angol Keller Helénnek, a múlt század végén élt vak és süket leánynak a vallomása: "A kút felé indultunk. A kutat övező bodzacserje illatától vezetve. Valaki vizet szivattyúzott a kútból, és tanítónóm a cső alá tette a kezemet. Míg a hűsítő víz egyik kezemre csurgott, a másikkra kibetűzte a "water" szót; először lassan, majd gyorsan. Csendesesen álltam s feszült figyelemmel követtem ujjai mozgását. Egyszerre csak átcikázott rajtam valami, feledésbe merült, ködös elmosódott emlék, a visszatérő gondolkodás villáma, és valamelyes módon nyilvánvalóvá lett előttem, hogy water azt a csodálatos, hűs valamit jelenti, mely kezemre folyt. Ez az eleven szó feltámasztotta lelkem életét; világosságot, reményt, örömet öntött belém, s megszabadított bilincseimtől! Még mindig voltak korlátaim, de korlátok, melyeket idővel már el tudtak távolítani.

Eltávoztam a kúttól, telve csupa tanulási vágygal. Minden dolognak megvolt a megjelölése, és minden megjelölés új gondolatot ébresztett" (Vértes O. 1954, 26).

Az elemi percepció egység

Talán az egyik legrégebb problémája a beszédmegértésnek; sokan alapkérdésnek tekintik. Mind a mai napig többféle hipotézis ismert ezzel kapcsolatban -- a kísérleti eredmények sem adtak egyértelmű választ. Az alábbiakban ismertetjük a legfontosabb nézeteket, egyben kritikájukat is adjuk. Majd saját eredményeink alapján próbálunk a kérdésben állást foglalni.

Sokáig a legelfogadottabb és a legszéleskörűbben ismert nézet a fonéma méretű beszédmegértési egység volt. Ez a hipotézis azon alapszik, hogy az akusztikai folyamat tükrözi a fonémák egymásutániségát. A percepció kísérletek eredményei azonban azt sugallják, hogy igen nehéz a fonéma elhatárolása az akusztikumban (szegmentálás, ld. újabban: Csisztovics 1980, 67--68). Számos kísérleti tapasztalat támasztja alá, hogy az elsődleges döntés nem feltétlenül a fonémáról történik. Gyakran a fonémáról hozott döntést megelőzően az ember már felismeri a hallott jel jellemzőit (Csisztovics 1965; Flanagan 1965; Ladefoged 1967; Fant 1968; stb.). A fonéma méretű egység percepciója ellen szól, hogy nagyon nehéz (hacsak nem lehetetlen) megtanítani valakit azoknak a fonémáknak a felismerésére, amelyek az ő számára nem jelentenek fonemikus valóságot. (Ötven rövid közlést 11 különböző anyanyelvű személlyel mondattak be anyanyelvükön, s ezeket 4 angol anyanyelvű fonetikusnak kellett fonetikus jelekkel átírnia. Ötven százalékban teljesítettek jól -- tehát a jelölés megfelelt az adott nyelv hangzási szabályainak --, 50 százalékban azonban "tévedtek", vagyis anyanyelvi percepció bázisukon keresztül "értették" az elhangzott hangsorokat. Ezek a "tévedések" a fonológiai ismeretek hiányáról tanúskodnak, vö. Studdert-Kennedy 1980, 59.) Kísérletei alapján Ladefoged úgy nyilatkozik, hogy a fonéma méretű percepció egységet alátámasztja az a tény, hogy a beszédet alfabetikusan írjuk le, de ez nem bizonyítja azt, hogy a percepció egység a fonéma. Szerinte a közvetlen percepció a fonémánál nagyobb egységekben történik, amelyek azonban kisebbek, mint a komplett közlés (1967, 143, 149). A kísérleti személyeknek nehézségeik vannak az individuális beszédhangok fizikai sorrendjének meghatározásában.

"Click"-kísérletek néven (click = csattanás, kattanás) foglalják össze azokat a vizsgálatokat, amelyekben folyamatos beszédet hallgattatnak le a

kísérleti személlyel, de a nyelvi anyagot úgy "preparálják", hogy a közlés-folyamatba valamiféle nem beszédjellegű zajt, csattanást, kattanást iktatnak (a folyamatos beszéd általában mondat, ritkábban szöveg nagyságrendű). A hallgatóknak meg kell határozniuk azt a helyet, ahol a zajt hallották. Az eredmények szerint a kísérleti személyek képtelenek voltak a zaj pontos helyének a meghatározására (Ladefoged i.m.), a legtöbb esetben -- ha erre külön nem figyelmeztették őket -- észre sem vették a zajt, illetőleg a beszédhang helyettesítésének tényét, automatikusan korrigáltak (Szirtes 1983). Az ebből adódó következtetések egyike, hogy tehát a beszédfelismerés nagyobb egységeiben kell, hogy történjen. 2)

A fonéma méretű egységet a megkülönböztető jegyek 2) ken alapuló beszédpercepció "váltotta fel". Ez úgy működne, hogy a bejövő jelet minden esetben egy bináris kérdés fogadná, például: zöngés/zöngétlen, vokális/nem vokális stb. (Hörmann 1971, 61). A kérdéssorozat vég-eredményeként alakulnak ki a megkülönböztető jegyek összességéből a fonéma. A nehézség itt is a szegmentálásban gyökerezik. Ahhoz ugyanis, hogy az egyes fonémákat jellemző megkülönböztető jegyeket azonosítsuk, meg kell határozni a hangsorban azt az akusztikai területet, jelrészletet, amelyről ezeket a döntéseket meg akarjuk hozni. Az akusztikai jelnek és a nyelvi szegmensnek a megfeleltetése pedig -- mint láttuk -- nem egyszerű. (Nagyon elnagyolt megközelítésben ez azt jelenti például, hogy a fogkefe szóban (ejtésben: [fok:efe]), az első [k] mássalhangzónak nem ugyanazt a fonémát kell megfeleltetnünk, mint a másodiknak; illetőleg a hosszú [k:] mássalhangzó két különböző fonémát reprezentál, nevezetesen a /g/-t és a /k/-t. Ez csak akkor képzelhető el, hogy ha a megkülönböztető jegyes beszédfelismerés egyfajta nyelvi szabályrendszerrel szimultán "dolgozik". Ez azonban további problémákat vet(ne) fel.

A megkülönböztető jegyes elmélet hatása ismerhető fel Dzsaparidze hipotézisében (1973, 130). Eszerint a percepciós egység az etalon, 3) amely olyan pontokat tartalmaz, amelyek fizikai jellemzői egy osztályt, egy kategóriát alkotnak (ez a zóna), az érzékelt jelenségek pedig ezeknek a pontoknak felelnek meg.

A kísérleti eredmények relatív ellentmondásai arra késztették a kutatókat, hogy a beszédfelismerés egységeként nagyobb nyelvi szegmentumot jelöljenek meg. Ilyen például a má s s a l h a n g z ó -- m a g á n h a n g z ó - k a p c s o l a t -- CVC és VCV szótagokkal végzett vizsgálatok alapján (Borovičková--Malác 1968, 83--88).

4.) Jóval szélesebb körben terjedt el a s z ó t a g mint elemi percepciós egység (Lehiste 1972; Fujimura 1975; stb.). A szótag mint elemi egység mellett újabban a reakcióidős mérésekkel szolgáltatott bizonyítékokat (angol anyanyelvűekkel végzett kísérletekkel). E kísérletekben azt találták, hogy akkor a legrövidebb a reakcióidő, ha az ingerszekvencia valódi szótag, továbbá, hogy rövidebb a reakcióidő, ha például a pajta szóban a paj-t kellett felismerni, mint amikor a pa-t (Szirtes 1983, 394; a szerző a könnyebb megértés kedvéért hoz magyar nyelvi példát). A szótag egység mellett szól az a tény is, hogy a kisgyermek (az anyanyelv-elsajátítás igen korai szakaszában is) képes a szótagolásra, s amint motoros képességei lehetővé teszik, legalább szótagszámban igyekszik az utánzott szóhoz hasonló t ejteni (Gósy 1984c). Igaz ugyan, hogy ez a képesség a ritmusérzékeléssel függ össze, és csak alapul szolgál a nyelvi funkció majdani elsajátításához. Említésre érdemes mégis, mint egy természetesen meglévő adottság, amelyet a nyelv (bizonyos értelemben) felhasznál.

A szótag egység mellett szól az a tény is -- Fromkin kísérletei szerint --, hogy másként történik egy szótag eleji [b]-nek a felismerése és másként a szótagvégi ([bad]~[dab]). Ladefoged a szótag mint percepciós egység ellen foglal állást, mondván, hogy kísérleteinek egyike sem igazolta ezt a hipotézist. Továbbá a Fromkin említette fonetikai pozíciótól függő percepció magyarázatára két lehetőséget lát: a) vagy létezik egy legkisebb változatlan neurofiziológiai egység, amelynek nagyobb (ható)területe van, talán nagyobb, mint egy szótagé, vagy b) kell, hogy legyen egy teljesen különálló, tárolt (emlékezetben őrzött) utasítás a szótag eleji és a szótagvégi mássalhangzókra vonatkozóan (Ladefoged 1967, 71). További ellenérvek a szótag méretű egység ellen: nagyon soknak tűnik az elraktározandó szótag-egység; túl nagy az az akusztikai információtömeg, amely egy szótagot jellemez (még akkor is, ha csupán egyetlen beszélő ejtéssajátságainak akusztikai következményeivel számolunk!); az artikulációs vizsgálatok azt mutatják, hogy nagy a szerepe a hangkörnyezetnek (az oroszban például csak a CV, CCV és CCCV típusok invariánsak: Csisztovics--Kozsevnyikov 1965, 199). (Ld. ugyanakkor a fonémaszint működésének leírását, amely -- magyar nyelvi kísérletek alapján -- megerősíteni látszik a szótag-egység működését.)

5.) Nagyságrendben a következő -- számba jöhető -- szerveződési egység a s z ó, amelynek már a jelentéstartalmát is figyelembe kell vennünk. Elemi egységként azért képzelhető el nehezen, mivel az emlékezet valószínűtlenül nagy megterhelését jelentené az elraktározandó mennyiség és információ

(Osgood--Sebeok 1965).

A szónál nagyobb méretű egységek közül elsőként az ún. ^{6.)} "fonemikus frázis" említendő meg (Dittmann-Llewellyn 1967, 341). (A terminust eredetileg a "beszélt nyelv elemzéséhez" használták 1951-ben a fonémánál és a morfémánál nagyobb egység megnevezésére.) A fonemikus frázis ismerve a határjegye; nagyságrendje pontosan nem meghatározott: egy, de gyakrabban több szóból álló beszédrészlet.

Nagyobb méretű szerveződési egység a ^{7.)} "mondat", ezen belül is a "magmondat". Eszerint mondatstruktúrákként azonosítjuk a beszédet (Bever-Lackner-Kirk 1969); s ezekből következtetjük ki -- "lefelé és felfelé" -- a kisebb és nagyobb nyelvi szegmenseket. Ahhoz azonban, hogy egy nyelv valamennyi lehetséges mondatát az emlékezetben a megfelelő engramként elraktározzuk, elképzelhetetlenül nagy emlékezet-kapacitásra lenne szükségünk. Amerikai kutatók kiszámították, hogy ahhoz, hogy az angol frázisok összes lehetséges grammatikai variációját egyszer halljuk, az embernek 3×10^{20} frázist kellene 1 másodperc alatt felfognia, s ennek 100 évig kellene megszakítás nélkül folytatódnia (Miller--Galanter--Pribram 1960). (A magmondat-hipotézist is számos bíráló érte, itt csupán egyre utalunk, a kisgyermek anyanyelv-elsajátító folyamatának törvényszerűségeire.)

Egyre nagyobb tért hódít -- éppen a kísérleti eredmények relatív bizonytalansága, a felfogások sokfélesége, az érvek és ellenérvek sokasága miatt -- az a hipotézis, hogy a folyamatos beszéd megértése -- bizonyos ^{8.)} hierarchia megtartásával -- a "globális egész" azonosításával és értelmezésével történik (Wingfield 1975, 146--160). Wingfield véleménye szerint nem létezik egy ún. egyedülálló percepciós egység, hanem a folyamatosságok hierarchiája létezik, ahol minden egyes szintnek saját "mértéke, nagyságrendje" van (pl. szótag, morféma). A működést úgy kell elképzelnünk, hogy megtörténik például a döntés a fonemikus szinten, majd módosul-folytatódik a döntéssorozat a szótagok, morfémák, szavak stb. szintjén, de az utóbb lévő szint(ek) mindig visszautal(hat)nak a korábbira abban az értelemben, hogy meg is változtathatják a későbbi információ következtében a korábbi döntést. Ez a felfogás nem zárja ki azt a lehetőséget sem, hogy a fonémáról történt végleges döntés a jelentés aktiválása után történjen meg. Wingfield mások méréseire hivatkozva állítja, hogy például a szótag-szinten végbemenő analízis "ideje" a memóriában tárolt hallási képnek megfelelően kb. 100--250 ms, de mindenképpen 1 s alatti érték. Massaro szerint

a beszédészlelés első feldolgozási szakaszában egy ún. echoikus memória csak 250--300 ms tartamú akusztikus ingert képes rögzíteni. (Mindkét értéktartomány mind beszédhang, mind szótagszinten reális lehet.) A "késleltetett kötés elve" megnevezés tulajdonképpen a visszacsatolásnak egy újraértelmezett, módosított formáját takarja. Nem biztos ugyanis, hogy a beszédmegértés folyamatában a hallgatónak minden egyes szinten azonnali igen/nem (bináris) vagy bármilyen más döntést kell hoznia. Elképzelhető, hogy a működés olyan, hogy késleltetheti a döntést, amíg újabb információkat nem kap (Pisoni 1981, 256). A felfogás gyökerei nyilvánvalóan ahhoz a tényhez nyúlnak vissza, hogy a hallgató nem egy esetben a megelőző fonémáról csak az azt követő elemzése után hoz döntést (Csisztovics--Kozsevnyikov i.m.). A különbség az, hogy Csisztovicsék ezt a szótag nagyságrendű elemi egység bizonyítékául és a szegmentálás alapproblémájaként értelmezték. Pisoni úgy vélekedik, hogy jó beszédminőség esetén igenis azonnaliak a döntések; ha azonban a megértést zavaró körülmények nehezítik, akkor működésbe lép a késleltetett kötés. Ez az elv csak úgy érvényesülhet, ha nincs egyetlen univerzális percepciós egység. Pisoni úgy nyilatkozik, hogy az egység nagysága változik a megkülönböztető jegytől a szó vagy mondatkapcsolatokig, ugyanúgy, ahogy a mechanizmus szintjei is változnak (1981, 253). (Flanagan már 1965-ben úgy fogalmaz, hogy a beszédfelismerés egysége a megkülönböztetési feladatnak megfelelően változik.) Ennek a -- ma legelterjedtebb -- felfogásnak számos előnye van: mindenekelőtt az, hogy az egység kötöttsége nélkül állíthatók fel a megértési folyamatot reprezentáló modellek.

Az invariancia jelensége

A mai napig vitatott jelenség. Az "állandó" meghatározásának nehézsége a beszéd sokrétűségével függ össze. Látszólag annyiféle invariancia jelenség található a beszédben, ahányféle területen a beszéd maga vizsgálható. Eszerint volna akusztikai állandó, artikulációs állandó, nyelvi jel állandó, észlelési állandó stb. Az ennek következtében fellépő székszist mutatja Port álláspontja, amely szerint az invariancia minden ember (értsd: kutató) számára mást és mást jelent (1984, 1). Ahhoz, hogy a jelen beszédkutatási eredményei alapján ítélhessük meg a problémát, áttekintjük röviden, hogy mit jelent az "állandó" a klasszikus fonetikában, fonológiában.

Az első kísérlet egyfajta tudományos értékű invariancia-keresésre a fonetikus átírás volt. Ez a rendszer, amely szándéka szerint tükrözi az elhangzó beszéd elemi egységeit, ezáltal lehetővé teszi a folyamatos beszédnek magának a leírását, azonnal kétféle "állandót" is tartalmaz: a beszélők "fölötti" és a nyelvek "fölötti" állandót. Ez azt jelenti, hogy függetlenül az egyedi ember beszédétől, valamint a különböző nyelvek hangos megjelenítésétől, létezik olyan invariáns jellemző, amelynek segítségével a különbségek figyelmen kívül hagyhatók. A fonetikus lejegyzés (átírás) számára mindegy, hogy például a magyar [o] magánhangzót nő, férfi, gyermek, fiatal vagy öreg mondta, az átírásban mindig ugyanúgy, tehát állandó jelenségként jelenik meg. Ugyanígy érdektelen, hogy orosz, magyar, német vagy angol nyelven hangzott-e el a hang, az átírás változatlan marad. Annak ellenére, hogy az említett ejtések eredménye, fizikai valóságukban -- akár képzésüket, akár akusztikai szerkezetüket nézzük -- erősen különböznek egymástól (vö. 1. táblázat).

1. táblázat

		[o]	
		Frekvenciaszerkezet (Hz)	
		F1	F2
magyar ejtés	magánhangzó		
	férfi	400	840
	női	470	900
	3 éves gyermek	755	1665
beszédszintézis		400	875
idegen nyelvi ejtés	orosz	440	825
	német	420	880
	finn	380	700
	fonetikai hangszabvány	450	580

A hagyományos fonetika, amikor a képzés jellege szerint írja le a beszédhangokat, ugyancsak állandó sajátosságokat keres. Ezek a jellemzők (pl. labiális, zárhang, affrikáta) teszik -- a feltételezés szerint -- lehetővé, hogy pusztán hallás alapján az elhangzó beszéd hangjai, például az átírási szabályok szerint leírhatók. Sok esetben azonban különféle mozgássor (artikuláció) azonos vagy hasonló hallásélményt okoz, még egyetlen nyelven belül

is. Port (amerikai) angol példája a [r] problémája, amelynek az ún. retroflex és a 'bunched' ejtését hallás alapján lehetetlen elkülöníteni. A beszéd megfigyelésében járatlan magyar beszélő sem képes például a hibásan ejtett réshangot a helyesen képzettől elkülöníteni. A megoldás az, hogy ezekben az esetekben természetesen a hallásélmény kell, hogy alapul szolgáljon a leíráshoz. Ezzel azonban a nyelvek "fölötti" állandó leírhatóságáról kell lemondanunk, hiszen az idegen ejtéssajátságokat rendszerint az anyanyelvi percepció bázison keresztül észleljük, ez pedig -- sok kísérlet tanúsága szerint -- hamisítja a valóságot.

A fonetikai szimbólumok az i d ő z í t é s tekintetében is "állandóak"; amennyiben az egyes beszédhangok mind nyelvi, mind fizikai (!) időtartamát relatíve határozzák meg: a nyelv rendszerbeli sajátságainak figyelembevételével. Az ajtó szót például mindig ugyanúgy fogjuk fonetikusan átírni, függetlenül attól, hogy az a szó milyen konkrét időtartamban hangzott el.

A fonetikai szimbólumok nem vesznek tudomást a beszédhangok s z e r - v e z ő d é s é b ől a d ó d ó v á l t o z á s o k r ó l sem, vagyis ismét egyfajta állandót tartalmaznak. Ezt a szeriális invariancia jelenséget jól példázza a magyar [k]. A veláris felpattanó zárhangunkat mindig ugyanez a jel testesíti meg, noha a követő beszédhangtól függően a rá jellemző artikulációs mozgás (a képzési hely) változik, vö. kutya, király, könyv. A [k] frekvenciagócának "vándorlása" a példaszavakban: 860 Hz, 2450 Hz és 1450 Hz, nagy különbségeket mutat. Állandó az is, ahogyan az átírási rendszer ugyanazon a módon jelöli a hasonló képzési jellemzőt, például a részjellet a különböző résmássalhangzóknál (vö. [f], [s], [l]) vagy a zöngéesség tényét függetlenül a mássalhangzó típusától.

A klasszikus fonológia, amikor a f o n é m á k a t j e l l e m z ő j e g y e k e t határozza meg, ugyancsak állandókat próbál találni, amelyeknek kettős a feladatuk: az egyes beszélők és az egyes nyelvek fölötti állandóság biztosítása. A jakobsoni elmélet eredetileg nagy rokonságot mutat a hagyományos invariancia-felfogással; ekkor az artikuláció és az akusztikum között még nem érvényesül mindig az állandóság. Például a 'flatness' jegy egy sor különféle artikulációs mozgás együttese, amelyek közös ismérve, hogy a magasabb frekvenciák energiamaximuma csökkenő tendenciát mutat. Ilyen artikulációs mozgás lehet például ajakmozgás (a szájnýílás szűkülése: [s] → [f]), az ajak előrecsücsörödésének változása: [e:] → [ø:]) vagy a larynx lejjebb kerülése. Akusztikai hatásuk azonos (vagy közel azonos), így nyelv-

vileg azonosnak tekinthetők. A Jakobson, Fant, Halle által alkotott elméletben az észlelés és az akusztikum sokkal nagyobb jelentőségű, mint korábban volt. Ez odavezetett, hogy az artikulációs/audiomotoros szimmetria helyett az artikulációs/akusztikus szimmetria került előtérbe. A disztinktív jegyek kibővülése az akusztikai jegyek bekerülésével további absztrahálódást eredményezett, s ezzel az invariancia fogalmának átértékelése szükség-szerűvé vált. Különösen a "nyelvek fölötti" állandó kérdése maradt erősen bizonytalan. Chomsky és Halle sok kérdést úgy próbált megoldani, hogy egyfelől átutalta őket a kognitív szintre, másfelől pedig univerzálisnak mondta ki, ami egyben azt jelentette, hogy értelmetlen a további elemzésük. Port szerint a Chomsky--Halle-féle elméletben a korábbi fonetikai teret egy ún. mentális tér váltja fel, amelynek legfőbb jellemzője, hogy nem a jeltulajdonságot hangsúlyozza, hanem a beszélő/hallgató észleleti sajátosságait (1984, 11).

Az újabb fonológiai elméletek (pl. autoszegmentális fonológia) már nem is próbálják az entitásokat általános ismérvekkel definiálni; ehelyett külféle módszereket alkalmaznak a felvetett kérdések megoldására.

A nyolcvanas évek elejére nyilvánvalóvá vált, hogy az addig alkalmazott eljárások, műveleti megoldások, illetőleg jegykereső stratégiák mind-mind csak részlegesen képesek az invariancia kérdését megoldani. Hallgatólagosan ezért három különböző síkra utalták át az általános jelenségének kutatását, illetve a meghatározás lehetőségét; az artikuláció, az akusztikum és az észlelés síkjára, azzal, hogy igyekeznek párhuzamokat találni az egyes síkok között. Középpontba került a funkció kérdése, ami azzal járt együtt, hogy a beszélő/hallgató "ember" felől közelítettek. A kiindulás az észlelés, az első kérdés pedig az, hogy ennek az artikulációval vagy az akusztikummal szorosabb-e a kapcsolata (vö. artikulációs vagy akusztikus invariáns jegy). Delattre egyenesen úgy fogalmaz, hogy az ún. disztinktív jegy tulajdonképpen nem más, mint egy perceptuális jel, amelyet vagy artikulációs, vagy akusztikus sajátosságain át lehet vizsgálni, de mindenképpen invariánsként fog funkcionálni (1967, 4). Más szavakkal ez azt jelenti, hogy az invariáns jegyeknek vagy a gesztusok (artikulációs mozgások), vagy eredményük (az akusztikum) szintjén kell leírhatóknak lenniük. Tekintve, hogy ez egyforma, a leírási módszer szabadon választható. Port javaslata a következő. A hallgató számára természetesen az akusztikai jelnek, azaz a hallásélménynek kell állandónak lennie. Ez viszont feltételezi, hogy a különböző fiziológiai adottságok ellenére, u g y a n a z o n m o z g á s o k

u g y a n a z o n é l m é n y t k e l t i k , i t t t e h á t k i m u t a t h a -
t ó a z i n v a r i á n s t a r t a l o m . G y e r m e k n y e l v i v i z s g á l a t a i m t a p a s z t a l a -
t a a z o n b a n e l l e n t m o n d e n n e k a f e l t é t e l e z é s n e k : s o k e s e t b e n
a h h o z , h o g y a g y e r m e k a " s z ű k s é g e s " h a n g z á s t e l é r j e , a " s z ű k s é g e s " h a l l á s -
é l m é n y t k e l t s e a h a l l g a t ó s á g b a n , m ó d o s í t o t t , n e m r i t k á n e g é s z e n m á s a r t i k u -
l á c i ó s m o z g á s s o r t k e l l e l v é g e z n i e , m i n t a f e l n ő t t n e k . E n n e k o k a e l s ő s o r b a n ,
(d e n e m k i z á r ó l a g) b e s z é d s z e r v e i n e k a f e l n ő t t é t ő l k ű l ö n b ő z ő m o r f o l ó g i a i
f e l é p í t é s e (a z " e l t é r é s e k " k ö z ű l : a l á g y s z á j p a d m é r e t e i , a r c - é s o r r ű r e g i
v á l t o z á s o k , a v o k á l i s t r a k t u s h o s s z a) . A h a s o n l a t k é n t f e l h o z o t t M o r s e - a b c
s e m m e g g y ő z ő . E s z e r i n t u g y a n i s m i n d e g y , h o g y v a l a k i í r o t t p o n t o t é s v e s z ő t
v a g y p e d i g a h a l l o t t r ö v i d é s h o s s z ú i d ő t a r t a m ú j e l e k e t é r t e l m e z i , a z e r e d -
m é n y a z o n o s . C s a k h o g y a z a r t i k u l á c i ó é s a z a k u s z t i k u m k ö z ö t t n i n c s e g y - a z -
e g y e n e k m e g f e l e l é s , e z p e d i g m e g n e h e z í t i a k ö z ö s i n v a r i á n s m e g t a l á l á s á t . N e m
d e r ű l k i t o v á b b á , h o g y m e l y e s e t e k b e n k e l l a r t i k u l á c i ó s é s m e l y e k b e n a k u s z t i -
k u s i n v a r i a n c i á t k e r e s n ű n k .

Mindezek feloldására Port szemléletes hipotézist vezet elő (i.m.).
Eszerint az invariancia mint jelenség az egyes s í k o k (úgy mint artiku-
láció -- akusztikum -- észlelés) k ö z ö t t i t r a n s z f o r m á c i ó -
b a n k e r e s e n d ő é s f o g h a t ó m e g e g y é r t e l m ű e n . T ö b b ú n . f o n e t i k a i t e r e t t é -
t e l e z f ő l , a m e l y e k k ö z ö t t i n v a r i á n s t r a n s z f o r m á c i ó k b i z t o s í t j á k a k a p c s o l a -
t o t . Ú g y g o n d o l j a , h o g y e z e n a m ó d o n j o b b a n l e í r h a t ó k a b e s z é d h a n g o k , m i n t
a s z e g m e n s e k k é n t . A z e g y e s t r a n s z f o r m á c i ó k f ő j e l l e m z ő j e a z i d ő b e l i t u l a j -
d o n s á g (t e m p o r a l p a t t e r n) , s z e m b e n a z " i d ő n é l k ű l i " f o n e t i k a i t e r e k k e l .
M i é r t m á s f e l f o g á s a e z a f o n e t i k a i t é r n e k , m i n t a k o r á b b i a k ? H a g y o m á n y o s a n
a f o n e t i k a i t e r e t (a l e g t ö b b e l m é l e t s z e r i n t) a f o n e t i k a i j e g y e k v a g y s z e g -
m e n s e k m e g h a t á r o z o t t s z á m ú e g y e d e a l k o t j a . Ú j a d a t o k b i r t o k á b a n (p l . e g y
ú j a b b b e s z é d h a n g m e g j e l e n é s e e g y ú j a b b a n t a n u l m á n y o z o t t n y e l v a l a p j á n) ú j
j e g y e k k e r ű l h e t n e k e b b e a t é r b e , s e z á l t a l m a g a t é r i s b ő v ű l . A p o r t i é r -
t e l e m b e n h a s z n á l t f o n e t i k a i t e r e t t r a n s z f o r m á c i ó k h a t á r o z z á k m e g , s a z á l -
l a n d ó ú g y j e l e n t k e z i k , h o g y a t r a n s z f o r m á c i ó k m ű k ö d é s e k ö z b e n a t é r m a g a
v á l t o z a t l a n m a r a d . M i n d e h e z m á r c s a k a t r a n s z f o r m á c i ó k m e g h a t á r o z á s á r a
v o l n a s z ű k s é g . E z b i z o n y o s e s e t e k b e n v i l á g o s a n k ö r v o n a l a z o t t , m á s e s e t e k b e n
m e g l e h e t ő s e n b i z o n y t a l a n n a k t ű n i k . L e g e l f o g a d h a t ó b b a b b a n a m e g f o g a l m a z á s b a n ,
h o g y e g y f e l ő l a z a r t i k u l á c i ó s é s / v a g y a k u s z t i k u s j e g y e k e g y é r t e l m ű e n m e g -
f e l e l t e t h e t ő k b i z o n y o s é s z l e l e t e k n e k , i t t t e h á t á l l a n d ó v a g y i n v a r i á n s a
m o z g á s é s a p e r c e p c i ó v a g y a z a k u s z t i k u m é s a p e r c e p c i ó k ö z ö t t i t r a n s z f o r -
m á c i ó , é s é r d e k t e l e n , h o g y m i k o r m e l y i k e s e t t e l á l l u n k s z e m b e n . M á s f e l ő l

létezik beszélők és nyelvek közötti tér-transzformáció, amely mindig állandó abban az értelemben, hogy kiszolgálja, vagyis biztosítja a kommunikációs feladatot, az észlelést (1984, 14--16).

Ennek az új transzformációs fonetikai térnek nagy haszna az, hogy továbbviszi az invariancia-kutatást; az eddigi elméletekkel ellentétben nem újabb és újabb érveket sorakoztat fel az egyik vagy másik lehetőséggel kapcsolatban, hanem tudomásul veszi az "artikuláció/akusztikum" aszimmetriáját, és ebből kiindulva igyekszik megoldást találni. Nagy eredménye, hogy -- feltételezhetően az utóbbi időkben nagy hangsúlyt kapó beszédészlelési, beszédmegértési vizsgálatok hatására -- a funkciót próbálja elsősorban meghatározni. Ez a funkció pedig nem más, mint a beszélő/hallgató nézőpontja, vagyis az elhangzó beszéd felismerésének kritériumai. Egyetlen gyenge pontját látom mind az eszmeifuttatásnak, mind a megoldási kísérletnek. A "fonetika" tudományának több évszázados "rossz" hagyománya ütközik ki itt is: a beszédet a leírható legkisebb egységek alapján értelmezi, tartja felismerhetőnek, így a fonetikai térben csak az ennek megfelelő entitások léteznek. Ennek alapján pedig legfeljebb a szóig jut el; fel is teszi a kérdést, hogy vajon mi a szó az artikuláció és az akusztikum területén. Választ azonban nem tud adni rá. Modellje így egy meghatározott körön belül marad és csak itt alkalmazható.

Véleményem szerint az invariancia jelensége nem szűkíthető le a beszédhangok vagy még kevésbé meghatározott jegyek területére. Emellett az az érv szól, hogy ha egyértelműen meg is határoznánk a beszéd invariánsjegy-rendszerét, ez nem azt jelentené, hogy a beszéddel kifejezett gondolat kibontásához (megértéséhez) szükséges állandót találtuk meg, csupán egyet közülük. Elgondolásom szerint a beszédben többféle, valószínűleg több szinten létező invariáns tulajdonság van. A beszélő/hallgató célja mindig az elhangzó beszéd pontos azonosítása, függetlenül például az extralingvisztikai körülményektől. Ismeretes, hogy az észlelési/megértési stratégiánk mindig az aktuális feladatnak megfelelően változik. Ezt számos kísérlet igazolta; az ilyen jellegű vizsgálatok egyértelműen mutattak rá, hogy az invariancia is változó jelenség kell, hogy legyen, amely az aktuális kommunikációs feladatnak van alárendelve. Tegyük föl továbbá, hogy megfelelő vizsgálatokkal meghatároztuk egy elhangzó mondat invariáns tulajdonságait

(most nem részletezzük ezek jellegét). Tegyük fel továbbá, hogy megértési kísérlettel igazoltuk az elemzés helyességét. Majd ugyanezt a mondatot torzítva (zajjal fedve, frekvenciaszűrőn átengedve stb.) újra lejátszuk a kísérleti személyeknek. Egy bizonyos torzításfokig a két kísérlet eredménye azonos marad. Erre pedig csak egyetlen magyarázat van: megváltozott az invariáns jegy, a beszélő/hallgató mindegyik esetben mást és mást használt fel a megértési folyamatban.

Az invariancia definícióját mindezek után a következőképpen fogalmazhatjuk meg: az elhangzó beszédben invariáns(ak) az(ok) a jegy(ek) vagy jelenség(ek) vagy esemény(ek), amely(ek) az adott beszédhelyzetben a beszélő/hallgató számára az észlelést és/vagy a megértést biztosítja(k). Ezek az akusztikum, a nyelvi ismeret, az extralingvisztikai tényezők és az egyéni tudáshalmaz területein határozhatók meg, s valamennyien a kognitív szinten használnak fel. Az adott beszédhelyzet alapján meghatározható, hogy az akusztikumnak, az extralingvisztikai tényezőknek vagy az egyéni ismerethalmaznak jutott-e nagyobb szerep. A kiindulás azonban mindig az elhangzó jel akusztikai tulajdonsága. Valószínű, hogy a különböző szintű invariancia-jelenségek között létezik egyfajta hierarchia, illetve dominancia, amely a konkrét esetekben vizsgálható. A továbbiakban a különböző szintű -- legfontosabbnak tartott -- invariancia-jelenségeket tárgyalom.

Invariancia az ismeretek szintjén

Az ismeretszintű invariancia -- ha hierarchiát feltételezünk -- a legfelső szinten működik. Lényege, hogy a beszélő/hallgató a beszéd felismerésének, a jelentések azonosításának stb. birtokában sem lesz képes az üzenet egészének a fogadására, ha az az állandó hiányzik, amely a megértést, adott esetben az értelmét biztosítja. A legjobb példa erre a viccek egy bizonyos csoportja, ahol a vicc, vagyis a csattanó megértéséhez megfelelő ismeretek szükségesek; ezek hiányában az "üzenet" a legfelső szinten dekódolatlan marad. Például: "Meddig volt Ádám és Éva a paradicsomban? Válasz: Almaérésig." A válasz két-, illetve egyértelműsége az első emberpárnak a paradicsomból történt kiűzetésében, annak bibliai leírásában gyökerezik. Az ismeretszintű invariancia természetesen a szokásos kommunikációnak is többnyire alapvető feltétele. Nemcsak tudás jellegű ismeretekről van szó, hanem a beszélő/hallgató számára "ismert" tényekről is. (Ha például két ember párbeszédébe belehallgatunk, rövidebb-hosszabb ideig képtelenek vagyunk értel-

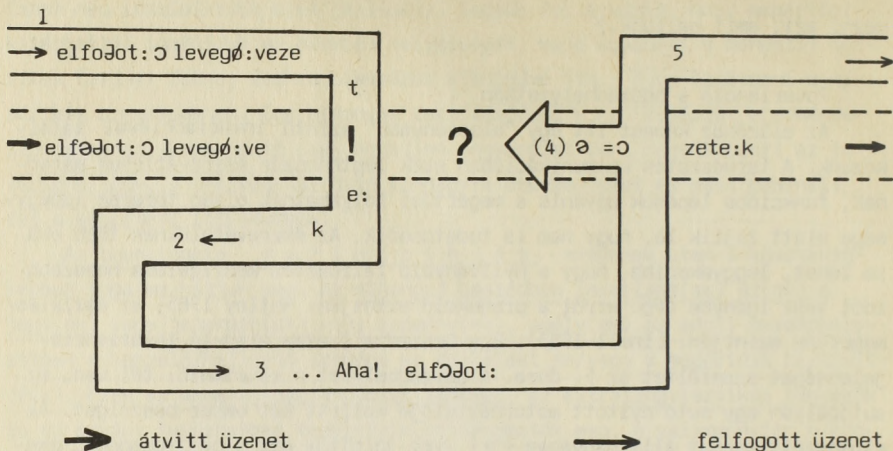
mezni azt, amit hallunk.)

Invariancia a beszédhelyzetben

Az előzőhöz képest itt egy "alacsonyabb" szintű invarianciával találkozunk. A természetes kommunikációban ezek legtöbbször észrevétlenek maradnak, funkcióba lépésük ugyanis a megértési folyamatnak olyan töredék időegysége alatt zajlik le, hogy nem is tudatosodik. Az észrevételüknek több oka is lehet, leggyakoribb, hogy a nyilvánvaló félreértés korigálása hosszabb időt vesz igénybe (vö. erről a produkció szintjén: Motley 1985; az észlelés/megértés szintjén: Linell 1983). Egy beszédhelyzetben alapuló invarianciajelenséget szemléltet az 5. ábra. A beszédhelyzet a következő: tél van, az autópályán egy autó nyitott motorháztetője mellett két ember beszélget. Az elhangzott mondat állítmányában [ə] -val jelöltük azt a beszédhangot, amelyik a "félreértést" előidézte. Ez nem azt jelenti, hogy a beszédhang artikulációján vagy akusztikus ismérvein múlt az észlelés, tehát hogy az invariáns jegyet ezekben kellene keresnünk. Az [ə] itt feltételezett oka a félreértésnek; amely azonban a beszédhelyzetben megjelenő invariancia következtében rendeződik. Nagyon gyakran nem a beszélő produkciójában, hanem a hallgató -- sokszor nyomon követhetetlen -- észlelési folyamatában adódnak a félreértést kiváltó okok (pl. más volt a hallgató elvárása a beszélőtől, vö. már Csisztovics--Kozsenyikov 1965; illetőleg a valószínűsített közlést észlelte/azonosította, nem a valóságosat). A példában beszédhelyzetben azonosítható mondatjelentés az, ami állandó: ti. a "vezeték nem fogyhat el"; mivel azonban tél van, a "fagyás" már szóba jöhet (nem is elemelve itt tovább azt, hogy kiváltóként a példában az eredetileg pongyolán megfogalmazott mondat is szerepet játszott: a beszélő arra gondolt, hogy a kicsapódó vízcseppek a hideg hatására ráfagytak a légvezetékre és eltömték azt).

Invariancia a kontextusban

Kontextuson itt a beszédhangok kapcsolódási sorozatait értjük, szavakat, szavak egymásutánját. Ez a kontextuális invariancia ismét szűkebb körben érvényesül, mint a megelőző, hiszen itt a beszédhelyzet nem feltétlen velejárója az invariáns tulajdonság működésének. Egy angol és egy magyar példával szemléltetem a jelenséget. Az angol mondat helyesírásban a következő: "The sheriff was shot through the heart". Kétféle ejtésvariációt közöl Abbott (1986, 302), hozzátéve, hogy az angol anyanyelvű beszélő/hallgató --



- 1 feltevés (az üzenet tartalmáról)
- 2 újravizsgálás
- 3 módosítás
- 4 mulasztás (tudatosítás)
- 5 újramegzés/javítás

5. ábra

Invariancia a beszédhelyzetben

az óriási ejtésbeli különbségek ellenére -- a mondatot mindkét esetben tökéletesen megérti. A két ejtésvariáció APhI-átírásban:

[diʃa'rifwazi'ʃota'trudihat]
[ðə'serifwəz'ʃotθruðəhɑ:t]

A magyar példa egy köznyelvi és egy nyelvjárási ejtés különbségeit mutatja: "jöttek egy páran, akik szoktak hozzájuk járni":

- a) [ɟt:ɛk ɛcpa:rɔn: ɟkik soktɔk y:hoz:jɔ ja:rɲi]
- b) [jɔt:ɛk ɛcpa:rɔn ɟkik soktɔ k hoz:a:juk ja:rɲi]

Feleslegesnek tűnik annak bizonygatása, hogy itt mindkét esetben valóban "beszédhangokon fölül" invariáns jelenség biztosítja a megértést.

Invariancia a beszédhangokban

Valamennyi közül a legjobban és a legtöbbet vizsgált lehetőség, sokak szerint az invariancia egyedül lehetséges megvalósulási területe. Pisoni definíciója a következő: "Egy x hangnak olyan akusztikai vagy tulajdonságjegy láncolata, amely minden kontextusban felidéri az adott hangot" (1981, 58). Ha csak és kizárólag az ennek a meghatározásnak megfelelő állandót fogadnánk el, akkor a fentebbi példában a jöttek és gyüttek szavakat nem lehetett volna egymásnak megfeleltetni. Ezzel nem azt mondjuk, hogy a definíció helytelen vagy használhatatlan, sőt; csupán visszautalva hangsúlyozzuk, hogy az invariancia jelensége szükségszerűen több szinten jelentkezik.

A Pisoni-féle definíció sugallja a feladatot: a fonetikusoknak nincs egyéb dolguk, mint megkeresni azokat a jegyeket -- bármik legyenek is azok --, amelyek a fenti követelményeknek eleget tesznek. Ehhez nagy segítség a fejlett artikulációs fonetika, a magas szintre jutott akusztikai fonetika, illetőleg a reneszánszát élő percepciós fonetika. Különböző kísérleti módszerek, legújabbán a beszéd szintézis (analysis by synthesis) nyújt sok lehetőséget e kérdés megválaszolására.

Megelőzően néhány elméleti jellegű, valójában azonban a gyakorlat számára alapvető kérdést kell átgondolnunk. Az invarianciának két ellenpólusa van: a redundancia és az irrelevancia jelensége. Mit jelentenek ezek a beszédészlelés szempontjából? A kétféle megnevezés itt kétféle jelenséget takar(hat). A redundancia "biztonsági többlet", az irreleváns pedig a "lényegtelen" (Karlgen 1968, 147), bár a fonetikában nagyon gyakran egymás szinonimájaként használják őket. A lényegtelen elhagyható, mert soha, semmilyen körülmények között nem képes az üzenet dekódolhatóságának biztosítására. A redundáns csak relatív feleslegességet jelent, amennyiben bizonyos körülmények között információhordozó jel(együttes)t képvisel. (Erre most csak egyetlen példa: csendes, zajtalan körülmények között jóval kevesebb akusztikai jel elegendő egy mondat megértéséhez, mint zajos vagy zavart körülmények között; utóbbi esetben több/többféle akusztikai paraméter hordozza azt a szükséges információt, amely a megértéshez nélkülözhetetlen.) Mindez újra a minimális egység problémájához vezet.

A kutatók nagyjából egyetértenek abban, hogy az akusztikai analízist a hallgató mintegy integrálja az összes már meglévő vagy kapott szükséges információval (itt jut nagy szerephez az ún. háttér-tudás, 'background knowledge'), s ez biztosítja a közlés tökéletes feldolgozhatóságát, értelmezését.

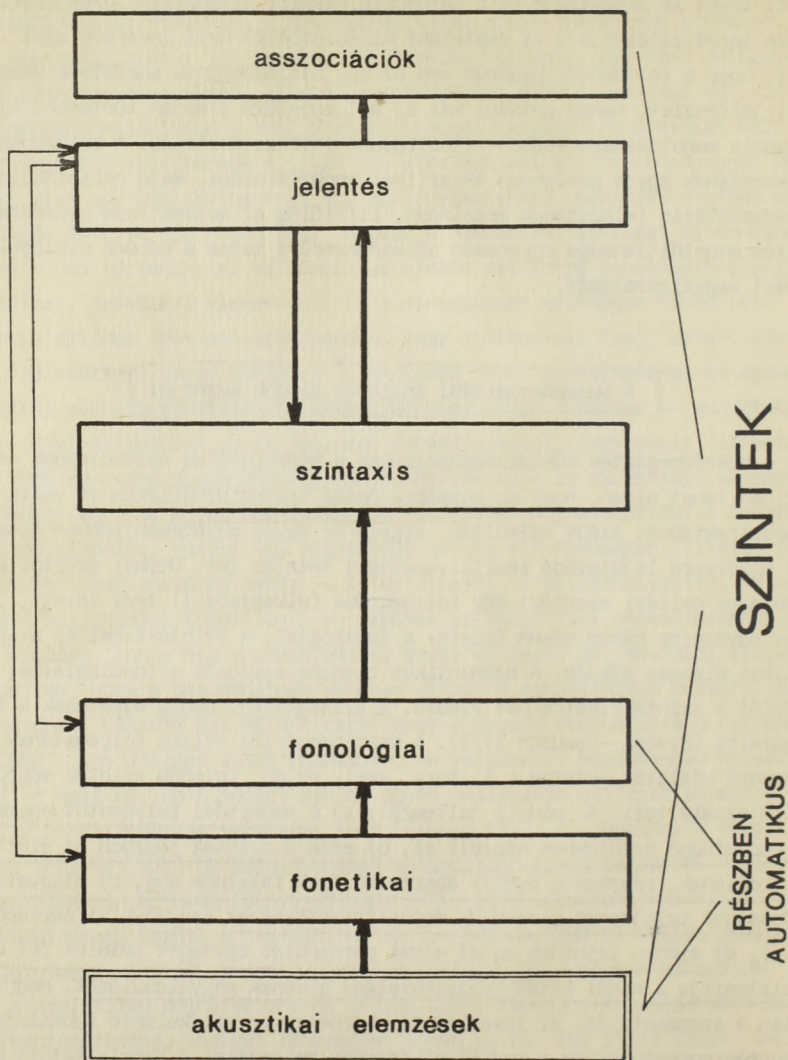
Egyfelől tehát különböző jellegű információkat (jelformációkat) tartalmazó "tömegeből" választhatja ki a beszélő/hallgató a számára fontos elemeket. Ezek egy része azonos az ugyanazon nyelvet beszélőknél, egy (más jellegű) része azonban választható, illetőleg egyénenként eltérő (például egy szövegfelismerési kísérletben, ahol mezőgazdasággal kapcsolatos szavakat kell megérteni, egy agrármérnök nyilván más stratégiával és ismeretanyaggal dolgozik, mint egy géplakatos vagy egy történelemtanár). Normális esetben mindenképpen létezik egy közös rész, amikor a dekódolási folyamat mindenkinél azonos módon történik; ez különösen az alsóbb szintekre vonatkozik (hallási, akusztikai, fonetikai elemzések). Ez a minimálisan szükséges közös rész változik állandó (pl. halláskárosodás) vagy pillanatnyi (zaj, zavaró körülmény, félreértés) zavar hatására. Ez a közös rész körülírhatóan és egzakt adatokkal meghatározhatóan látszik. Tekintve, hogy a beszéd -- akusztikailag -- egy hullámforma, a rá jellemző paraméterekkel (frekvencia, intenzitás, idő) szükségyszerűen behatárolható. Melyek azok az akusztikai paraméterek tehát, amelyek a beszéd felismerést alapvetően meghatározzák? A beszéd megértés globális, vagyis egy összefüggő szöveg elhangzásakor, illetve felismerésekor meglehetősen nehéz -- a tartalomtól, a beszédhelyzettől és a beszélő/hallgatótól függetlenül -- ezeknek az akusztikai paramétereknek a behatárolása. Ennek oka a redundancia; amely egyfelől ok, másfelől alapvető biztosíték, hogy a beszédüzenet valóban eljusson rendeltetési helyére, a hallgatóhoz. Lássunk két példát arra, hogy a beszéd redundanciája milyen nagyfokú! Egy átlagos angol mondatban, ha valamennyi magánhangzó helyett [ə] -t hallunk, de a mondat minden egyéb jellemzője változatlan marad, akkor a mondat megértés tökéletes lesz (Gimson 1962, 4). Ugyancsak angol példa, hogy egy 1000 Hz sáv szélességű mondat 1500 Hz-es középfrekvenciával 90 százalékosan érthető (Denes--Pinson 1973, 185). (Összehasonlításképpen: normál körülmények között a beszédet átlagosan 100 és 8000 Hz között halljuk.) Szemléletesen mutatják a beszéd információsűrűségét a következő adatok, amelyek kísérletileg meghatározhatók (Fülöp 1983, 12--15). A beszéd maximális információsűrűsége, amelyet lényegében sosem ér el, 100 bit/s lehet. Összehasonlításképpen: egy beszélgetés átlagosan 0,2 bit/s, de egy konvencionális beszélgetés átlagosan 0,05 bit/s-ra is csökkenhet (Lajtha 1964, 17). Mindehhez képest az agy információfeldolgozó képessége a beszéd megértésben 10^8 -- 10^9 bit/s (Gordos--Takács 1983, 251).

A beszédészlelés kutatásának fellendülése eredményezte, hogy egyre jobban előtérbe került az ún. akusztikai invariancia. Ez a beszéd szegmentális

szintjén, a beszédhangok területén keresi az invariáns jegye(ke)t. Megerősödött tehát az akusztikum és a percepció közötti szimmetria. Szem előtt tartva természetesen azt az észlelési kísérletekben nyilvánvalóvá vált tényt, hogy a felismerés gyakran nem az egy beszédhangnak megfelelő akusztikai jelrészlet, hanem például már az azt megelőző alapján történik. Az invariancia megoldáskeresésében elsődlegessé vált az észlelés. A beszédhangok szerkezetének egyre pontosabb akusztikai meghatározása, majd napjainkban a beszédszintézis feladatának megoldása, illetőleg az automatikus beszédfelismerés megoldatlansága egyenesen szükségszerűvé tette e kérdés mielőbbi korrekt megválaszolását.

A beszédmegértési folyamat alsóbb szintjei

A beszédmegértés működésmechanizmusa a hierarchikus építkezésnek megfelelően lehet olyan, hogy az egymásra épülő transzformációkon át válik egyre absztraktabbá, avagy szimultán, vagyis az egyes szinteken időben (csaknem) egyszerre lejátszódó részfolyamatokat tételez fel. Utóbbi szerint az elsődleges hallási elemzést egy felismerési (elképzelési) terv követi, amely egyszerre három síkon folyik: a fonológiai, a szintaktikai és a szemantikai elemzés síkján. A szemantikai elemzés egyfelől a fonológiaiakra, másfelől a szintaktikaiakra hat vissza, a szintaktikai pedig ugyancsak a fonológiaiakra (Pisoni--Sawusch 1975). A beszédmegértés teljes folyamatának sémaszerű vázlatát mutatja a 6. ábra, amely az ún. felsőbb szintek működését is szemlélteti. A modell jellemzői: a) a megértési folyamatot egymással összefüggő szintekben képzelel el, b) ezek a szintek részben az élettani hierarchiának, részben a nyelvi absztrakciónak felelnek meg, c) alapvető fontosságot tulajdonít a nyelvi meghatározottságnak (anyanyelvi percepciós bázis), d) minden szintnek saját elemi percepciós egységet tételez fel és e) biztosítja a késői kötés (visszaütalás) elvének megvalósulását. Megoldatlan a szegmentálás, az invariancia kérdése, az artikuláció funkciója, a szupraszegmentumoknak a modell működésébe helyezése. A "késleltetett kötés elve" megnevezés a visszacsatolásnak egy újraértelmezett, kissé módosított formáját takarja (Pisoni 1981, 256). Nem biztos ugyanis, hogy a hallgatónak azonnali döntéseket kell hoznia. Valószínű, hogy a működés olyan, hogy késleltetheti a döntést, amíg újabb információkat nem kap. (A gyökerek nyilvánvalóan oda nyúlnak vissza, hogy a hallgató nem egy esetben a megelő-



6. ábra

A hierarchikus beszédmegértési modell felépítése és működése

ző fonémáról csak az azt követő elemzése után hoz döntést: Csisztovics--Kozsevnyikov 1965).

A hallásról

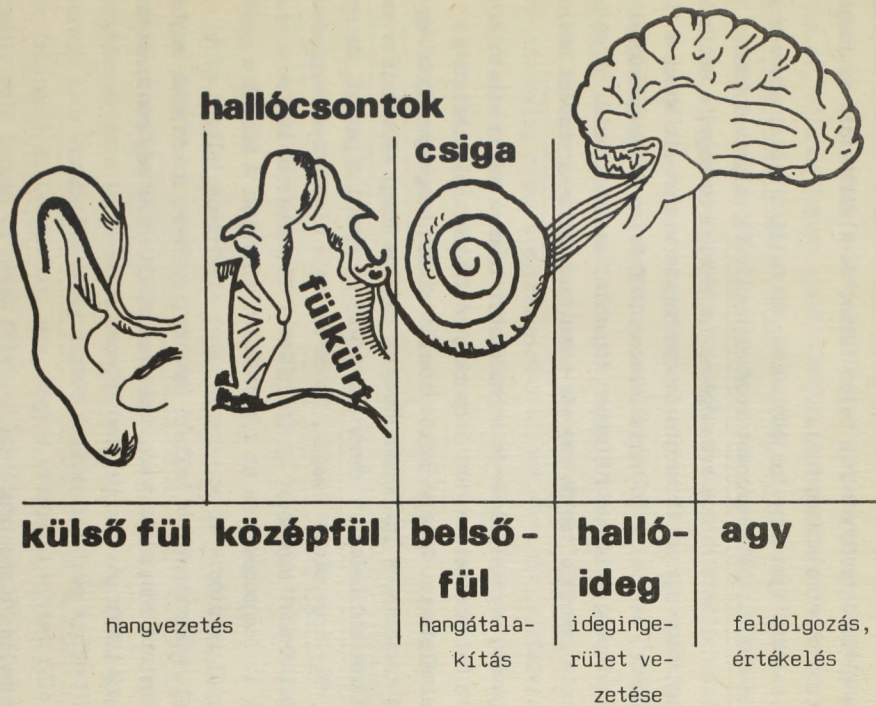
A beszédmegértés a halláson alapszik. Mondhatjuk-e akkor, hogy a hallás folyamata egyenlő a megértési folyamattal? Azt jelenti-e ez, hogy aki jól hall, annak a megértése is kifogástalan? A válaszadást megelőzően határozzuk meg a hallás fogalmát! Hallásnak azt a folyamatot nevezzük, amelyben hallószervünk környezetünk hangjelenségeit érzékeli, továbbítja, feldolgozza (Halm 1963, 9). A hallásnak két fő része van: a perifériás és a centrális hallás (nem kisebb jelentőségű természetesen a közvetítő idegpályák szerepe sem). A perifériás hallás a hallószerv és a megfelelő idegpályák működése, a centrális pedig az agy megfelelő központjaiban történő feldolgozás. Sajátos, minden más akusztikai jelenség feldolgozásától eltérő működést igényel a beszéd felismerése és megértése. A hallás épsége még nem jelenti a beszéd feldolgozás épségét is; akár a perifériás, akár a centrális hallásműködést tekintjük. Kimondható és hangsúlyozandó is egyben, hogy a hallás nem egyenlő a beszédmegértéssel, a hallási rendszer tökéletes működése nem biztosítja az elhangzó beszéd pontos azonosítását, megértését és értelmezését! Hangsúlyozni azért fontos, mert ez általában csak bizonyos betegségtípusok (például afáziák) esetében válik nyilvánvalóvá.

A hallás kérdésével már az ókori népek is foglalkoztak. A legrégebbi orvosi dokumentumok szerint, a sumérok ékírásos feljegyzéseiből megtudjuk, hogy a fül az akarat szerve. Az egyiptomi Ebers-papiruszon olvasható (i.e. 1500-ból), hogy a fül nemcsak a hallás szerve, hanem része a légzőrendszernek. Az egyébként magas színvonalon álló ókori kínai orvostudomány szerint a fül a vesék ablaka (feltehetően a fülzsír kellemetlen íze miatt). A görögök nézete szerint érzékszerveinkben a külvilág tükröződik, s azt tartották, hogy hasonlót a hasonlóval érzékelünk. Ez odavezetett, hogy elképzelésük szerint a rezgő levegőből álló hangot a dobüregben lévő levegő segítségével fogjuk fel. Pitagórasz -- akit elsősorban a róla elnevezett tétel révén ismerünk -- volt az első, aki az érzetet mérhető fizikai mennyiséggel hozta kapcsolatba; felfedezte ugyanis a hangmagasság és a különböző hosz-

szúságú rezgő húrok közötti összefüggést (i.e. VI. század). A későbbi századok újabb és újabb kutatásai révén egyre többet tudtak meg a fül anatómiájáról és a hallási folyamatról. Azt, hogy a fülnek melyik a hangot felfogó része -- a csiga (cochlea) -- valószínűleg egy Thomas Willis nevű oxfordi egyetemi tanár fedezte fel a XVII. században. A fizikai akusztika fejlődése nagymértékben elősegítette a hallásélettan sok kérdésének megválaszolhatóságát. Jelentős felfedezések születtek a XIX. és a XX. században. Békésy György kutatásaival a modern halláselméletek megteremtőjévé vált; eredményeiért 1961-ben Nobel-díjat kapott. Ugrásszerű fejlődést jelentett az elektroakusztika felhasználhatósága; a régi mechanikai-akusztikai fogalmakat elektroakusztikai fogalmak váltották fel (vö. Halm 1963).

A fül a hallási rendszer hangérzékelő része, amelyben a hangjelenség idegjelenséggé alakul. A fület külső, közép és belső fülre osztják (7. ábra). A külső hallójárat a hangtér rezgéseit a dobhártya felé közvetíti. Akusztikailag mint kis rezonátor szerepel. A hang innen a középfül felé halad, amelynek részei: a dobüreg, a cssecsnyúlvány és a fülkürt. A dobüregben három csontocska van, amelyeket alakjukról kalapácsnak, üllőnek és kengyelnek neveznek. Két nyílás vezet a belső fülbe: az ovális és a kerek ablak. A kerek ablakot a dobhártya zárja el, amely csúcsával befelé, kissé ferdén helyezkedik el a külső hallójárat végén. Békésy mérései szerint nem egyszerűen egy kifeszített hártya; rugalmasságát inkább a mögötte lévő levegőpárnának köszönheti, amely a dobhártyát és a kerek ablakot is védi a hirtelen légnyomásváltozásoktól és a túl erős mély hangoktól. Meg kell említenünk a dobüregi izmokat, amelyek a hallócsontokon tapadnak és a hangintenzitással arányosan megfeszülnek. Erősebb hang hatására a középfül izmai összehúzódnak. Hallásról akkor beszélünk, ha a hangrezgés a hallójáraton át a dobhártya--hallócsont rendszeren keresztül vezetődik a belső fülbe. A középfül feladata, hogy a levegő hangrezgéseit a belső fül folyadékára vezesse. Végeredményként azt mondhatjuk, hogy a középfül úgy működik, mint egy illesztő transzformátor, összehangolja a külső levegő és a csiga akusztikai sajátosságait. A belső fül legfontosabb része az ún. Corti-szerv. Ez a hallóideg végkészüléke, amelyben a fizikai hangjelenség idegjelenséggé alakul át. Mindez a külső és a belső szőrsejtek működésén alapszik, amelyek nagy védtetésben vannak: a szervezet legkeményebb csontjában helyezkednek el, és folyadékban úsznak. A testünket érő mechanikai hatások ezért nem zavarhatják őket.

Természetes beszédhelyzetben a hang levegőrezgések útján éri el a fü-



7. ábra

A hallásmechanizmus főbb részeinek sémája

let (ez a légvezetési hallás). A hallójáraton keresztüljutva mozgásba hozza a dobhártyát. Ez, valamint a hallócsontláncolat (kalapács, üllő, kengyel) átveszi és részben átalakítja a rezgéseket, amelyek a belső fülben lévő csigára jutnak. Itt a mechanikai inger, a rezgés, idegi ingerületet vált ki. Ezt a hallóideg, illetve az agy hallópályái a nagyagykéreg megfelelő központjáiig vezetik. Itt történik a felfogás. (A koponyacsontokat ért hangrezgések közvetlenül is átvehetőnek a belső fülbe, és kiválthatják az idegi ingerületet. Ez a csontvezetési hallás.)

Az ember hallása igen nagy terjedelmű. Az ép hallású fiatal ember a 20 és 20 000 Hertz (Hz) közötti hangokat érzékel(het)i. Ez valamivel több, mint 10 oktáv. A hallható hangok erősségének terjedelme is nagy. A decibel azt mutatja meg, hogy egy adott hang erőssége hogyan viszonylik ahhoz a leghalkabb hanghoz, amelyet a jó hallású ember 1000 Hz-es hangmagasságnál meghall. A hangerősség érzése a fülünkben logaritmikusan változik. 85 dB az a határ, amely felett a tartós zaj már maradandó halláskárosodást okoz (vö. Tarnóczy 1984).

A hangmagasság érzékelésének és megkülönböztetésének módja vitatott. E szempontból a halláselméletek két csoportra oszthatók: helyelméletekre és frekvenciaelméletekre. A helyelmélet szerint minden hangmagassághoz egy megfelelő hely tartozik, s így minden hangnak sajátos idegi képviselője van a csigában. A frekvenciaelmélet szerint az agykéreg elemzi a hangot, de nem egy vagy több hely ingerülete alapján, hanem az egész alaphártya rezgésével keletkezett hullámképet elemezve. Az ingerületet továbbvezető sejtek a neuronok, amelyek az idegdúcokat és az idegpályákat alkotják. A kódjel a neuronokon keresztül vezetődik az agykéregbe, ahol elektromos jelként mint kéri potenciál, kimutatható. A hallott hang felismerése, a kódjelek megfejtése magasabb szintű idegtevékenység, amelynek működése ma még pontosan nem ismeretes (Pauka 1980).

Az emberi hallás vizsgálata

Az ember ép vagy csökkent hallásának vizsgálata nagy hagyományra, sokféle módszerre tekinthet vissza. Minél többet tudunk a folyamatról, annál nagyobb a pontos vizsgálhatóság lehetősége is. A hallásmérés lehet szubjektív; itt igen fontos a beteg közreműködése és a vizsgáló személy jártassága; és objektív, ebben az esetben a vizsgálat céljának megfelelő műszerek jel-

zéseire épül a szakvélemény (pl. impedanciamérő, BERA). A szubjektív audiometriához a következő vizsgálatok tartoznak: a) súgott, illetve társalgó beszéddel történő hallásvizsgálat (hallástávolság megállapítása); b) hangvilla alkalmazása; c) tisztahang-audiometria; d) ún. küszöbfeletti vizsgálatok és e) beszédaudiometria. Ezeknek a vizsgálatoknak a sikeres elvégzéséhez nagymértékben szükség van a gyermek értékelhető részvételére; a megfelelő eredmények elérésének feltétele, hogy a gyermek a feladatot megértse, jól értse és hajlandó legyen együttműködni. A tapasztalat az, hogy -- különösen óvodáskorban, illetőleg enyhén retardált gyermekek esetében -- e vizsgálatok többsége csak megközelítőleg végezhető el. Legkevesebb nehézséget a hangvilla alkalmazása jelenti, valamint a súgott/társalgó beszéddel végzett hallásvizsgálat. A tisztahang-audiometria eredménye igen sokszor bizonytalan, nemegyszer magának a feladatnak az elmagyarázása is akadályba ütközik. Mindezek után ellentmondásnak tűnhetik, hogy a szubjektív audiológiai vizsgálatok közül mégis a szinuszaudiometriát használják leggyakrabban. Gyermekekben viszonylag ritka a küszöbfeletti vizsgálatok alkalmazása; a beszédaudiometriát elsősorban a hallókészülékkel történő ellátáskor használják.

A jó hallás alapvető feltétele a tökéletes beszédészlelésnek és az erre épülő beszédmegértésnek. Halláskárosító tényezők -- sajnos -- az élet minden területén jelen vannak (zajok, betegségek, balesetek stb.), ezért különösen fontos lenne a hallás épségének meghatározott időközönkénti ellenőrzése, gyermekkorban és (panasz esetén) felnőttkorban egyaránt.

A továbbiakban röviden ismertetjük az említett hallásvizsgálati eljárásokat, valamint a szűrési gyakorlatot.

a) Súgott és társalgó beszéddel történő hallásvizsgálat. -- A hallásvizsgálat legegyszerűbb formája, eredménye alapján következtetni lehet a beszédmegértésre is. A mérés csendes helyiségben végzendő, ahol a vizsgált gyermek és a vizsgáló személy között 6 méter távolság kialakítható. A gyermek féloldalt ül (vagy áll), úgy, hogy a vizsgálni kívánt füle a vizsgáló személy felé legyen (a másik füle -- lehetőség szerint -- egy segítő elfedi, egyúttal eltakar(hat)ja a gyermek szemét, nehogy a "szájról olvasás" segíthesse a megértést). Hat méter távolságból ún. tartalékkevegővel súgjuk a szavakat (egy és két szótagú értelmes szavak, illetve számok), a gyermek feladata azok hangos ismétlése. E vizsgálat előnye, hogy általános kép alkotható a gyermek hallásáról (beszédészleléséről); hátránya -- különösen szűrés-kor --, hogy a kívántnál lényegesen nagyobb mértékben alapul mind a vizsgáló

személy, mind a gyermek szubjektivitásán; a sűgás "hangerejének" mértékét csak hosszú gyakorlással lehet megközelítõleg azonosra kialakítani. Ha a gyermek az 1 méteren belüli sűgást nem érti, akkor ún. társalgó beszéddel vizsgálunk. A módszerrel nem fedhetõ fel minden hallászavar.

b) Hangvilla alkalmazása. -- A hangvillákat 1802-ben Chladni használta elõször hallásvizsgálatra. A vizsgálati eredmény felvilágosítással szolgál a hallászavar feltehető helyérõl, ezáltal diagnosztikailag fontos adatokat kaphatunk. Rendszerint két (egy magas és egy mély) hangvillát használunk. A vizsgálattal egyfelõl a két fül csontvezetéses hallásának egymáshoz való viszonyáról tájékozódhatunk (Weber-vizsgálat), másfelõl elkülöníthetõ egymástól a vezetéses és az idegi eredetű halláscsökkenés (Rinne-vizsgálat). A lecsengési idõ mérésével vagy normál füllel történõ összehasonlításával a hallásküszöb tájékozódó meghatározására is alkalmas.

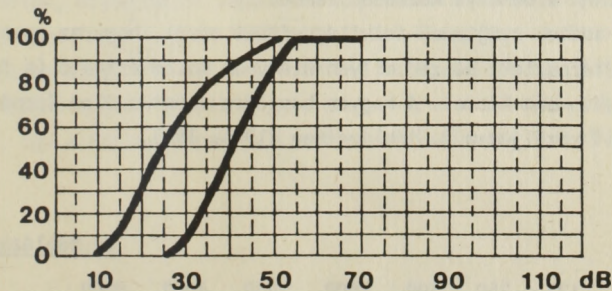
c) Tisztahang-audiometria. -- Eszköze az audiometer, amellyel az egyes frekvenciákon éppen észlelt szinuszhang aktuális intenzitásértékének jelölésével felrajzolható az ún. küszöbgörbe. Ép hallás esetén a görbe (camera silentában mérve) hozzávetõlegesen 0 dB-nél húzódik. (0 dB-nek azt a szintet tekintik, amely hangnyomásban $0,0002 \text{ din/cm}^2$, energiában $10\text{-}16 \text{ watt/cm}^2$ 1000 Hz-en, vö. Halm 1963, 40.) E vizsgálatnál különbözõ decibelértékeken eltérõ frekvenciájú szinuszhangot juttatnak fülhallgatón keresztül -- felváltva -- a gyermek jobb oldali és bal fülébe. A gyermek feladata, hogy (kézfeltartással, hanggal vagy más, játékos formában) jelezze, ha a szinuszhangot meghallotta. A különbözõ frekvenciájú (125, 250, 500, 1000, 2000, 4000, 6000, 8000 Hz-es) szinuszhangot egyrészt légvezetéssel (fülhallgatót át), másrészt csontvezetéssel (megfelelõ vibrátorral) juttatják a gyermekhez. Ez az ún. lég-csont-köz ellenõrzésére szolgál.

A tisztahang-audiometria a beszédmegértéshez szükséges finomabb mûködések (mint amilyen például a felpattanó zárhangok képzési helyének pontos azonosítása vagy a zöngés és zöngétlen mássalhangzók elkülönítése) esetleges hibájáról azonban igen kevésbé (vagy egyáltalán nem) szolgál felvilágosítással. A tisztahang-audiogram korlátozottan tájékoztat a halláskárosodásnak a beszédpercepcióra gyakorolt hatásáról is (Risberg--Agelfors 1984, 59).

d) Küszöbfeletti vizsgálatok. -- További speciális kérdések megválaszolásához ún. küszöbfeletti vizsgálatok szükségesek. Ezek elsõdleges célja a hallásromlás helyének minél pontosabb behatárolása: azoknak a betegségeknek az elkülönítése a belsõ fülben, amelyek a Corti-szervhez, illetõleg a

hallóidegekhez köthetők (Fleischer 1976, 12--13).

e) Beszédaudiometria. -- A beszédmegértés vizsgálatára szolgál. Különböző intenzitásértékeken szavakat kell a vizsgált személynek felismernie és visszamondania. A beszédaudiometria atyja a magyar Hőgyes Endre. Már a múlt században is kevesellték ugyanis a beszédmegértés mérése nélküli hallásvizsgálatokat. A gyermekek beszédaudiometriás vizsgálata 1924-re nyúlik vissza, és Fletcher nevéhez fűződik (vö. minderre: Götze 1974, 19). A tapasztalat az, hogy a második életévben a szó alkalmassá válik hallásvizsgálatra (Götze 1974, 85). A gyakorlat azt mutatja, hogy a gyermek sokkal jobban reagál a szavakra, mint a szinuszhangokra (Hahlbrock 1970, 97). A kisgyermekkorú beszédaudiometriához kétféle teszt is rendelkezésre áll (Götze 1974; Farkas--Gósy--Hirschberg 1983). A kísérleti eredmények szerint a 3--6 évesek megértési küszöbe (egyszótagú, ismert jelentésű szavak alkalmazásával) 40 dB-en van, itt 50 %-os megértés. A 100 %-os megértés 60 dB-nél következik be (Farkas--Gósy--Hirschberg 1983). A 8. ábra a felnőttek és a gyermekek felismerési arányait mutatja az intenzitás függvényében. Mindkét beszédaudiometriás vizsgálat anyagát egyszótagú értelmes magyar szavak képezték, de gyakoriságuk eltérő volt, tekintettel a gyermekek és a felnőttek eltérő szókincsére. A 10-10 szóból álló szósorok mindkét anyagban fonetikailag kiegyenlítették és egymással cserélhetők: nehézségi fokuk nagyjából azonos volt (részlet vevő beszédhangok, jelentés stb.).



8. ábra

Felnőttek (bal oldalon) és gyermekek (jobb oldalon)
megértési görbéje

A tisztahang-halláshoz relatíve nagyon kevés működő rost kell. Az olyan komplex hangjelenség felfogásához, mint a beszéd: nagyon sok (Götze i. m. 158). Láttuk, hogy azonosan ép hallás (ép küszöbgörbe) mellett mekkora az eltérés a beszédmegértési teljesítményben; bizonyos centrális zavarok rejtve is maradhatnak, mert ép a küszöbgörbe. Idős korban előbb jelentkezik a normáltól való eltérés a beszéd-audiogramon, s csak aztán mutat eltérést a tisztahang-audiogram.

f) Szűrés. -- A gyermek hallásának és beszédmegértésének szűrése gyakorlatilag nem megoldott. (A hagyományos eljárás az, hogy a nagycsoportos óvodás gyermekek hallásszűrését audiológus asszisztens, helyenként audiológus orvos, ritkábban védőnő végzi el. Ez azonban nem terjed ki az összes hatévesre, s a szűrés meg- vagy meg nem történte lényeges különbségeket mutat az ország földrajzi megoszlásának tekintetében. Nem hivatalos adatok szerint az iskolába lépő gyermekeknek csak mintegy 30--40 %-a esik át ilyen szűrésen.) A beszédmegértés fejlettségét egyáltalában nem vizsgálják.

A hallásszűrés az esetek többségében mindössze három, esetleg négy (!) frekvenciát érint: az 500, az 1000, a 2000 és a 4000 Hz-et (az óvodai zajtól függően általában 20--30 dB-en). Nincsen hivatalosan megállapított kötelező szűrési frekvencia, például Kollár csak hármát kíván meg szűrésnél: az 500, az 1000 és a 4000 Hz-et (1977, 29). A szűrés eredménye ennek megfelelően meglehetősen durva, kisebb zavarok rejtve is maradhatnak; egyáltalán nem foglalkozik a beszédmegértéssel, és a kevésbé intelligens vagy erősen szorongó gyermekek eredménye kérdéses lehet.

Kísérleti adatok nyújtanak felvilágosítást arról, hogy az egyes frekvenciasávokon áteresztett beszéd milyen arányban marad érthető (a frekvenciasávot oktávcsávokban mérve). A magyar beszédre vonatkozólag Tarnóczy mérési eredményeit közöljük a 2. táblázatban (1984, 194).

2. táblázat

Sávközép (Hz)	125	250	500	1000	2000	4000	8000
A magyar beszéd érthetősége (%)	2	13	18	22	22	20	3

Az adatok azt mutatják, hogy az 500, 1000 és 4000 Hz-en összegezve, 60 %-

osra tehető a megértés, amely az alkalmazott nyelvi anyagtól függően változó lehet. A 125--250, 2000 illetőleg 6000--8000 Hz-en esetleg meglevő károsodás a beszédmegértésben komoly nehézségeket okozhat; jól tettenérhető az írás--olvasás elsajátításában és az idegennyelv-tanulásban, illetve a televízió-, rádióadások megértésében, a telefon közvetítette kommunikációban. Nem lenne mellőzhető tehát a "teljes" hallástartomány ellenőrzése már a szűréskor.

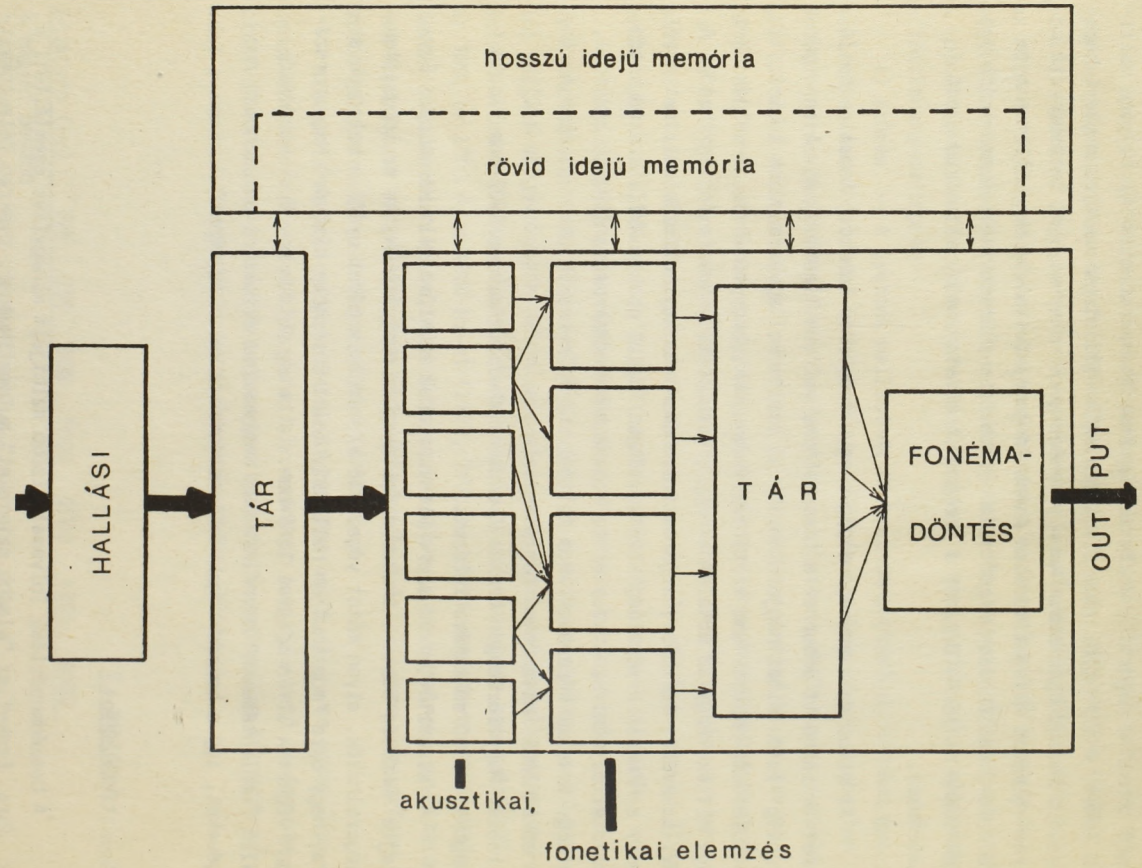
Az orvos és a szurdopedagógus együttműködése lehetővé teszi a hallásban károsodott embernek a társadalomba való beilleszkedését. A reneszánsz koráig "Isten által megjelöltnek", az örültekkel egyenrangúnak tekintették a süketnéma embert. Nem tartották érdemesnek, hogy tanítsák, neveljék őket, úgy hitték, hogy az értelemnek úgyis csak csekély fokára fejleszthetők. A XV. században Agricola, a jeles fríz humanista tesz először említést arról, hogy a némaság a süketség következménye. Mégis, csak a XVII. század végén kezd elterjedni az a szemlélet, hogy a süketnémák megtaníthatók beszélni, s hogy értelmi képességük nem feltétlenül alacsonyabb, mint az ép embereké. A süketnémák intézményes foglalkoztatása a XVIII. század végére nyúlik vissza. Magyarországon Cházár András nevéhez fűződik az 1802-ben Vácott megnyílt első süketnéma intézet.

Az elvesztett érzékszerv visszanyerése lehetővé vált korunkban. Egyfajta süketség esetén: amikor a Corti-szerv nem működik, de az idegsejtek ingerelhetők, olyan műtét végezhető el, amelyik módot nyújt a betegnek arra, hogy újra "megtanuljon hallani". Kelet-Európában elsőként Magyarországon végeztek ilyen műtétet 1985-ben. (A beteg fülébe egy 8 elektródából álló "implantátumot" operálnak be, amely képes érzékelni a különböző rezgéseket, így a beszédhangok felismerésére alkalmassá tehető.)

Akusztikai szint

A beszédmegértési folyamat alsóbb szintjeit kinagyítva szemlélteti a 9. ábra. Ezeket az "alsóbb szinteket" automatikusnak, gyakran félautomatikusnak minősítik; jobbnak találom a részben automatikus meghatározást. Az automatikus így inkább arra vonatkozna, hogy a hallgató általában nincs tudatában azoknak a működéseknek, amelyek a morféma-, illetve szófelismerést megelőzik. Kísérleti helyzetben az automatizmus mértéke

9. ábra
A beszédmegértési folyamat alsóbb szintjei



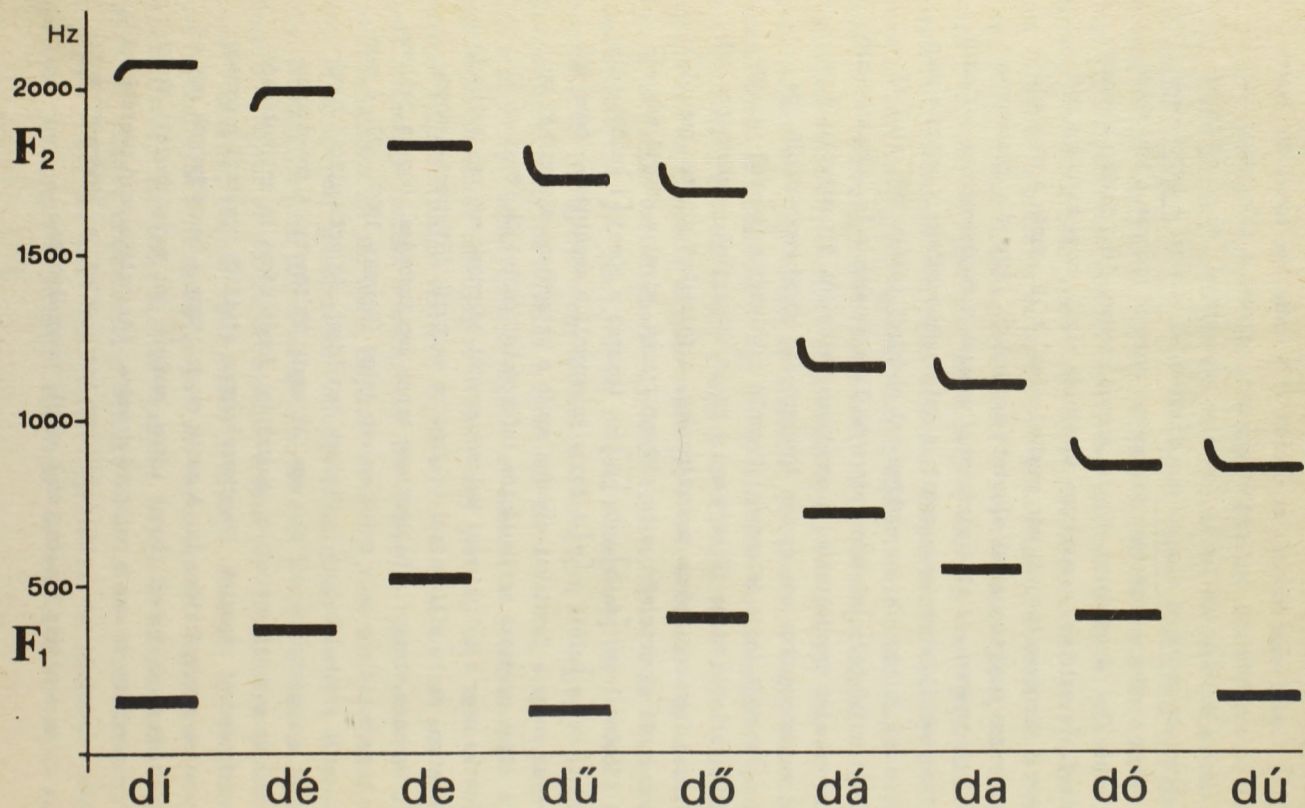
csökkenthető (a jelentés kiiktatásával), és nyomon követhető, hogy az alsóbb szinteken lejátszódo működéssorozat csak részben automatikus.

Az input vagy bemenet maga a beszéd, az az akusztikai jelsorozat, amely a nyelvi információt tartalmazza. A feldolgozás első lépcsője az ún. elsődleges hallási elemzés, ahol néhány előzetes döntés történik a frekvencia, az intenzitás és az idő tekintetében. Ezek a döntések nagyjából behatárolják a felfogott (észlelt) akusztikai jel hullámtulajdonságait; észleljük például, hogy zene vagy beszéd, ének vagy hangszer, gyors vagy lassú, magas vagy mély, halk vagy erős stb. (sőt egyszersmind egyéb szubjektív ítéleteket is alkotunk: kellemes, kellemetlen). Ezek az eredmények az észlelési információtárban összegződnek, ahonnan a beszédfeldolgozás indul. Az itt felgyűlt adatok alapján történik az akusztikai elemzés, amelytől kezdve lép be az emlékezet is a feldolgozásba. Tegyük fel, hogy az elhangzott beszédészlet egyetlen szó volt: sír. Az akusztikai elemzéssel választ kapunk arra, hogy a hangsor tartalmazott-e zöngés hangot vagy nem; vokális elem volt-e jelen vagy nem; észlelhető-e energiakonzentráció valamilyen frekvencián vagy nem; ha igen, akkor hol; a frekvenciaváltozások miként jelentkeznek az időben; hol volt intenzitáscsúcs (volt-e); milyen volt annak a lefutása; stb. (Feltételezésünk szerint bináris döntések sorozatával valósul meg az akusztikai elemzés. A szakirodalom tanúsága szerint azonban még e tekintetben sincs egyetértés a kutatók között. Massaro például úgy találja, hogy a jellemző jegyek nem binárisak és a feldolgozás inkább folyamatos jellegű, vö. Summary 1980, 87.) A sír szó példájánál maradva: a jelsorozat akusztikai elemzése a következő eredményeket hozza (itt csak a jellemző értékeket közöljük; a percepciósfeldolgozás ennél jóval részletesebb):

- a) a jel teljes időtartama 640 ms;
- b) a jel első 30 %-a nem tartalmazott periodikus rezgéseket az 500 Hz alatti tartományban, a további 70 %-ban igen;
- c) az első 30 %-ban 1500--2500 Hz táján, 56 %-ban 300 és 2200 Hz-nél, az utolsó 14 %-ban pedig 400, 1800 és 2600 Hz körül észlelhető energiakonzentráció;
- d) a középső, 56 %-os időintervallumban nincs zörej, csak periodikus rezgés;
- e) az intenzitás a felfutási és a lecsengési szakasz között két ponton csökken lényegesebben, mintegy 10--20 dB-nyit.

Az akusztikai elemzés adataira épül a fonetikai osztályozás. Ekkor döntési szabályok alapján nyelvészeti-fonetikai szempontból ítéljük meg a bejövő jelet. A sír szót azért választottuk példának, mert nehézség nélkül megoldható a szegmentálás problémája. Nehéz meghatározni azt, hogy egy folyamatosan változó akusztikai jelsorozatban, amilyen a beszéd, melyik és mekkora rész jelenti egy adott fonéma realizációját. Nincs ui. egy az egynek megfelelés az akusztikai jelben levő információ és a nyelvi analízisből kapott szegmentumok között (Pisoni--Sawusch 1975, 16; Csisztovics 1980, 186). A kísérletek azt mutatják, hogy egy adott jelrészlet tartalmazhat információt különböző fonetikai szegmensekről (pl. egy bilabiális nazális vagy egy [j] és egy magánhangzó kapcsolata); illetve igen különböző akusztikai jeleket gyakran azonos nyelvi információként észlelünk (pl. a különböző magánhangzók előtt álló felpattanó zárhangot, vö. Gósy 1981). Más megfogalmazásban: van néhány artikulációs konfiguráció, amely akusztikailag stabil abban az értelemben, hogy kis artikulációs változásnak kis akusztikai változás a következménye, a nagyak pedig hasonlóan nagy. A legtöbb artikulációs konfiguráció azonban instabil, mivel hasonlóan vagy ugyanolyan kis artikulációs változásoknak nagy akusztikai változás a következménye és megfordítva (vö. 27. old.). Mindez felveti a szegmentálási (elhatárolási) problémákat: felfogható-e egy speciális -- beszédre kialakított -- hallási folyamatnak, amely tartalmazza a fonetikai osztályozást; avagy a szegmentálás a fonetikai osztályozásnak magának az akusztikus következménye. A kérdés megválaszolását nehezíti az a tény, hogy gyakran nem a konkrét értékek alapján kell döntéseket hoznunk, hanem viszonyok alapján (pl. egy hang észlelt hangossága nemcsak a specifikus intenzitásától függ, hanem pl. az intenzitásváltozástól is, amely megelőzi és követi, vö. Darwin: Summary 1980, 84). Az akusztikai jelsorozatban mindig vannak olyan "pontok", amelyek könnyen azonosíthatók paramétereik alapján (pl. a magánhangzók ún. tiszta fázisai). Amikor az akusztikai elemzés eredményei összegeződnek, vannak már olyan részletek, amelyek egyértelműen, s olyanok, amelyek nagy valószínűséggel bejósolhatók mint beszédhangok (vö. Markov-lánc). Kimondható, hogy a szegmentálás a fonetikai szinten kezdődik, majd a fonológiai szinten "fejeződik be" állandó visszacsatolás mellett (ill. szimultán működésekkel valósul meg). A sír szó azért nem vet fel szegmentálási nehézségeket, mert az egy adott időszakaszra jellemző frekvencia- és intenzitásstruktúra egy a-

dott beszédhangra jellemző: az első 30 % egy [ʃ] típusú réshang, a további 56 % egy [i:] típusú magánhangzó, az utolsó 14 % pedig egy tremuláns jelenlétére utal. A beszédhangok egy részét azonban, például a felpattanó zárhangokat, ezen a szinten nem tudjuk ugyanígy azonosítani. Ha az adott beszédjel hullámtulajdonságai alapján nem dönthető el azonnal a beszédhang minősége, akkor mintha a megértési folyamatot vezérlő program a következő utasítást adná: "ha bizonytalan vagy az azonosításban, lépj tovább a következő szintre". És valóban: a rendszer továbblép azzal, hogy bizonyos előfeltevéseket a kérdéses hangról már magával visz. A 10. ábrán a magyar [d] hangnak az összes magánhangzóval alkotott kapcsolata látható a frekvencia és az idő függvényében. Az alsó vízszintes vonalak a magánhangzóra jellemző alacsony frekvenciájú energiamaximumot, a felső vízszintesek a második nagy energiamaximumot jelképezik, az ezekhez bal oldalról csatlakozó, alulról vagy felülről hajló (az [e] esetében nincs hajlás) vonalak a [d] mássalhangzó jellegzetes zörejét szemléltetik. A sematikus ábrázolás jól mutatja, hogy valamennyi magánhangzóhoz más és más frekvenciájú zörej kapcsolódik, de mindegyik a [d] hangot fogja jelenteni! Hasonló a helyzet a [b]-vel is; így az akusztikai szinten hiába történt meg a rövid időtartamú -- zárhangra jellemző -- zörej frekvenciájának meghatározása, a fonetikai szinten nem állapítható meg erről egyértelműen a beszédhang típusa, mivel a követő mássalhangzótól függően lehet [d, b] vagy [g] is. Ugyanaz a zörej [i] előtt például a [b]-t, de [e] előtt a [d]-t fogja jelenteni! A megoldás -- némelyek szerint -- az, hogy a fonetikai elemzés során a beszédhang akusztikai hullámformája által nyújtott artikulációs információt észleljük (Best et al. 1981), illetve hogy a felismerési kulcsok együtt jelennek meg egy artikulációs gesztusban, amely pillanatról pillanatra változik (Studdert-Kennedy 1980, 55). Massaro szerint a hallgatónak arról van tudomása, hogy a jellemző jegy a beszédcsatorna mely szintjén van jelen (Summary 1980, 87). A szegmentálást tehát artikulációsan igyekeznek megoldani. Az artikuláció hangsúlyozásának a megértésben két oka van. Az egyik történeti. A fonetikusok (nyelvészek) az eszközfonetikában, akusztikai fonetikában is artikulációs megkülönböztetéseket végeztek. Következésképpen átvitték ugyanezt a perceptív folyamatok magyarázatára is. A másik ok az, hogy a beszédhangok egy olyan hangforrás segítségével jönnek létre, amelyek jól meghatározott akusztikai következményei vannak a hallgató számára. (Az artikuláció szerepét a megértésben -- mint láttuk -- különösképpen két elmélet, a libermani motoros teória és az analízis szintézissel modell hangsúlyozza.



10. ábra

Az F2-átmenet mozgása a [d] mássalhangzó magánhangzós kapcsolataiban (az F1-átmenet mozgását nem jelöltük)

A fonetikai elemzés eredménye ismét egy tárba kerül, nevezhetjük fonetikai tárnak. A beszédmegértés kutatása az utóbbi években sok módszert és jellemzőt kölcsönzött az információfeldolgozási modellektől. Ennek eredményeként épültek be a modellekbe a tárolási struktúrák, amelyek alapvető részei a megértési folyamatnak. Így nyílik lehetőség arra, hogy a feldolgozáshoz mindig az éppen szükséges adatokat hívjuk elő, de az összes elemzési eredmény megőrződik a folyamatműködés időtartamára (a tárolás biztosítja az esetleges korrigáláshoz szükséges pluszinformációkat). A fonetikai szinten hozott döntési eredmények az adott nyelvre jellemző beszédhangok megnevezései. Erről a működésről tájékoztat az az EEG-vel végzett kísérlet, amely szerint csak a beszédhangok felismerésekor jelentkezett nagyobb amplitúdójú kiváltott válasz a bal agyfélteke fölött (Szirtes--Diekmann--Rothenberger--Jürgens 1981, 171). Egy másik kísérletben speciálisan előkészített magánhangzók, zaj és szinuszhangok kiváltotta agyi potenciálokat hasonlítottak össze. Az eredmények szerint a magánhangzó-sorozattal előidézett görbék itt is nagyobb amplitúdójúak voltak, mint a másik két akusztikai ingerre kaptak (Spreng--Young 1978). A fonetikai szint működésére vonatkozólag végzett kísérleteket Pisoni (1973). Négy mássalhangzót választott ki a bilabiális felpattanó zárhangok közül, amelyek egymástól a VOT-ben, tehát a zöngékezési időben különböztek. A hangokat C_1 , C_2 , C_3 és C_4 jelekkel látta el. A VOT ideje a C_1 esetében: 0,00 s, a C_2 -nél: -0,02 s, a C_3 -nál: -0,04 s és a C_4 -nél: -0,06 s volt. A kísérleti személyek meg tudták különböztetni a négy mássalhangzót, vagyis a "hallási memóriában" elkülöníthetők voltak eltérő akusztikai sajátágaik alapján (mint a /b/ és a /p/ realizációi). Megnevezni azonban csak kétféle mássalhangzót tudtak: a [b]-t és a [p]-t. A C_1 és C_2 képviselte zárhangokat [b]-nek, a C_2 és C_3 -mal jelöltek pedig [p]-nek feleltették meg. (A tesztelés CV hangkapcsolatokban történt, ahol a V az [a] magánhangzó volt.)

Fonológiai szint

A részben automatikus feldolgozási szakasz utolsó lépcsője: a fonéma-döntés meghozatala. A beszédhangokat itt, a fonológiai szinten soroljuk be a megfelelő fonémaosztályokba. Itt dől el, hogy -- a hangképzés szempontjából: hátul képzett, ajakkerekítéses, rövid magánhangzó, amely a középsőnél kissé magasabb nyelvállással képzett; akusztikailag 360 Hz-es

első formánssal és 800 Hz-es F2-vel jellemezhető -- az /o/ vagy az /u/ fonéma kategóriájába tartozik. Avagy például a jéggpálya szóban, ahol az első szótag végén zöngétlen [k] mássalhangzót ismertünk fel a fonetikai szinten, valójában a zöngés /g/ megvalósulásáról van szó, azaz a szóösszetétel első szótagja nem a jék hangsor, hanem a jég szó. A fonémaszint mondatok azonosításakor is aktívan jelen van; bár ekkor a szemantikai relációk már elsődlegesek. Például: az Éva bátyja allergiás lett a [mę:stó:l] mondat kétséget hagy a hallgatóban afelől, hogy vajon a méz vagy a mész okozott-e allergiát. A példák azt sugallják, hogy a fonémadöntést lényegében az értelmezés határozza meg. Ez részben igaz, hiszen ismert olyan nézet, amely szerint a fonémáról történt döntés a szöfelismerés után jön létre, azaz a szó azonosítása erősíti meg a korábbi feltételezéseket. A fonológiai szint működése azonban elsősorban az anyanyelvi fonémarendszeren alapszik.

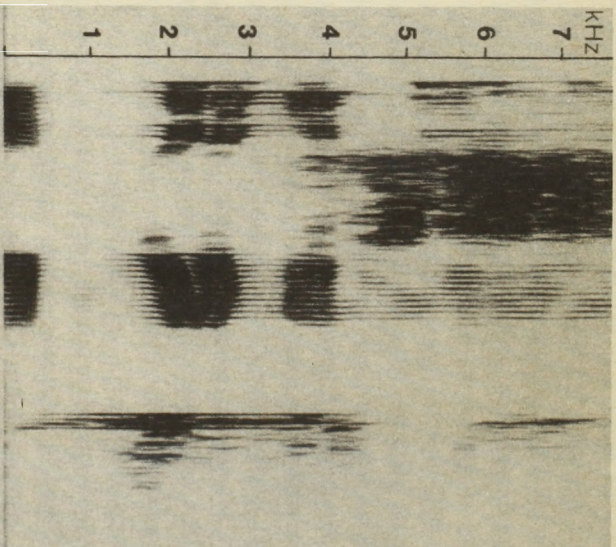
A fonémáról történő döntés, azaz a megfelelő fonémakategóriába sorolás a beszédhangok egy részénél megtörténhet izolált megjelenéskor is, vagyis ha a beszédhangot önmagában ejtve halljuk (pl. [o:, s, r]). A beszédhangok egy másik részének felismeréséhez azonban szükség van a hangkörnyezet bizonyos mértékű részvételére is, mivel -- az adott beszédhangra jellemző mértékben -- a hangkörnyezet hordozza a jellemző jegyeket (más kifejezéssel: kontextustól függő és kontextustól független észlelés, vö. Ladefoged 1976). A hangkörnyezet meghatározó szerepét mutatja, hogy például a rövid időtartamú szünetet grúz és angol anyanyelvűek beszédhangként azonosították (Dzsa-paridze 1971, 7; Clark--Clark 1977, 195.) A szünet és a [ʃ] időtartamának rövidítése jelentésváltozást eredményezett a következő mondatban: "Did anybody see the gray ship?" helyett "Did anybody see the great chip?"-et értettek (Studdert-Kennedy 1980, 53). A fonémadöntés meghozatalához a következő hangkörnyezeti (kontextuális) hatásokat kell minden esetben figyelembe venni: a) a követő (esetenként a megelőző) beszédhang minősége, b) a szótagban elfoglalt helyzet, c) a szótag szerkezete, d) bizonyos szupraszegmentális tényezők. Ismert olyan vélemény, amely az emberi beszéd megértésében -- már a legalsó szinteknél is -- azt tartja a legfontosabbnak, hogy a kontextus alapján működik (Kuwabara 1982, 175).

A fonémadöntést megelőző akusztikai és fonetikai eredményeket összefoglaló tár felfogható egy számítógép operatív memóriájaként, amelyből a felhasználó csak az éppen szükséges adatokat hívja elő a további feldolgozáshoz. Felvetődik a kérdés, hogy a megértési mechanizmus miért nem használja fel az összes elemzési eredményt döntéseihez. A beszéd nagymértékben redun-

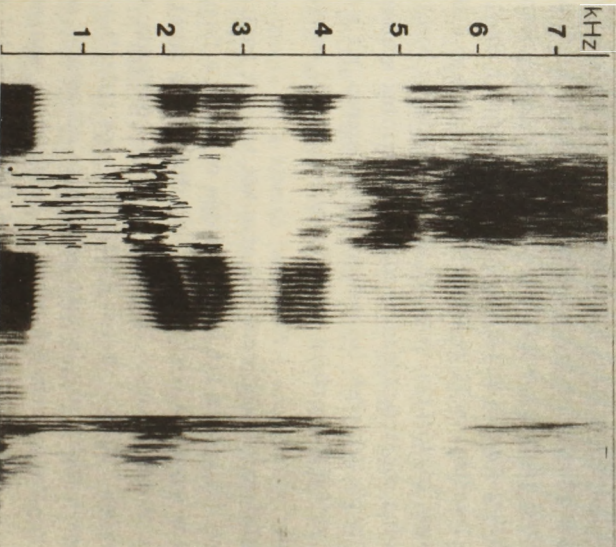
dáns, azaz a beszéd akusztikai hullámformája sok többletinformációt tartalmaz. A többletinformáció jól megragadható és kimutatható az akusztikumban; valójában az artikuláció, a beszédképzés következménye. Rendkívül nagy a szerepe a beszédmegértésben. Tudjuk például, hogy egy zöngétlen [s] mássalhangzó jellemző zörejegőcai az ún. felsőbb frekvenciákon helyezkednek el, mintegy 6000--8000 Hz-en (Fónagy--Szendé 1969; Olasz 1982, 66), sávszélességük az ejtéstől függően 1000--3000 Hz. A [s] azonban mintegy a 100--12000 Hz-es (esetenként még magasabb) tartományban is tartalmaz a mássalhangzóra jellemző zörejjösszetevőket, amelyeknek intenzitása azonban kisebb. A mássalhangzó azonosításában ezeknek fontos szerepük lehet. Az intenzív zörejegő eltűnésével, annak funkcióját a kevésbé intenzív, eredetileg többletinformációként jelentkező komponensek veszik át, és biztosítják erősebb torzulás esetén is a mássalhangzó pontos felismerését. Ez a magyarázata annak, hogy magasfrekvenciás hangcsökkenésben szenvedő angol anyanyelvű hallássérültek jól azonosították a [s, z] mássalhangzókat, noha hallásromlásuk éppen azokat a frekvenciákat érintette, ahol ezeknek a hangoknak az intenzív zörejegőca található (a tanulmány szerzői egyébként adások maradtak a magyarázattal, vö. Lawrence--Byers 1977). Speciális hangszínképelemzési eljárással, 700-as hangspektrográffal kiírtattuk a [s] többletinformációt tartalmazó zörejjösszetevőit az szik szóban (11. ábra). Az ábrán látható (esetenként még annál is több) akusztikai információ gyűlik össze például a [s]-ről az akusztikai-fonetikai térben. A fonémadöntéshez ebből a térből elegendő a legjellemzőbb, tehát az éppen szükséges adatokat előhívni. A tárolási rendszer azt biztosítja, hogy az alsóbb szinteken hozott döntések a mind időben, mind absztrakcióban távolabb eső szintek eredményeinek ismeretében korrigálhatók legyenek (mindebben fontos szerep jut a hosszú idejű memória éppen aktivált és rövid idejű (vagy közvetlen) emlékezetnek nevezett részének).

A fonológiai szint működését tekintve két kérdést kell megválaszolnunk: a) milyen jellegű az itt hozott döntés és b) mi a fonológiai szint elemi percepciós egysége. Általánosnak tekinthető az a nézet, hogy a fonémáról hozott döntés bináris. Ez összefügg a neuronműködéssel, amely ugyancsak 'igen/nem' típusú, hiszen a neuronok vagy pulzálnak (kisülnek), vagy nem. A hasonlatként már idézett számítógép-működésnek megfelelően azt mondhatjuk, hogy analóg beérkező jelet (a beszédet) a neuronműködésnek megfelelő digitális jelfeldolgozás fogadja. A fonológiai szint elemi percepciós egységének nagyságrendjét a szótagban jelölik meg (Lehiste 1972;

a [s] mássalhangzóról készített normál
hangszínkép



a [s] mássalhangzó többletinformációját is
tartalmazó speciális hangszínkép



11. ábra Az iszik szó hangszínképe (férfi ejtés alapján)

Fujimura 1975; Oden--Massaro 1978, vö. még Szirtes 1983, 394).

Véleményem szerint a felfogásban két probléma is meghúzódik. Az egyik a szótag bizonytalan definíciója, a másik pedig az egyes nyelvek eltérő szerkezete, eltérő hanghosszúsága, az egyszótagú értelmes szavak aránya stb. A beszédmegértés szempontjából inkább az utóbbinak van nagyobb jelentősége (ezt látszanak igazolni Fujimura, Macchi és Streeter eredményei is: 1978). Nyilvánvaló, hogy a szótag mint a fonémadöntés alapja mást jelent az angol anyanyelvűnek, ahol igen sok a jelentéses egyszótagú hangsor (vö. a szótári alapalakok és a használati előfordulás különbségét: Cutler 1987), és mást a finn anyanyelvűnek, mivel a finnben alig található egyszótagú szó.

Kísérleti eredményeim szerint a fonémadöntés alapja minimálisan egy magán- és egy mássalhangzóból, illetve egy mással- és egy magánhangzóból álló hangkapcsolat kell, hogy legyen, amely lehet egy magán- és több mássalhangzós szekvencia is. A nyelv maga -- ezen belül -- eleve hordoz magában bizonyos korlátozásokat, amelyek a megértési mechanizmus működésében érvényesülnek (vö. a percepció bázis kialakulása). Ilyen például a magyarban, hogy nem lehet döntési alap a magánhangzó +[h] kapcsolat abszolút szóvégen, mert ebben a helyzetben e mássalhangzót nem ejtjük (a méh hangsor jelentései közül kivételt jelentenek az 'anyaméh' és a MÉH mozaikszó, ezekben a szóvégi [h]-t ugyanis kiejtjük). Libermanék már idézett kísérlete (1957a,b) óta, amelyet azóta más nyelvekben is megismételtek, ismeretes, hogy a [b,d,g] azonosítása az őket követő magánhangzótól függ. Oden és Massaro éppen ezért vélekedik úgy, hogy a hosszú idejű memória nem tartalmazhatja a /d/ fonémát egyetlen prototípus formájában, mivel a [du] és a [di] között például akkora a különbség, hogy a megfeleltetés nem képzelhető el ugyanazzal a prototípussal. Mivel nyelvileg mégis ugyanazon fonémáról van szó, a döntésnek feltétlenül szótag-nagyságrendűnek kell lennie, illetve a prototípus minősége kell, hogy változzon (1978, 188). Még egy érveléssel alátámasztható a hangkapcsolat (szótag) mint döntési alap. Ez a koartikulációs következményekkel, a kontextus hatásával hozható összefüggésbe. A kutatók egyetértenek abban, hogy a magánhangzók felismerhetősége vizsgálható hangkörnyezet nélkül is. Nincs egyetértés ugyanakkor a tekintetben, hogy vajon a magánhangzó korrekt azonosításához szükség van-e egyáltalán hangkörnyezetre, és ha igen, milyen mértékben. A folyamatos szövegből kivett magánhangzók felismerésére kapott eredmények ugyanis eltérőek: egy részük szerint a magánhangzó tiszta fázisának informá-

ciója ez esetben is elegendő a biztos felismeréshez, más részük szerint viszont a hangkörnyezet elősegíti, pontosítja az azonosítást (Fry--Abramson 1962; Ochiai--Fujimura 1971; Kent 1975; Strange et al. 1976; Macchi 1980; Kuwabara 1982; stb.).

Folyamatos szövegből szegmentált magánhangzók azonosítása

Kísérletet végeztem, hogy megállapítsam, a különböző magyar magánhangzók más-más hangkörnyezetből kivágva, milyen mértékben őrzik meg a beszédhang biztos azonosításához szükséges akusztikai információkat. (Nem volt célom, hogy minden magánhangzó minden releváns fonetikai helyzetben és hangkörnyezetben többször is előforduljon.)

Anyag és módszer

Férfi bemondóval magnetofonszalagra rögzítettünk egy 60 szóból álló összefüggő szöveget (felolvasásban). A szöveg közömbös tartalmú, a felolvasás érzelemmentes volt. A kísérlethez ebből a szövegből a hangspektrográf szegmentátorának segítségével magánhangzókat "vágunk ki", az [a:] és az [e:] kivételével csak nyelviileg rövid időtartamúakat. Összesen 25 magánhangzót választottunk ki különböző fonetikai helyzetekből, különböző szótagszámú szavakból. A szegmentáláskor a magánhangzók tiszta fázisát hagytuk meg, ügyelve arra, hogy átmeneti részek ne maradjanak. Az így kapott magánhangzókat random sorrendben magnetofonszalagra rögzítettük és 20 kísérleti személynek (egyetemi hallgatók) lejátszottuk. Feladatuk a magánhangzók betűjelének leírása volt. A szegmentált magánhangzókat intenzitás-görbéjük és oszcillogramjuk alapján elemeztük, meghatároztuk viszonylagos intenzitásukat és fizikai időtartamukat.

Eredmények

A 3. táblázat a magánhangzók felismerését mutatja százalékos megoszlásban, a 4. táblázatban a korrekt azonosítás csökkenő arányában az intenzitás- és időtartamértékek láthatók.

3. táblázat

Magánhangzó	Fonetikai helyzet		szóvég (%)
	szó eleje	szó belseje	
a:	100	93	
ɔ	100	24	
ɔ		62	
u		33	
ø	100	87	

y		46	
e	13	93	90
e:		26	
i		69	

4. táblázat

Azonosítás (%)	Magánhangzó	Időtartam (ms)	Intenzitás (dB)
100	ø	70	32
100	ɔ	110	26--35
100	a:	100	18--24
100	e	90	18--3
93	e	70	35
93	o	100	18--28
93	a:	90	18--28
93	i	80	32
87	ø	85	32
80	e	80	30
73	i	60	26
70	o	90	34
53	y	70	30
47	u	75	32
40	y	80	28
40	ɔ	65	35
40	i	90	29
33	ɔ	80	28
27	o	70	36

26	e:	100	33
20	u	45	30
13	e	70	36
-	o	70	36
-	i	70	28

Az eredmények értékelésében tudatosan óvatosak vagyunk, mivel 25 magánhangzó adataira támaszkodunk csak. Úgy tűnik azonban, hogy a folyamatos szövegből szegmentált magánhangzók felismerése lehet egyértelműen korrekt és bizonytalan, sőt felismerhetetlen is. A kérdés tehát az, hogy melyek azok a tényezők, amelyek meghatározzák a korrekt vagy nem korrekt azonosítást. Felsorolásban (rangsorolás nélkül) eredményeink szerint ezek a következők: a) a magánhangzó tiszta fázisának fizikai időtartama, b) az aktuális hangkörnyezet, c) a magánhangzó spektruma: a fonéma realizációs lehetőségein belül hol helyezkedik el (középső részen vagy szélsőbb pontokhoz közel) és d) valószínűleg a magánhangzó-minőség: a nyelvemelkedés fokát illetően. E feltevéseket az adatok a következőképpen igazolják. A 70 % alatti azonosítások jobbra a 80 ms-os vagy annál rövidebb időtartamú magánhangzókra adódtak. A kevésbé korrekt felismerések inkább a felső nyelvállású -- akusztikailag az alacsony F1-gyel rendelkező -- magánhangzókra érintették. Az intenzitás -- úgy tűnik -- mint elsődleges kulcs nem érvényesül. A hangkörnyezet hatását feltételezik, illetőleg a formánsszerkezettel kapcsolatosak a fentebbieknek látszólag ellentmondó adatok. Például a 80 ms-os [ɔ] felismerése 33 %, a 70 ms-osat egyáltalán nem voltak képesek felismerni, pedig intenzitása mintegy 4 dB-lel nagyobb volt. Első szótagi [i] felismerése 73 %-os, utolsó szótagi csak 40 %-os. A formánsszerkezetben, illetőleg a belső intenzitásvizonyokban jelentkező különbség eredményezte a szükséges szó [e] és az esetben szó első [e]-je közötti azonosítási különbséget. Az első esetben 93 %-os, a másodikban csak 13 %-os a helyes felismerés. A formánsszerkezetben és a belső intenzitásvizonyokban jelentkező eltérést a hangszíneképek alátámasztották. Ezek az eredmények azt igazolják, hogy a hangkörnyezet részt vesz a magánhangzók korrekt felismerésének biztosításában, a fonéma-döntés egysége tehát itt is nyilvánvalóan nagyobb, feltételezhetően hang-

kapcsolat nagyságrendű szekvencia.

Az akusztikumban miként jelentkezik a hangkörnyezeti hatás? Két kísérletsorozat eredményeinek összevetésével próbáltam e kérdésre választ kapni. Mindkét esetben különböző frekvenciákon szűrtem egyrészt izolált ejtésű magánhangzókat, másrészt VC és CV típusú természetes hangkapcsolatokat (ugyanazon bemondó ejtésében). A magánhangzók felismerésére kapott adatokat vettem össze. Eszerint: hangkapcsolatban általában 40--50, néhány esetben 80--90 %-kal jobb volt a magánhangzók azonosítása, mint izolált ejtésben. A kontextus tehát elősegíti a fonémadöntést. A számadatok azonban arra nem adnak választ, hogy miben ragadható meg a kontextus részvétele. Ha elfogadjuk azt a feltételezést, hogy a kontextus "funkciója" tulajdonképpen a koartikulációs hatásban gyökerezik, akkor ennek jelentkeznie kell az akusztikai hullámformában is. Hangszínképeken megvizsgáltam a nagy percepció el-térést mutató magánhangzókat izolált és hangkapcsolati előfordulásban. A kísérleti személyek az [u]-t például izolált ejtésben egyáltalán nem ismerték fel, az [o]-t 15 %-ban. Hangkapcsolatban (a [z] mássalhangzóval) az [u]-t 80 %-ban, az [o]-t 85 %-ban pontosan azonosították. A magánhangzók akusztikai szerkezetében a lényeges különbség nem az átmenetek meglétében, illetőleg hiányában jelentkezik, hanem abban, hogy nagyobb intenzitással szerepelnek a magánhangzó tiszta fázisában azok a komponensek, amelyek izolált ejtésben nem vagy oly gyenge intenzitásúak, hogy a felismerésben gyakorlatilag alig vesznek részt. A többletinformációt jelentő komponensek érvényre jutása természetesen nem független a hangkapcsolat típusától, a fonetikai helyzettől. Például a szi azonosítása jobb eredményt adott, mint az isz-é.

Mindezek meggyőztek arról, hogy a fonológiai szinten hozott döntés egy magánhangzó és egy vagy több mássalhangzó kapcsolata alapján történik. Ezek a döntések azonban -- a beszédhangok különbözősége miatt -- lehetnek abszolút biztosak és kevésbé biztosak is, mivel az azonosíthatóságot a képzés, a koartikuláció, illetve az ezek következtében kialakuló akusztikai jelsorozat aktuális paraméterviszonyai határozzák meg.

Mindezek alapján elmondható, hogy a fonémaszinten lejátszódó művelet tulajdonképpen egy "többválasztásos feladatsor" megoldása. A választási lehetőségek a következők:

a) a beszédhang egyértelműen besorolható a megfelelő fonémaosztályba, például [a:], [j], [f],

b) több beszédhang ugyanabba a fonémakategóriába kerül, például a mák szó [m]-je, a hamvas mássalhangzó-kapcsolata,

c) mássalhangzó-kapcsolatban lévő beszédhangról el kell dönteni, hogy fennállhat-e a hasonulás vagy az összeolvadás jelensége,

d) a feltételesen azonosított, valószínűsített beszédhangról el kell dönteni, hogy valójában melyik fonémának felel meg, például [b,d] vagy [g]. Megpróbáltuk ábrázolni a fonémadöntés műveletsorozatát. Megjegyezzük, hogy egy viszonylag egyszerűbb jelfeldolgozást választottunk példánkknak. A 12. ábrán egy magánhangzó/mássalhangzó/magánhangzó típusú hangsor fonémaosztályba sorolásának feltételezett mechanizmusa látható. A folyamatára jellemű rajzról a következők olvashatók le. Ahol a 'fonéma' szó látható, ott a kategorizálás, tehát a megfelelő fonémaosztályba sorolás megtörtént. Ahol a művelet nem végezhető el egyértelműen, ott a megadott szabályok szerint kell a rendszernek eljárnia. Ez bináris kérdések sorozatával oldható meg. Az első beszédhangról megtudjuk, hogy magánhangzó. A másodikról is megkérdezzük, hogy magánhangzó-e. Ha a válasz nemleges, akkor mássalhangzót talált a rendszer. Ezt követi a szabály alkalmazása: a beszédhang megfeleltethető-e közvetlenül a fonémának vagy nem. Látható, hogy a 'nem' válasz két lehetőséget rejt magában: a nazálisok osztályát és a zárhangokét. Az 'igen' válasz mögött a nyelv azon szabályai állnak, amelyek a szomszédos hangok egymásra hatásával kapcsolatosak. Ezekben az esetekben a program a következő hang elemzésére utasítja a rendszert. Példánkban a következő hang magánhangzó, így ennek fonémaosztályba sorolása után a megelőző hangé is egyértelműen elvégezhető lesz.

Négy egy szótagú szóval illusztráljuk a feltételezett működést, ezek: ima, Ida, Ila, Ica. E szavak választását az indokolta, hogy ezekben az esetekben jól szemléltethető az ábrán látható működéssorozat lefolyása. Természetesen további hangsorok is "létrehozhatók", például inya, abó, esó, áru. A folyamatára nem jelzi külön az egyes beszédhangokra vonatkozó észlelési stratégiákat ugyanazon fonémaosztályon, sőt ugyanazon beszédhangcsoportokon (például nazális vagy magánhangzó) belül sem. A "befejezetlen utak" a megadott szabályok szerint folytathatók hosszabb hangsorok, akár mondatok azonosítására is. (Az egyes beszédhangosztályokon belüli differenciálás -- szándékunk szerint -- implikálva van.)

MAGYAR BESZÉDHANGOK AZONOSÍTÁSA

Flanagan írja, hogy a nyelvi elemek: fonémák, szótagok stb. vizsgálata segít azoknak a faktoroknak a meghatározásában, amelyekről azonosításuk függ (1965). A magyar fonetika utóbbi másfél évtizedes fejlődése, a Lazicius után (1944) gyakorlatilag folytatás nélkül maradt akusztikai fonetika megerősödése, a fejlett technika beépülése a kutatások kiszolgálásába; megteremtette a lehetőséget a kísérleti fonetika, azon belül pedig a perceptiókutatás kialakulására.

Felismerési kulcsok

Az invariáns jegyet ún. felismerési kulcsokkal fejezhetjük ki. Ezek az egzaktan kifejezhető felismerési kulcsok akusztikai-fonetikai kulcsok (Fujisaki 1980; Pisoni 1981). Sokszor tulajdonság-, illetve jegykeresőként is találkozunk ugyanezzel a fogalommal (Studdert-Kennedy 1980). Amikor a felismerési kulcsokat határozzuk meg, tulajdonképpen az akusztikum és a perceptió közötti összefüggéseket tárjuk fel a három paraméter segítségével. Ebben az értelemben az akusztikus invarianciát a következőképpen definiálhatjuk: az adott beszédrészletnek azon jellemzője, amely minimálisan szükséges a helyes felismeréshez; a beszédrészlet vagy egyetlen hang, vagy maximálisan két beszédhang kapcsolata legyen. A kísérleti eredmények azt mutatják, hogy gyakran nem egyetlen, hanem két vagy több felismerési kulccsal is számolnunk kell (ún. 'multiple cues': Oden--Massaro 1978, 178). Van olyan nézet, amely kontextustól függő és változó felismerési kulcsokról beszél, illetőleg elsődleges és másodlagos kulcsokról (Stevens 1975). A másodlagos felismerési kulcsok akkor jutnak szerephez, ha az elsődleges kulcsok eltűnnek vagy annyira torzulnak, hogy használhatatlanok. A másodlagos kulcsok lényegében redundáns tényezők. Gyakori ugyanis, hogy a beszédhullám különféle szimulán felismerési kulcsokat tartalmaz, bár közülük egy is elegendő az azonosításhoz. Megjegyzendő ugyanakkor, hogy ezek az elméletek és vélemények

még csak említést sem tesznek irreleváns elemekről; aminek feltehetően az oka, hogy a kutatás mai állása szerint nem állapítható meg egyetlen akusztikai elemről sem, hogy redundáns-e vagy irreleváns-e az észlelés számára. Csisztovics írja, hogy amikor a fonetikusok "kulcsokat" keresnek az azonosításhoz, főként átmeneteket találnak. Sok esetben pedig egészen másoknak (is) fontos szerepük van, például a formánsátmenet irányának, időszervezeti állandóknak, vö. 1980, 67.

A beszédet jellemző három akusztikai paraméter mint felismerési kulcs funkciójáról számos kísérletet végeztek. Mindezek alapján a leglényegesebb eredmények a következőkben foglalhatók össze: a) a hallási rendszer eleve rendelkezik bizonyos kategorizáló képességgel, amely már újszülötteknél is kimutatható nyelvi jelek közvetítésekor. Két tény bizonyítja, hogy ekkor még a nyelvi tapasztalattól független észlelés történik. Az egyik, hogy például tengerimalacokkal végzett kísérletek is hasonló eredménnyel zárultak (itt az evolúció, illetve a hallószervi megfelelés eredményez azonosságot). A másik az, hogy az újszülöttek valamennyien (függetlenül leendő anyanyelvüktől) azonos módon reagáltak a kísérleti nyelvi anyagra -- például a zöngés/zöngétlen mássalhangzók megkülönböztetésekor --, noha az egyes nyelvek beszélt formája e tekintetben is nagy eltéréseket mutat. b) A beszédhangok akusztikai szerkezete előjelzi a felismerési kulcsokat; egyesek -- például a magánhangzók -- ún. tiszta fázisukban meghatározhatók, mások -- mint a felpattanó zárhangok -- csak bizonyos hangkörnyezetben jellemezhetők. c) Az egy nyelven belüli beszédhangok felismerési kulcsa természetesen legkülönbözőbb, mind az akusztikai paraméterek kihasználtsága, mind pedig a redundanciafok tekintetében. d) Vannak felismerési kulcsok, amelyek csak egyetlen nyelven belül bírnak jelentőséggel, mások "nyelvfüggetlenek" abban az értelemben, hogy ahol ugyanazon beszédhangok realizálódnak, ott a kulcsok is azonosak. e) Egzakt adatokkal megadták a magánhangzók, illetve a mássalhangzók felismerési kulcsait több nyelvben (a magyarra ld. Gósy 1986). f) Meghatározták azt a frekvenciasávot, amely az ún. beszédfrekvenciákat tartalmazza, s amely szükséges például a beszélő azonosításához. Ez a 200--3500 Hz-es tartomány.

Mindezek fontos eredmények, de szinte egyikük sem támadhatatlan. A legnagyobb nehézségek közül kettőt emelek ki: a beszédhangok elsődleges/másodlagos felismerési kulcsának meghatározási nehézségeit és az ún. beszédfrekvencia-sáv bizonytalansági tényezőit. Az eddigiekből nyilvánvalóvá vált, hogy a beszédhangok egyetlen vagy egyik felismerési kulcsa a frekvencia

paraméterével fejezhető ki. Különösen vitathatatlanak tűnt ez a magánhangzók észlelési vizsgálataiban. Viszonylag régóta ismeretes, hogy az első két formáns meghatározza a magánhangzó hangszínét, biztosítja az egyértelmű azonosítást. Pontosabb megfogalmazás szerint a két első formáns egymáshoz való viszonya látja el ezt a funkciót (Fry--Denes 1958). A két formáns meghatározása valóban a magánhangzó-észlelés alapja, de szükséges még az időtartam megjelölése is. Az időtartam ugyanis bizonyos korlátok között nagy hatással van a magánhangzó-minőség percepciójára. A változatlan frekvenciaszerkezet és a változó időtartam között az alábbi szabályszerűség érvényesül a magyarban (a tendencia nyelvfüggetlen). A csökkenő időtartam hatására a változatlan akusztikai szerkezetű magánhangzók magasabb nyelvállású hangminőségekbe csapnak át; és megfordítva: a növekvő időtartam az alacsonyabb nyelvállású magánhangzó képzetét (azonosítását) eredményezi. Felvetődik tehát, hogy a magánhangzók esetében az időtartam (mint fizikai és nem nyelvi időtartam) elsődleges (kiegészítő) vagy másodlagos tényező. Hasonló a helyzet a résmássalhangzók esetében is.

Mind a telefontechnikában, mind az audiológiai gyakorlatban a 200/900--3500 Hz-es sávot tekintik a beszédre jellemzőnek (az átvitelben alsó határuk a 900 Hz-et adják meg, amely mellett a beszélő személye még felismerhető). Mindennapi tapasztalat ugyanakkor, hogy az akusztikailag ebben a tartományban "lezajló" telefonbeszélgetések gyakran nehezen érthetők, különösen, ha nevek azonosításáról van szó vagy hirtelen témaváltás történik. (Mindez nem magyarázható minden esetben azzal, hogy a megadott frekvenciatartomány a valóságban meglehetősen változó, továbbá, hogy egyéb körülmények is nehezíthetik a telefonálást.) Másfelől, az az ember, akinek 3500 Hz-től kezdve erőteljes halláscsökkenése van, komoly gondokkal küzd a természetes kommunikációban is. Azt mondhatjuk, hogy jó körülmények között, feltételezett ép hallás és ép beszédmegértési folyamat esetén a 200--3500 Hz-es frekvenciasáv tartalmaz annyi felismerési kulcsot a folyamatos beszédre, hogy az elhanyagolható hibaszázalékkal érthető marad. (A [s] és a [ʃ] elkülönítése ebben a tartományban bizonytalan.) A legújabb kísérleti eredmények arra engednek következtetni, hogy hasonló sajátságokkal más frekvenciatartományok is rendelkeznek a felsőbb frekvenciák területén (vö. 5. táblázat). Ez ismét továbbgondolásra készíti az invariáns jegyek akusztikai értelmezésével kapcsolatosan (vö. Gósy 1987).

5. táblázat

beszédrészlet	Elsődleges akusztikai kulcsaiktól megfosztott beszédegységek észlelése/megértése (%)	
	2200--2700 Hz	2700--3300 Hz
magánhangzók	80,7	78
logatomok	78	75
szavak	98	95

A magánhangzók percepció vizsgálat

A magánhangzókat jellemző akusztikai jel sajátosságai a következők: frekvenciaspektrum, formánssávzsélesség, formánsintenzitás és időszerkezet. A legmegbízhatóbb eredmények mellett tanúskodnak, hogy döntően a spektrális eloszlás a meghatározó a magánhangzók percepciójában (Miller 1951, 63). Bármely magánhangzót az első két formánsa alapján fel lehet ismerni (Flanagan 1965, 212; Simons 1980, 29). Pontosabb az a meghatározás, amely a magánhangzók azonosítását a formánsok (egymáshoz viszonyított) relatív távolságában jelöli meg (Fry--Denes 1958, 43).

Az alaphangmagasság változása -- a felismerés szempontjából -- nem lényeges tényező (Miller 1951, 73--4). (Ennek az az akusztikai magyarázata, hogy az alaphangmagasság változása -- normál férfi- és női hangot tekintve -- nem befolyásolja lényegesen a frekvenciaszerkezetet.) Mégis ismertek olyan kísérleti eredmények, amelyek szerint a magánhangzók felismerésében fontos szerep jut az alaphangmagasságnak. Scheffers magánhangzókkal végzett megkülönböztetési feladatában a kísérleti személyek jól teljesítettek, ha azonos volt az alapfrekvencia, ha azonban különböző volt -- s a különbség fél hangnál kisebb -- a válaszok elbizonytalanodtak (1979).

Fontos tényezője a magánhangzó-azonosításnak az időtartam; számos kísérlet ismert a kvalitás és a kvantitás viszonyáról a különböző nyelvekben (Cohen--t'Hart 1963; Onishi 1967; Ainsworth 1972; Scott 1976; Remington 1977; Scheffers 1979; stb.).

Magyar magánhangzók felismerésével kapcsolatosan Magdics végzett kí-

sérleteket (1970). A nyelvi anyagot 11 izoláltan és egyszavas mondatokban előforduló magánhangzó adta (nem vizsgálta az [i:, e:, u:] hangokat). Eredményei szerint a formánsfrekvenciának van a legnagyobb hatása a felismerésre (az intenzitás és az alaphangmagasság mellett). Ugyancsak az első két formánsban jelöli meg az elsődleges akusztikai kulcsot, de fontosnak tartja a második formáns intenzitását is. A harmadik formáns, az intenzitás és az alaphangmagasság változása nehezen meghatározható, bizonytalan magánhangzókat eredményezett. Ezek tehát, úgy tűnik, másod- vagy még inkább harmadlagos felismerési kulcsok. Egyáltalán nem vizsgálta ugyanakkor az időtartam szerepét.

Két kísérletsorozatot terveztem meg és folytattam le az összes magyar magánhangzó percepcióis elemzéséhez: elsősorban a frekvenciaszerkezetet, másodsorban a frekvencia- és az időszerkezet összefüggését vizsgáltam.

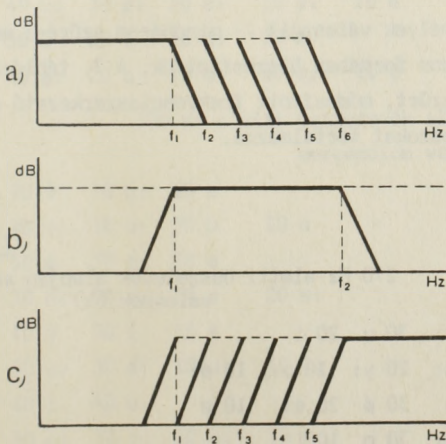
A frekvenciaszerkezet szerepe a felismerésben

A kísérletsorozat célja annak vizsgálata, hogy a frekvenciaszerkezetben bekövetkező változás milyen hatással van a magánhangzók felismerésére, illetőleg hogy a frekvenciaszerkezet egyes részletei milyen információt hordoznak. Mi a magasabb indexszámú formánsok szerepe -- ennek vizsgálata azért is indokolt, mivel a magánhangzók azonosításában betöltött szerepét nemigen tanulmányozták (vö. Fant 1973, 168).

Anyag és módszer

Természetes ejtésben, férfi bemondóval, izolált magánhangzókat rögzítettünk magnetofonszalagra. Az összes magyar magánhangzó normatív ejtésben, random sorrendben szerepelt: [ő, u:, a:, e, i, o, y:, ű, e:, i:, ő:, y, o:, u]. Ezt a hanganyagot 36 dB/oktáv meredekségű F-J Electronics típusú frekvenciaszűrőn engedték át úgy, hogy: a) az első esetben alul áteresztő, b) a második esetben sáváteresztő és c) a harmadik esetben felül áteresztő szűrővel dolgoztunk, a műszer lehetőségeinek megfelelően a 270, 330, 390, 470, 560, 680, 820, 1000, 1200, 1500, 1800, 2200, 2700 és 3300 Hz-es levágási frekvenciákkal (13. ábra). Ez azt jelentette, hogy az első (a) esetben 270--1200 Hz-ig -- minden frekvenciaértéken -- a tizennégy magánhangzónak az adott érték alatt komponensei maradtak meg, attól felfelé pedig 36 dB/oktáv volt a csillapítás. A második (b) esetben a sáváteresztő szűrő

alsó frekvenciájának beállítását változtattuk 270 és 2200 Hz között, a felső levágási frekvenciát állandó értéken, 2700 Hz-en tartottuk. Így fokozatosan szűkülő sávon át eresztettük át az egyes magánhangzókat: 270--2700 Hz, 330--2700 Hz, 390--2700 Hz, 470--2700 Hz stb. k ö z ö t t . A harmadik (c) esetben 1000 Hz-től felfelé az adott értékektől magasabb frekvenciákon lévő komponensek maradtak meg, például az 1000, az 1500 vagy a 2000 Hz-en f e l ü l i e k . A szűrő levágási meredeksége itt is 36 dB/oktáv volt. Az alul áteresztő szűrő használatával így 9, a sáváteresztő szűrővel 22, a felül áteresztő szűrővel pedig 7 sorozatot kaptunk. Összesen 28x14 -- módosított formánsszerkezetű -- magánhangzót vizsgáltunk. Az eredeti és a szűrt magánhangzókról 700-as hangspektrográffal hangszíneképeket (8 kHz-ig, széles sávú szűrővel, magas kiemeléssel) és amplitúdómetseteket (8 kHz-ig, keskeny sávú szűrővel, magas kiemeléssel) készítettünk (akusztikai adataikat ld. a 6. táblázatban). A 28 szűrt magánhangzó-sorozatot magnetofonszalagra rögzítettük és 10 fő 28--40 év közötti felnőttel (7 különböző alkalommal) lehallgattattuk. A nagyobb fokú szűrések következtében a magánhangzók specifikus intenzitása természetesen csökkent, ezért ezekben az esetekben a lejátszó magnetofont nagyobb hangerőre állítottuk. Egy kísérleten belül a hangerősséget nem változtattuk. A válaszokból összesen 28x14x10 adatot kaptunk, vagyis majdnem 4000-et. Ezeket PDP 11/34-es számítógépre vittük és kiértékelítettük.



13. ábra

A magánhangzók szűréséhez felhasznált szűrőállások:
alul áteresztő (a), sáváteresztő (b), felül áteresztő (c)

6. táblázat

Magánhangzó	Fizikai idő- tartam (ms)	F1	F2	F3	(Hz)
a:	430	750	1300	2900	
ɔ	360	600	950	2800	
o	240	450	850	2900	
o:	420	440	800	2900	
u	240	250	700	-	
u:	420	250	600	-	
ø	240	450	1600	2300	
ø:	380	440	1800	2400	
y	210	250	2000	2500	
y:	420	250	2100	2600	
i	200	220	2700	3400	
i:	340	220	2700	3600	
e:	400	400	2400	3050	
ę	380	500	2000	2800	

Eredmények

A kísérleti személyek válaszait -- mindhárom szűrési mód esetében külön-külön -- táblázatos formában összesítettük. A 7. táblázat az alul át-eresztő szűrővel megszárt, módosított frekvenciaszerkezetű magánhangzókra adott percepciós válaszokat tartalmazza.

7. táblázat

Eredeti magánhangzó	270 Hz alatti összetevők alapján adott percepciós válaszok (%)
ɔ	50 u: 30 u 20 i
a:	60 u: 20 y: 10 o: 10 ø:
o	40 o 30 ø 20 e: 10 e
o:	40 o: 30 o 10 i
u	50 u 20 u: 10 i
u:	20 u: 10 o 10 u 10 y 10 i 10 i:
y	40 u 40 i 10 y

y:	20 a:	20 i	20 i:	10 y:	10 ø:		
i	40 u	30 y	10 i				
i:	30 i	30 i:	10 u	10 y	10 y:		
e	50 u:	30 u	10 y				
e:	30 ø:	10 o	10 o:	10 i	10 i:	10 e	10 e:
ø	70 u	20 o:					
ø:	30 o:	20 ø:	10 a:	10 o	10 u:	10 e:	

330 Hz alatti összetevők alapján adott
percepciós válaszok (%)

o	50 u:	20 o	20 ø:				
a:	60 y:	10 ø:	10 o:				
o	50 o	40 ø:	10 o:				
o:	50 ø	30 ø:	10 o	10 e:			
u	60 u	20 y	10 i	10 ø			
u:	50 u:	30 y:	20 ø:				
y	40 y	30 i	20 u	10 u:			
y:	60 u:	20 ø	10 y:	10 ø:			
i	50 u	40 u	10 i				
i:	40 u:	40 y:	10 i:	10 ø:			
e	60 u:	20 o:	20 y:				
e:	40 u:	30 ø:	10 o:	10 y:	10 ø		
ø	50 o	20 u	10 u:	10 ø			
ø:	40 ø:	20 o:	10 o	10 e:	10 ø		

390 Hz alatti összetevők alapján adott
percepciós válaszok (%)

o	70 o	10 o:	10 e				
a:	40 o:	30 u:	20 u	10 o			
o	50 ø	30 e:	20 o				
o:	30 o:	30 ø:	20 e	20 e:			
u	40 u	30 i	10 ø				
u:	50 u:	30 ø:	10 o	10 y:			
y	60 i	40 u					
y:	50 u:	30 i:	10 o:	10 y			
i	80 i	10 u	10 y				
i:	60 u:	10 y	10 y:	10 i	10 i:		
e	60 o:	10	10 o	10 u	10 u:		

e:	40 e:	10 o:	10 u:	10 y:	10 i:	10 ø	10 ø
ø	30 o	30 u	10 o:	10 ø	10 ø		
ø:	50 e:	20 o:	20 ø:				

470 Hz alatti összetevők alapján adott
percepciós válaszok (%)

ɔ	90 ɔ	10 o:					
a:	40 ɔ	40 o:	10 u	10 u:			
o	40 o	20 ɔ	20 e	10 y	10 ø		
o:	60 o	30 ø:	10 e:				
u	70 u	30 u					
u:	40 o:	20 u:	20 e:	10 ø:			
y	60 i	30 u	10 y				
y:	40 i:	30 u:	10 o:	10 y:	10 i		
i	40 u	40 i	20 y				
i:	70 u:	10 u	10 i	10 i:			
e	60 ɔ	30 o:	10 o				
e:	40 e:	20 o:	20 ø:	10 y:	10 ø		
ø	40 o	30 ɔ	20 e	10 ø			
ø:	60 o:	20 e:	10 ø	10 ø:			

560 Hz alatti összetevők alapján adott
percepciós válaszok (%)

ɔ	90 ɔ	10 e					
a:	90 ɔ	10 a:					
o	50 o	40 ø	10 ɔ				
o:	60 o:	30 ø:					
u	60 o	40 u					
u:	60 o:	10 o	10 u	10 u:	10 ø:		
y	70 u	20 i	10 ø				
y:	60 u:	30 i:	10 u				
i	60 u:	30 i	10 i:				
i:	60 u:	20 i:	10 u	10 y:			
e	80 ɔ	10 a:	10 e				
e:	70 o:	10 o	10 e:	10 ø:			
ø	40 ø	30 o	20 ø:	10 o:			
ø:	70 o:	10 o	10 ø:				

680 Hz alatti összetevők alapján adott
percepciós válaszok (%)

ɔ	90	ɔ	10	a:	
a:	100	ɔ			
o	70	o	20	o:	10
o:	100	o:			
u	60	u	40	o	
u:	70	o:	30	u:	
y	70	u	20	i	10
y:	70	u:	20	ɔ:	10
i	70	u	10	o:	10
i:	70	u:	10	u	10
e	90	ɔ	10	e	
e:	80	o:	10	o	10
ɔ	60	ɔ	20	o	20
ɔ:	80	o:	10	e:	10

820 Hz alatti összetevők alapján adott
percepciós válaszok (%)

ɔ	100	ɔ			
a:	100	ɔ			
o	80	o	10	e	10
o:	90	o:	10	a:	
u	70	u	20	o	
u:	80	o:	10	ɔ	10
y	60	u	20	i	10
y:	30	o:	30	u:	20
i	50	ɔ	20	o	20
i:	50	u:	30	u	10
e	100	ɔ			
e:	90	o:	10	e:	
ɔ	50	o	20	ɔ	10
ɔ:	50	o:	30	o	20

1000 Hz alatti összetevők alapján adott
percepciós válaszok (%)

ɔ	100	ɔ		
a:	80	ɔ	20	a:

o	80 o	10 ɔ	10 o:
o:	80 o:	10 a:	10 o
u	90 u	10 o	
u:	80 o:	20 u:	
y	70 u	20 i	10 y
y:	40 u:	20 o:	20 e: 10 y: 10 i:
i	80 u	10 i	10 e
i:	60 u:	30 u	10 o:
e	100 ɔ		
e:	90 o:	10 e:	
ø	50 o	20 ø	20 ø: 10
ø:	90 o:	10 o	

1200 Hz alatti összetevők alapján adott
percepciós válaszok (%)

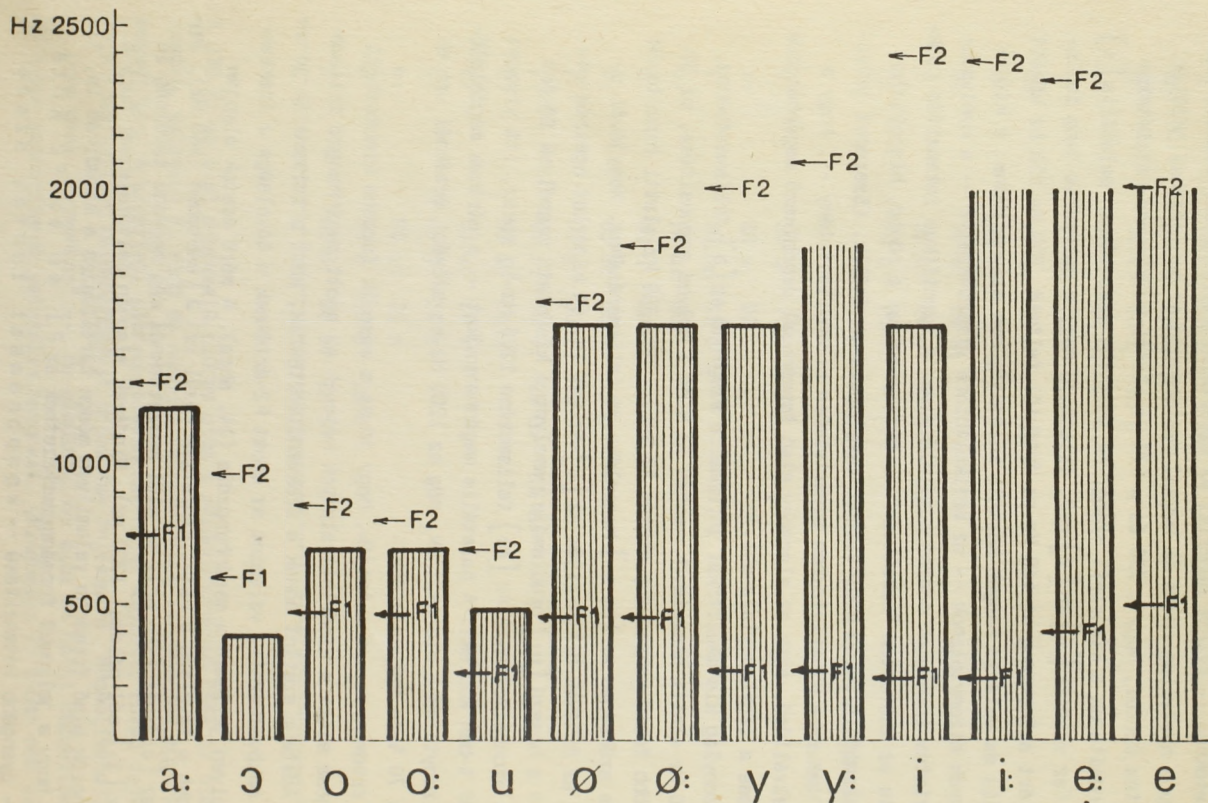
ɔ	10 ɔ		
a:	80 a:	20	
o	90 o	10 e	
o:	90 o:	10	
u	60 u	40 o	
u:	60 o:	30 u:	10 o
y	50 y	30 i	20 u
y:	40 i:	30 y:	20 u: 10 e:
i	50 u	50 i	
i:	80 u:	20 u	
e	80 ɔ	20 e	
e:	80 o:	20 o	
ø	50 ø	30 o	20 e
ø:	60 ø:	20 o:	20 e:

A 270 Hz-es alul átteresztő szűrés után a magánhangzókból csak az alaphangfrekvencia, valamint az első és második felharmonikus (vagy annak egy része) maradt meg. A kísérleti alanyok elsősorban hátul képzett, rövid vagy hosszú felső nyelvállású magánhangzókat próbáltak felismerni. Megemlítendő, hogy -- noha csak "nyomokban" --, de elkülönülnek elől és hátul képzettek; 30--40 %-ban azonosítanak palatális hangokat. (Mindössze 15 %-ban nem azonosít-

tottak beszédhangot.) Az alul áteresztő szűrő sávjának növelésével egyre több akusztikai információ jelenik meg az eredeti magánhangzóból. Ez -- bizonyos esetekben -- szinte nehezíti az egyértelmű felismerést. Gyakori például, hogy a nyelvileg rövid és hosszú magánhangzókat pontatlanul jelölik. Magyarázatra szorul, hogy a 390 és a 470 Hz alatti komponensek meglétekor miért azonosítanak a kísérleti személyek jelentős százalékban palatális [i] hangokat (az eredeti [i] és [y]-re), illetőleg hogy a palatális hang felismerése miért szűnik meg az 560 Hz-es értéktől felfelé. 350 és 550 Hz között a palatális és veláris hangok közötti különbség nem szignifikáns: a kísérleti alanyok bizonytalanok -- az artikulációra visszavezetve -- a nyelvmozgás vízszintes irányának pontos megítélésében. A specifikus intenzitás csökkent volta az időtartam "becslését" is befolyásolja: a rövid, bizonytalan hangot valamiféle felső nyelvállású magánhangzónak ítélik. (Szerepet játszhat ebben az -- a kísérletekben egyértelműen bizonyított tény --, hogy a felső nyelvállású, azaz az alacsony első formánssal rendelkező magánhangzók felismerése a legbizonytalanabb.)

Viszonylag kis akusztikai információ elegendő az [ɔ] felismeréséhez, már a 390 Hz alatti összetevők esetén 70 %-os a helyes azonosítása, ez 80 Hz-cel több formánsmaradvány esetén 90 %-osra nő. 780 Hz alatti összetevők jelenléte szükséges az [o, o:] hasonlóan jó felismeréséhez. Nagyjából ugyanezek az értékek biztosítják az [u]-értést is. (A kísérlet egészen végigvonult a hosszú [u:] problémája, amelyet a kísérleti személyek főként [o:]-nak azonosítottak.) Az [a:] felismerése 1200 Hz-ig rossz, itt viszont rögtön 80 %-ban korrekt. A palatális magánhangzókat -- a vártak megfelelően -- bizonytalanul azonosítják, még az 1200 Hz-es szűrési értéknél sem érik el a 70 %-os helyes arányt.

Az eredmények azt mutatják, hogy jóval a második formáns kimutatható megjelenése előtt a kísérleti alanyok képesek az adott magánhangzó felismerésére. Adataim szerint annak a frekvenciaértéknek, ahol a felismerés jó arányban bekövetkezett, valamint az adott F₂-értéknek a távolsága a frekvenciatengelyen jellemző a magánhangzóra (14. ábra). A mért adatok alapján legnagyobb a távolság a hátul képzettek közül az [ɔ] esetében (560 Hz), legkisebb az [a:]-nál (100 Hz); az elől képzetteknél még nagyobb távolság is akad (az [i:] -nél például 1200 Hz). Mindezek alapján feltételezem, hogy a magánhangzók első formánisa valamilyen módon tartalmazza a másodikat is, illetőleg hogy a jellemző formánsszerkezetnek az F₁ és az F₂ közé eső -- gyengébb intenzitású -- komponensei fontos szere-



14. ábra

Alul áteresztő szűrővel torzított magánhangzók felismerése
 (A felül zárt téglalapok a 70 %-os pontos azonosítást jelzik a megfelelő frekvenciaértéknél, a nyitott téglalapok az ennél gyengébb felismerést.)

u	80 u 20 o
u:	60 o: 40 u:
y	100 y
y:	100 y:
i	100 i
i:	100 i:
e	100 e
e:	100 e:
ø	100 ø
ø:	100 ø:

390--2700 Hz közötti összetevők alapján
adott percepciós válaszok (%)

ɔ	100 ɔ
a:	100 a:
o	100 o
o:	100 o:
u	100 u
u:	100 o:
y	100 y
y:	100 y:
i	100 i
i:	100 i:
e	100 e
e:	100 e:
ø	100 ø
ø:	100 ø:

470--2700 Hz közötti összetevők alapján
adott percepciós válaszok (%)

ɔ	100 ɔ
a:	100 a:
o	100 o
o:	100 o:
u	80 u 20 o
u:	100 o:
y	90 y 10 ø
y:	80 y: 20 ø

i	90 i	10 e
i:	90 i:	10 e:
e	100 e	
e:	100 e:	
ø	100 ø	
ø:	100 ø:	

560--2700 Hz közötti összetevők alapján
adott percepciós válaszok (%)

ɔ	100 ɔ		
a:	100 a:		
o	100 o		
o:	100 o:		
u	60 u	40 o	
u:	60 ɔ	20 o:	20 u:
y	50 y	50 ø	
y:	60 ø:	30 y:	10 y
i	90 i	10 e	
i:	90 i:	10 e:	
e	100 e		
e:	100 e:		
ø	100 ø		
ø:	100 ø:		

680--2700 Hz közötti összetevők alapján
adott percepciós válaszok

ɔ	100 ɔ		
a:	100 a:		
o	90 o	10 o:	
o:	90 o:	10 o	
u	70 u	30 o	
u:	70 ɔ	30 u:	
y	80 y	20 ø	
y:	50 y:	30 y	20 ø:
i	90 i	10 e:	
i:	80 i:	10 i	10 e:
e	100 e		
e:	100 e:		
ø	100 ø		

ø: 80 ø: 20 ø

820--2700 Hz közötti összetevők alapján
adott percepciós válaszok (%)

ɔ 100 ɔ

a: 100 a:

o 50 o 50 ɔ

o: 70 o: 30 ɔ

u 40 ɔ 40 u 20 o

u: 100 ɔ

y 100 y

y: 80 y: 20 ø

i 80 i 20 i:

i: 90 i: 10 ɔ

e 100 e

e: 90 e: 10 i:

ø 90 ø 10 e

ø: 60 ø: 40 ø

1000-2700 Hz közötti összetevők alapján
adott percepciós válaszok (%)

ɔ 100 ɔ

a: 100 a:

o 70 o 30 ɔ

o: 70 o: 20 ɔ 10 o

u 40 u 40 o 10 u:

u: 70 ɔ 20 u: 10 o:

y 100 y

y: 90 y: 10 ø:

i 80 i 20 i:

i: 100 i:

e 100 e

e: 100 e:

ø 100 ø

ø: 90 ø: 10 ø

1200-2700 Hz közötti összetevők alapján
adott percepciós válaszok (%)

ɔ 100 ɔ

a: 100 a:

o	90 o	10 o:		
o:	90 o:	10 o		
u	50 u	20 u:	10 o	
u:	60 u:	20 u	10 o	10 o:
y	90 y	10 y:		
y:	80 y:	10 y	10 o:	
i	80 i	20 i:		
i:	90 i:	10 i		
e	100 e			
e:	100 e:			
ø	100 ø			
ø:	90 ø:	10 ø		

1500-2700 Hz közötti összetevők alapján
adott percepciós válaszok (%)

o	90 o					
a:	90 a:	10 e				
o	30 o	30 u	10 o:	10 i	10 e:	10 ø
o:	70 u:	10 o:	10 u	10 i:		
u	40 u	20 i				
u:	40 u	40 i:	10 a:	10 u:		
y	100 y					
y:	70 y:	20 y	10 ø			
i	80 i	20 e:				
i:	90 i:	10 i				
e	60 e	20 ø:	10 y:	10 i		
e:	80 e:	20 i:				
ø	90 ø	10 y				
ø:	90 ø:	10 ø				

1800-2700 Hz közötti összetevők alapján
adott percepciós válaszok (%)

o	90 o	10 a:				
a:	90 a:	10 e				
o	30 o	30 u	10 o:	10 i	10 e:	10 ø
o:	70 u:	10 o:	10 u	10 i:		
u	40 u	20 i				
u:	40 u	40 i:	10 a:	10 u:		
y	100 y					

y:	70 y:	20 y:	10 ø:	
i	80 i	20 e:		
i:	90 i:	10 i		
e	60 e	20 ø:	10 y:	10 i
e:	80 e:	20 i:		
ø	90 ø	10 y		
ø:	90 ø:	10 ø		

2200-2700 Hz közötti összetevők alapján
adott percepció válaszok (%)

ɔ	100 ɔ			
a:	100 a:			
o	70 o	10 ɔ	10 o:	10 u
o:	70 o:			
u	50 u	10 i:		
u:	20 i:			
y	90 y	10 y:		
y:	70 y:	20 ø:	10 y	
i	100 i			
i:	90 i:	10 u:		
e	60 e	20 y:		
e:	100 e:			
ø	100 ø			
ø:	90 ø:	10 ø		

A táblázatban közölt adatok azt mutatják, hogy a magánhangzók nagyobb részének felismerése e sávon belül végig pontos. Ez megegyezik azzal az eredménnyel, amely szerint -- noha a kísérleti módszer bizonyos fokig eltérő volt -- az akusztikai jel (a burkológörbe) közepe elegendő információt tartalmaz a felismeréshez (Miller 1951, 72--73; megjegyzendő, hogy itt folyamatos beszédet vizsgáltak). Néhány adatra érdemes részletesebben is kitérni.

Az [o,o:] és [u] azonosítása a 820--2700 Hz-es tartománytól kezdve bizonytalanodik el. Ugyanakkor 1200 és 2700 Hz között átengedve a jelet, az [o]-t 90 %-ban, az [o:] -t ugyancsak, az [u]-t 70, az [u:] -t pedig 80 %-ban azonosítják helyesen. Ez azért feltűnő, hiszen ezeknek a hangoknak az első formánása jóval a megadott alsó érték a l a t t helyezkedik el. Ez ismét azt

a megállapítást támasztja alá, hogy a frekvenciaeloszlás végig jellemző a beszédhangra, valamint hogy a felsőbb frekvenciákon lévő formánsok, hasonlóan az F1- és F2-höz, biztosítják a korrekt felismerést. Más magánhangzók esetében is tapasztalható ez a "funkcióátvétel": az F2 és az F3 veszi át az F1 és az F2 szerepét, például az [e:,i] esetében, ahol mindkét első formáns az 1000 Hz-es érték alatti. A 2200 és 2700 Hz-es 500 Hz szélességű sávban az elöl képzett magánhangzók azonosítása pontos, valamennyinek a második formánása 2000 Hz fölé esik, vagy erősen megközelíti ezt az értéket. Az F3 fontos szerepét mutatja az [e] -- a többihez képest kissé gyengébb -- 70 %-os azonosítása. Itt ugyanis -- a többihez viszonyítva -- a 2800 Hz táján lévő harmadik formáns kissé gyengébb intenzitású.

Ha a magánhangzót jellemző első két formánsból csak az egyik áll rendelkezésünkre, s így kell a hangot azonosítanunk, akkor az adott nyelvre jellemző hangrealizációktól függően egy vagy több magánhangzót fogunk felismerni (vö. meghatározott frekvenciájú szinuszhangok beszédhangként történt azonosításáról: Fant 1973, 182). Például, ha az akusztikai szerkezetben egyetlen 400 Hz-es formáns van, akkor azt nagy valószínűséggel [o]-nak, [ø]-nek, esetleg [e]-nek fogjuk ítélni; illetőleg ha csak 2300 Hz táján van összetevő, akkor az nagy valószínűséggel az [i] vagy az [e:] benyomását fogja kelteni. Az a tény tehát, hogy a szűrőzései kísérletben az alanyok pontosan felismerték és megkülönböztették az [i:] -t az [e:] -től ott, ahol gyakorlatilag hiányoztak az első formánsok -- ugyancsak ahhoz a következtetéshez vezet, hogy az F2 és a fölötte lévő formánsok, illetőleg ezek viszonya éppolyan jellemző a magánhangzóra, mint az F1 és F2 helye, illetőleg viszonya. A második és harmadik formánsoknak a felismerésre gyakorolt hatását mutatja, hogy például hozzájárult az [u] és az [y] megkülönböztetéséhez (Fant 1973, 168). Az első két formáns jelenlétében a magasabb indexszámú formánsok mint redundáns elemek vannak jelen, az F3 frekvenciájának változtatása gyakorlatilag nincs hatással a fonémadöntésre (Gósy--Graaf 1983). A második és harmadik formánsok szerepét vizsgálta Fujimura a három felső nyelvállású svéd magánhangzó, az [i, y, u] felismerésére vonatkozólag (1967). A szintetizált magánhangzókkal végzett kísérlet eredménye szerint az F2, illetve az F3 (konstans F1, F4 és F5 esetén) biztosítja a hang korrekt azonosítását, de nem függetlenül a formánsfrekvenciáktól, illetőleg a formánsok távolságától.

c) A harmadik kísérletsorozatban 1000 Hz-től felfelé szűrtük a magán-

hangzókat úgy, hogy az adott értékek fölötti teljes frekvenciatartomány megmaradt. A tesztelés adatait a 9. táblázat tartalmazza.

9. táblázat

Eredeti magánhangzó 1000 Hz fölötti összetevők alapján adott
percepció válaszok (%)

ɔ	100	ɔ			
a:	100	a:			
o	90	ɔ	10	o	
o:	90	ɔ	10	o:	
u	100	ɔ			
u:	100	ɔ			
y	100	y			
y:	50	y:	30	ɔ:	10 y 10 ɔ
i	90	i	10	e:	
i:	80	i:	20	i	
e	100	e			
e:	100	e:			
ɔ	100	ɔ			
ɔ:	90	ɔ:	10	ɔ	

1200 Hz fölötti összetevők alapján adott
percepció válaszok (%)

ɔ	100	ɔ			
a:	100	a:			
o	100	ɔ			
o:	80	ɔ	20	o:	
u	70	ɔ	20	u	10 o
u:	80	ɔ	10	u	10 i:
y	90	y	10	ɔ	
y:	60	y:	20	y	20 ɔ:
i	100	i			
i:	70	i:	30	i	
e	100	e			
e:	90	e:	10	i:	
ɔ	90	ɔ	10	y	
ɔ:	90	ɔ:	10	y:	

1500 Hz fölötti összetevők alapján adott
percepciós válaszok (%)

ɔ	100 ɔ				
a:	100 a:				
o	50 ɔ	20 i:	10 u	10 i	10 ø
o:	40 i:	30 u:	20 o:	10 ɔ	
u	60 i	30 u	10 ɔ		
u:	90 i:	10			
y	80 y	20 ø			
y:	50 y:	40 ø:	10 u:		
i	60 i	40 e:			
i:	70 i:	20 i	10 e:		
e	90 e	10 ø			
e:	50 e:	40 e	10 i:		
ø	70 ø	30 ø:			
ø:	90 ø:	10 ø			

1800 Hz fölötti összetevők alapján adott
percepciós válaszok (%)

ɔ	90 ɔ	10 e:			
a:	90 a:	10 ø:			
o	60 i	20 i:	10 ɔ	10 ø:	
o:	60 i:	10	10 u	10 i	10 e:
u	70 i	20 u	10 i:		
u:	100 i:				
y	70 y	20 ø	10 i:		
y:	50 y:	30 ø:	10 ɔ	10 u:	
i	80 i	20 e:			
i:	70 i:	10 i	10 e:	10 ø:	
e	90 e	10 ø:			
e:	60 e:	30 e	10 i:		
ø	90 ø	10 ø:			
ø:	80 ø:	10 y	10 ø		

2200 Hz fölötti összetevők alapján adott
percepciós válaszok (%)

ɔ	60 ɔ	20 e	20 e:
a:	60 ø:	20 a:	20 ø
o	60 i	40 i:	

o:	100	i:	
u	100	i	
u:	100	i:	
y	100	y	
y:	100	y:	
i	70	i	30 i:
i:	90	i:	10 i
e	40	e	40 ø: 20 y:
e:	90	e:	10 i:
ø	50	y	50 ø
ø:	60	ø:	40 y:

2700 Hz fölötti összetevők alapján adott
percepciós válaszok (%)

ɔ	40	ɔ	20	i	20	i:	
a:	90	ø:	10	a:			
o	90	i	10	i:			
o:	100	i:					
u	100	i					
u:	100	i:					
y	100	y					
y:	90	y:	10	ø:			
i	100	i					
i:	90	i	10	i:			
e	30	y	30	y:	30	ø:	10 e
e:	80	e:	20	i:			
ø	40	y	20	i			
ø:	50	y:	50	ø:			

3300 Hz fölötti összetevők alapján adott
percepciós válaszok (%)

ɔ	50	e:	40	i:			
a:	60	e:	30	i:	10	ø:	
o	80	i	20	i:			
o:	100	i:					
u	90	i	10	i:			
u:	90	i:	10	o:			
y	100	y					
y:	80	y:	20	y			

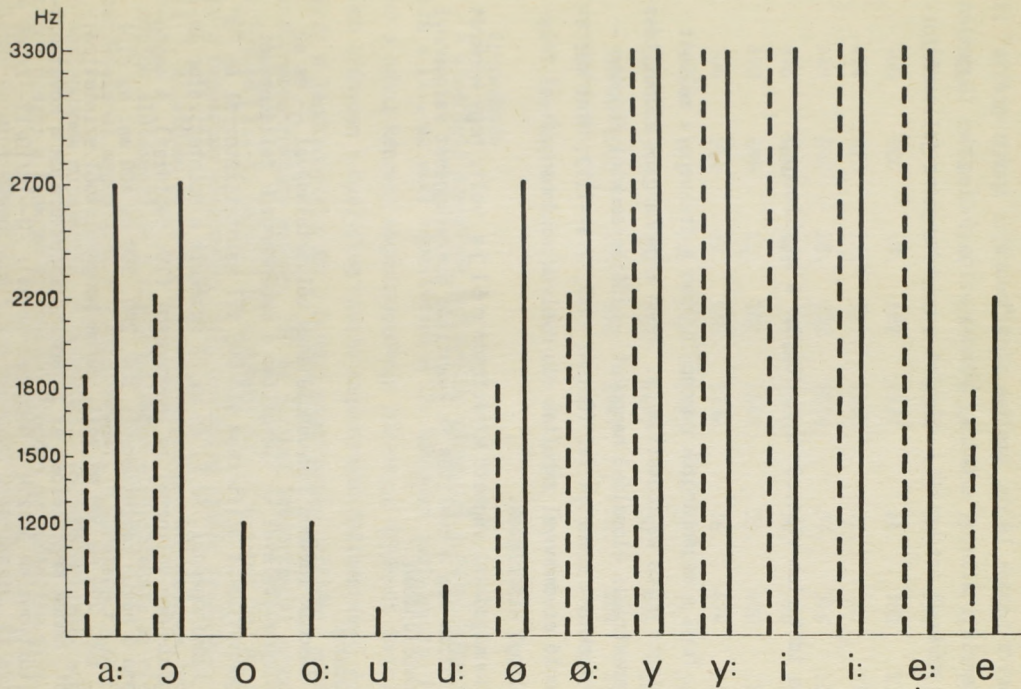
i	80 i	20 i:			
i:	90 i:	10 i			
e	50 ø:	20 y:	10 y	10 e:	10 ø
e:	90 e:	10 i:			
ø	60 i	30 y	10 ø		
ø:	60 y:	40 ø:			

A táblázatosan összefoglalt percepciós eredmények azt mutatják, hogy több magánhangzónk korrekt felismerése ilyen nagyfokú torzítás mellett is biztosított. Várhatóan a magas frekvenciájú F2-t tartalmazó hangok ezek: az [i, i:, y, y:]. Mind a négy beszédhang az F2 fölött a 3000-5000 Hz-es tartományban további közepes intenzitású komponenseket tartalmaz. Az eddigi eredmények meggyőztek arról, hogy ezek a magasfrekvenciás összetevők és egymás közti viszonyaik teszik lehetővé a biztos felismerést. Felvetődik azonban az, hogy miképpen lehetséges a labiális/illabiális megkülönböztetés enél a -- percepció szempontjából -- nagymértékben "károsodott" formánsszerkezetnél. A regisztrátumok azt mutatják, hogy a labialitásra, illetve az illabialitásra vonatkozó akusztikai információt a részösszetevők intenzitása biztosítja. (Teljes formánsszerkezet esetén Magdics az F2 frekvenciájának változásában jelöli meg az ajakműködésnek megfelelő akusztikai következményt; 1965.) Ismeretes: az illabiális magánhangzók intenzívebbek, mint a labiálisok (Fónagy 1966, 231--244) -- fizikai sajátságukból ez azért adódik így, mivel a labiálisok második formánsa valamivel alacsonyabb frekvenciájú, mint az illabiálisoké, tehát az F1 és F2 közel kerül egymáshoz, és mindkettő relatíve alacsonyan van (vö. például Bolla 1978), valamint azért is, mert az ajakkerekítésekénél nagyobb felületű az a nyílás (ajkak), ahol a levegő kitódul. A kísérletben felhasznált magánhangzók hangszíneképpen ugyanakkor jól látható, hogy az illabiálisok magasabb indexszámú -- főként az F2 -- formánsainak relatíve nagyobb az intenzitása a labiálisokéhoz képest. Az F1-t tekintve a labiális magánhangzóké az intenzívebb az illabiálisokkal szemben. Mindegyiknél az F2 intenzitása nagyobb, mint az első formánsé, tehát a hang specifikus intenzitásának kialakításában az F2-nek van nagyobb szerepe.

Az alsó nyelvvállású [e] és a középső nyelvvállású [ø, ø:] -k felismerése a 2200 Hz-es szűréstől bizonytalanodik el, az illabiális [e:] -é azonban

végig korrekt. Ugyancsak a 2200 Hz-es értéktől nem ismerhető fel jól a hátul képzett [ɔ, a:] sem. 1000 Hz-től kezdve gyengül a többi veláris magánhangzó azonosíthatósága. Amilyen mértékben torzul a formánsszerkezet, s többé-kevésbé bizonytalanodik a felismerés, jól megfigyelhető az időtartam megítélésének arányosan növekvő bizonytalansága. A nyelvileg rövid és hosszú hangok cserélődnek egymással a percepcióban, ami azt mutatja, hogy valamilyen hangminőség nagyjából egyöntetű megítélését a frekvenciaösszetevők lehetővé teszik, de a megfelelő időtartam kiválasztásához ez már nem elegendő.

E c) kísérlet adatait összevettem a b) eredményeivel. Arra kerestem választ, hogy vajon az ugyanazon értéktől, de nem ugyanazon sávban megszűrt hangok percepciójában tapasztalható-e különbség. Feltevésem az volt, hogy eltérés esetén azok a magánhangzók azonosíthatók jobban, ahol a szűrési értéktől felfelé megmaradt a frekvenciatartomány. Az adataim azonban nem ezt erősítették meg. Több magánhangzó esetében ugyan alig tapasztaltam a percepcióban különbséget (20 %); de kevés kivétellel mindig ott kaptam pontosabb válaszokat, ahol 2700 Hz fölött nem volt a hangban összetevő. Néhány esetben ([ɔ, a:, y:, i, ø]) csak egy-egy sorozatban jelentkezett lényegesebb eltérés. A korrekt válaszok javára 40--50 %-nyi az eltérés. Több sorozatban jelentős volt a különbség az [o, o:, u, u:]-knál; a rosszabb azonosítás itt is a teljesebb szerkezet megléte esetén mutatkozott. Például amíg az 1000 és 2700 Hz között megmaradó összetevők esetén a kísérleti alanyok 70 %-ban felismerik az [o, o:] -kat, addig az 1000 Hz fölött meglévő teljes formánsszerkezet esetében 90 %-ban [ɔ] -t hallanak. Hasonlóan az 1200 és 2700 Hz közötti sávban 70 %-os az [u] felismerése, ugyanakkor az 1200 Hz fölötti valamennyi komponens jelenlétében csupán 20 %-os (15. ábra). A lehetséges magyarázatot a következőkben látom. A magánhangzókat alkotó összetevők intenzitása adja a hang specifikus intenzitását. Az egyes komponensek észlelése azonban nemcsak a konkrét intenzitásértékektől, hanem (a frekvenciaértéken túl) az elfedési jelenségektől is függ. Tudjuk, hogy az elfedés jelensége az, hogy egy erősebb hang jelenléte megnehezíti vagy lehetetlenné teszi egy gyengébb hang észlelését. Az elfedő hang a gyengébb (elfedett) hangra nézve az elfedés idejére hallásveszteséget okoz (Halm 1963, 83; Pauka 1980, 96--102; Tarnóczy 1984, 130--137; stb.). (Az elfedés valószínűleg a hallóideghez tartozó jelenség.) A formánsszerkezet megváltozásával egyidejűleg számolnunk kell olyan elfedési jelenségekkel, amelyek korábban -- a hang teljes formánsszerkezetének ész-



15. ábra

A sáváteresztő (folyamatos vonal) és a felül áteresztő szűrő (szaggatott vonal) alkalmazásával kapott magánhangzók percepciósi eredményei (a 60 %-on felüli pontos azonosítás figyelembevételével)

lelésekor -- gyakorlatilag nem jelentkeztek. Minderre persze csak a percepció adatokból következtethetünk vissza; de magyarázója lehet annak, hogy a több akusztikai komponens nemhogy megkönnyíti, hanem ellenkezőleg: nehezíti a felismerést. A magasabb frekvenciák "járulékos" összetevői -- a domináns, a felismerést egyértelműen biztosító komponensek hiányában -- olyan konfigurációt hoznak létre, amelyek mint "neurális spektrográfia" ismeretlenek a kísérleti alanyok számára. (Mindez valószínűleg nem független a frekvenciasáv konkrét értékeitől, valamint a magánhangzóra jellemző formánsképtől.)

A frekvencia- és időtartam-változás összefüggése a magánhangzók azonosításában

Említettük, hogy a magánhangzók percepciójában a frekvencia mellett szimultán elsődleges kulcs az időtartam is. Japán magánhangzók azonosításakor ugyanolyan spektrális eloszlású hangokat -- időtartamuktól függően -- eltérő hangminőségekként ismertek fel (Onishi 1967, 714--715). Mint már utaltunk rá, hasonló eredménnyel zárultak más nyelvek magánhangzóival folytatott ilyen jellegű kísérletek.

Magyar magánhangzókkal végzett kísérletem célja volt, hogy konkrét értékekkel határozzam meg a minőség és kvantitás összefüggését valamennyi magánhangzó percepciójában.

Anyag és módszer

Az első és második formáns frekvenciájának változtatásával -- és az idő faktor figyelembevételével -- vizsgáltam a magánhangzók felismerését. Szintetizált magánhangzókat állítottam elő OVE III szintetizátorral és PDP 11/34-es típusú számítógéppel. Az F1 és az F2 értékeit a szintetizátor lehetőségeinek megfelelően változtattam lépésenként (10. táblázat). A magánhangzók időtartama a hátul képzettekénél: 100, 200, 300 és 400 ms, az elől képzettekénél 50, 100, 200 és 300 ms volt. Három hangszeletből szintetizáltam minden hangot, ennek megfelelően az intenzitásstruktúrájuk a következőképpen alakult (törtvonalak jelzik a hangszeleteket): /0--10/10/12--2/ dB a hátul és /0--10/10--14/12--2 dB az elől képzettekénél. Az alaphangmagasságot állandó értéken, 116 Hz-en tartottam; változatlan volt az F3 és az F4 értéke is: 1800 és 3500, illetve 2800 és 3500 Hz.

10. táblázat

Szintetizált magánhangzók formánsszerkezete

	F1	F2		F1	F2		F1	F2	(Hz)
1.	300	700	14.	599	1000	27.	389	1600	
2.	327	734	15.	635	1037	28.	436	1554	
3.	356	755	16.	654	1069	29.	490	1510	
4.	378	777	17.	673	1099	30.	550	1467	
5.	400	800	18.	692	1131	31.	300	2329	
6.	424	823	19.	713	1165	32.	327	2329	
7.	449	848	20.	734	1234	33.	356	2329	
8.	476	872	21.	755	1300	34.	378	2264	
9.	490	898	22.	206	1849	35.	400	2198	
10.	504	924	23.	238	1796	36.	424	2136	
11.	534	951	24.	275	1745	37.	449	2175	
12.	550	979	25.	317	1695	38.	476	2016	
13.	582	1000	26.	356	1647	39.	504	1958	

Eredmények

A kísérleti személyek percepciós válaszait a hátul képzett magánhangzókra a 11., az elől képzettekre a 12. táblázat tartalmazza. Mind az elől, mind a hátul képzett magánhangzók nál 300 ms-os értéknél fordulnak elő nagyobb arányban a nyelvileg hosszú magánhangzó-felismerések. Az [i, i:] hangpárnál a legbizonytalanabb a hosszú magánhangzó jelölése: mindössze egyetlen formánsstruktúra esetén, 300 ms fizikai időtartamnál azonosítanak 60%-ban hosszú [i:] -t, hasonló az eredmény az [u, u:] -nál is. Az adat mutatja, hogy a magyarban viszonylag kevés minimális szópár található, amelyek az [i] ~ [i:], illetőleg az [u] ~ [u:] magánhangzókban különböznek (pl. irat ~ írat). Az adatok alátámasztják a szakirodalomban olvasottakat: az első két formáns maradéktalanul meghatározza a magánhangzót, a konstans F3 és F4 a percepcióban semlegesnek mutatkozik. A frekvencia- és időszerkezeti adatok sajátosan függenek össze: a konkrét értékektől függően az időtartam mint szimultán elsődleges és mint másodlagos kulcs egyaránt jelentkezik. Ez azt jelenti, hogy ugyanazon frekvenciaadatok a fonémakategóriák határait másképpen jelölik ki az időtartamtól függően -- ekkor az időtartam szimultán el-

sődleges kulcs (vö. 16. ábra). A magyar magánhangzók legjellemzőbb formánsadatain belül azonban az időtartam nem játszik ilyen szerepet: ekkor másodlagos felismerési kulcs. Jelentős eltérést e tekintetben 100 és 200, valamint 300 és 400 ms fizikai időtartam esetén tapasztaltam; a 200 és 300 ms-os magánhangzók percepciójában minimális volt a különbség. Hasonló kísérletsorozatában közel ugyanilyen eredményekre jutott angol magánhangzók tesztelésével Ainsworth (1972). Következtetése szerint az időtartam hatással van a magánhangzók felismerésére, különösen azokban a határesetekben, ahol csupán az F1- és F2-értékek alapján történő döntéssel a magánhangzók nem azonosíthatók egyértelműen (651). Az időtartam mint felismerési kulcs funkcióváltozásait, illetőleg az azonosított magánhangzónak megfelelő frekvenciatartományokat a 13. táblázat tartalmazza. (A táblázatot a 70 %-on felüli egyöntetű döntések alapján állítottam össze, tekintet nélkül a nyelvileg rövid és hosszú magánhangzókra.)

13. táblázat

Időtartam (ms)	u/u:	Azonosított magánhangzók értékei (Hz)			
		o/o:	ɔ	a:	
100	300--378	400--534	550--673	692--755	
200	300--356	400--490	530--654	692--755	
300	300--356	378--490	530--654	692--755	
400	300--327	400--476	504--635	673--755	

Időtartam (ms)	i/i:	Azonosított magánhangzók értékei (Hz)			
		ɛ:	e	ø/ø:	y/y:
50	300--356	400	476--504	356--550	206--317
100	300--327	378--400	476--504	389--550	206--317
200	300	356--400	450--504	356--550	206--317
300	300	356--400	450--504	356--550	206--275

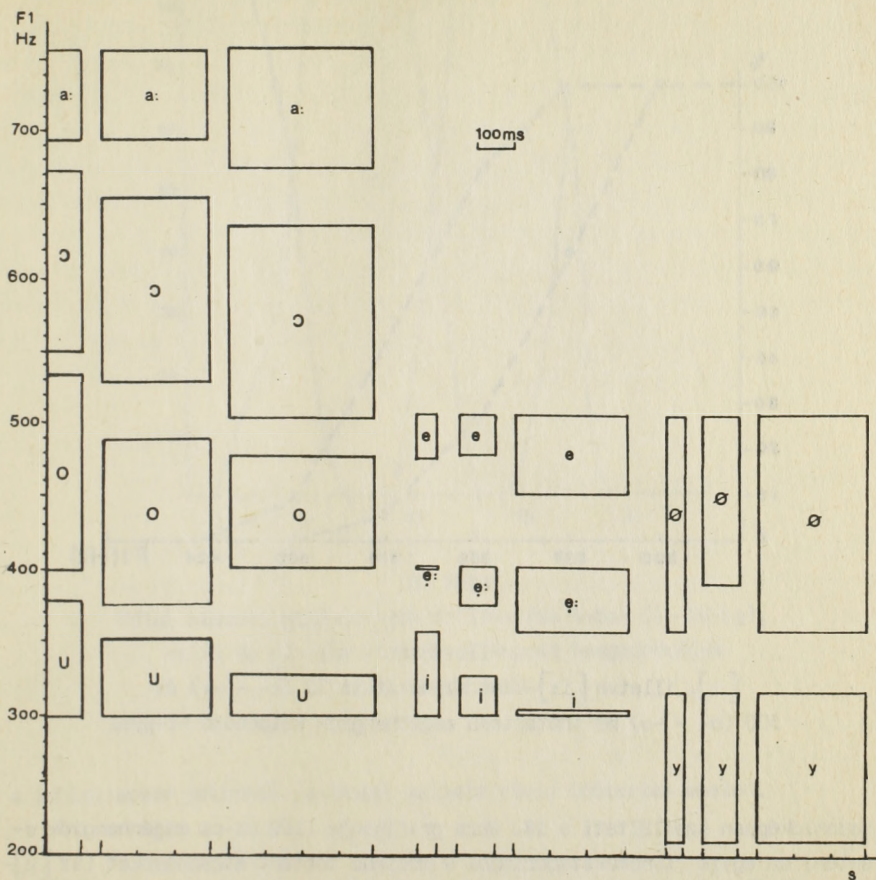
A magánhangzó időtartamának növekedésével az egyes fonémakategóriákban az első formáns frekvenciasávja szűkül. Az [u]-ra kapott adatoknál például 50 Hz-nyi, az [o]-nál csaknem 60 Hz az eltérés; az alsó és legalsó nyelvállású hangzóknál ez a "szűkülés" kisebb mértékű. Csaknem változatlan

Magán- hangzó F1-e (Hz)	Az azonosított magánhangzók															
	100 ms				200 ms				300 ms				400 ms			
	u/u:	o/o:	ɔ	a:	u/u:	o/o:	ɔ	a:	u/u:	o/o:	ɔ	a:	u/u:	o/o:	ɔ	a:
300	100/	-	-	-	100/	-	-	-	94/6	-	-	-	40/60	-	-	-
327	83/	17/	-	-	94/	3/3	-	-	86/7	7/	-	-	20/43	14/23	-	-
357	86/	13/	-	-	83/	17/	-	-	66/6	22/6	-	-	23/37	10/30	-	-
378	86/	13/	-	-	47/	50/3	-	-	34/3	50/13	-	-	33/27	10/30	-	-
400	33/	67/	-	-	3/	97/	-	-	/3	83/14	-	-	17/17	26/40	-	-
424	6/	94/	-	-	-	100/	-	-	-	90/10	-	-	10/	33/57	-	-
449	-	100/	-	-	-	76/	24	-	-	86/14	-	-	-	30/33	37	-
476	-	100/	-	-	-	73/	27	-	-	80/14	6	-	-	40/30	30	-
490	-	100/	-	-	-	80/	20	-	-	66/6	28	-	-	40/16	44	-
504	-	76/	24	-	-	40/	60	-	-	47/3	47	3	-	3/21	76	-
534	-	63/	37	-	-	27/	73	-	-	10/3	84	3	-	3/7	80	10
550	-	24/	76	-	-	-	100	-	-	-	100	-	-	/3	90	7
582	-	-	100	-	-	-	100	-	-	-	100	-	-	-	90	10
599	-	-	100	-	-	-	100	-	-	-	100	-	-	-	90	10
635	-	-	94	6	-	-	92	8	-	-	83	17	-	-	80	20
654	-	-	80	20	-	-	73	27	-	-	73	27	-	-	43	57
673	-	-	57	43	-	-	47	53	-	-	47	53	-	-	20	80
692	-	-	6	94	-	-	-	100	-	-	10	90	-	-	6	94
713	-	-	-	100	-	-	-	100	-	-	-	100	-	-	-	100
734	-	-	-	100	-	-	-	100	-	-	-	100	-	-	-	100
755	-	-	-	100	-	-	-	100	-	-	-	100	-	-	-	100

12. táblázat

Magán- hangzó F1-e (Hz)	50 ms				100 ms				200 ms				300 ms			
	i/e	y/y:	ø/ø:		i/e	y/y:	e	ø/ø:	i/e	y/y:	e	o/ø:	i/e	y/y:	ø/ø:	e
206	3/23	-	-	-	100/	-	-	-	53/47	-	-	-	27/73	-	-	
238	3/17	-	-	-	100/	-	-	-	53/47	-	-	-	23/77	-	-	
275	/20	-	-	-	100/	-	-	-	53/47	-	-	-	10/87	/3	-	
317	-	-	-	-	93/	-	7/	-	47/27	-	3/23	-	17/30	/50	-	
356	-	-	-	-	30/	-	67/	-	7/	-	30/63	-	/3	3/93	-	
389	-	-	-	-	-	-	100/	-	-	-	40/60	-	-	20/80	-	
436	-	-	-	-	-	-	100/	-	-	-	53/47	-	-	30/70	-	
490	-	-	-	-	-	13	87/	-	-	13	57/30	-	-	40/47	13	
550	-	-	-	-	-	13	87/	-	-	20	57/23	-	-	30/47	23	

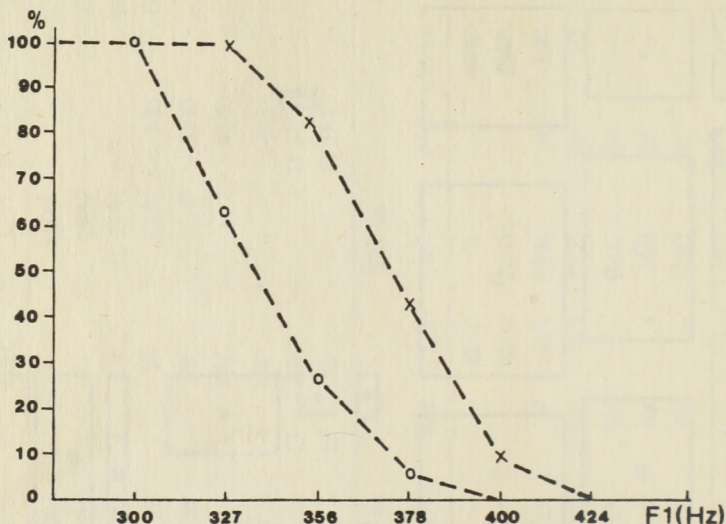
Magán- hangzó F1-e (Hz)	50 ms			100 ms			200 ms			300 ms		
	i/i:	e/e:	e/ø	i/i:	e/e:	e/ø	i/i:	e/e:	e/ø	i/i:	e/e:	e/ø
300	100/	-	-	100/	-	-	83/10	7/	-	40/60	-	-
327	100/	-	-	93/	7/	-	50/10	27/13	-	23/40	10/27	-
356	83/	10/7	-	53/	30/13	3/	20/3	47/30	-	7/20	27/47	-
378	43/	37/17	3/	17/	53/17	13/	3/3	60/23	10/	3/3	43/50	-
400	10/	57/17	17/	10/	53/13	23/	-	60/17	23/	-	47/33	20/
424	-	50/10	40/	-	50/10	40/	-	37/7	53/	-	37/20	43/
449	-	43/10	47/	-	30/7	63/	-	23/7	70/	-	23/10	67/
476	-	23/3	73/	-	13/3	77/7	-	10/3	80/7	-	-	97/3
504	-	17/	73/10	-	13/	70/17	-	7/	80/13	-	-	90/10



16. ábra

Az első formáns frekvenciájának és az időtartamnak
az összefüggése a magánhangzók felismerésében

nok az értékek a labiális palatális magánhangzóknál. A 17. ábrán az [i] és [i:] -ként azonosított magánhangzók percepciós arányait látjuk 50 és 300 ms időtartamú hangok esetében. (Megjegyezzük, hogy ez az "éles" elkülönülés a percepcióban csak akkor jelentkezik, ha fonemikus különbség áll a háttérben.)

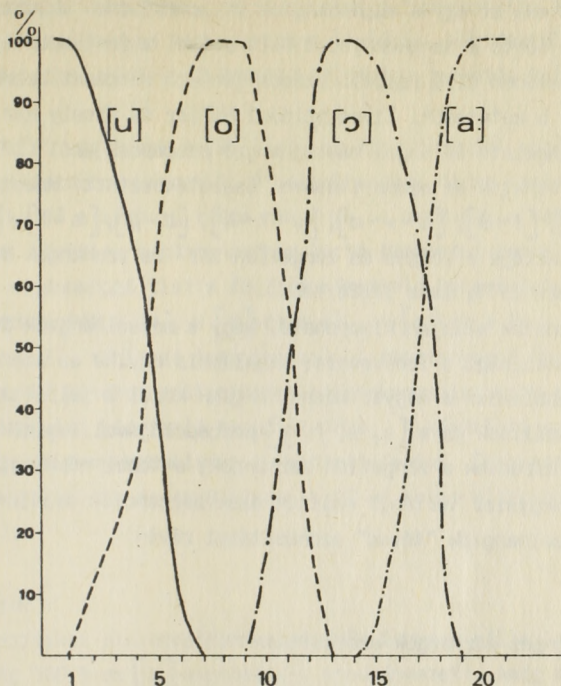


17. ábra

[i], illetve [i:] -ként azonosított 50 (x - - x) és 300 (o - - o) ms időtartamú magánhangzók különböző F1-gyel

Hasonlóképpen szemlélteti a 18. ábra grafikonja (100 ms-os magánhangzók esetén) az egyik fonémakategóriából a másikba történt átcsapásokat (az [u]-tól az [a:] -ig). A grafikon egyben a kategoriális magánhangzó-felismerés jellemzésére is alkalmas.

Igen magas F1 esetén keveredik az [e] és az [ø] a felismerésben. A labiálisokra jellemző formánsstruktúra mellett -- kis százalékban ugyan, de -- azonosítanak [e]-t 490 és 550 Hz-es első formánssal. Az illabiálisokra jellemző szerkezet esetében pedig hasonló, 476 és 504 Hz-es F1 mellett ítélnék [ø]-t. Igen alacsony első formáns pedig [i] és [e] percepciót eredményezhet



18. ábra

Hátul képzett magánhangzók felismerése %-ban (1--20-ig),
 az F1 és F2 változtatásával kapott hangminőségek
 (9. táblázat)

a labiálisokra jellemző szerkezet mellett rövid időtartam esetén.

A mássalhangzók percepció vizsgálat

A mássalhangzók felismerését is a frekvencia, az intenzitás és az idő paraméter, illetve ezek jellegzetes kombinációja határozza meg. A különféle kísérleti eredmények azt sugallják, hogy az akusztikai jel sajátágaiból

fakadóan más és más az egyes magánhangzók felismerésének akusztikai kulcsa (vagy kulcsai). Döntő a hangkörnyezet hatása (ld. a fonémadöntés különböző beszédhangok esetében című részt). Számos tényező szerepét azonban egyértelműen tisztázták a kutatások. Ilyen például Miller és Nicely már klasszikusnak számító vizsgálata 16 angol mássalhangzó percepciójáról (1955); ebből szemléletesen láthatók az azonosításban "összetévesztett" mássalhangzók csoportjai: [m~n], [f~θ], [b~v~d], [p~t~k], [d~g], [s~ʃ], [z~ʒ]. Továbbá a VOT funkciója a zöngés és zöngétlen zár- és réshangok felismerésében (Clark--Clark 1977; Lane 1968; stb.).

Mindezen adatok alapján kimondható, hogy a mássalhangzók felismerésében is döntő fontosságú a frekvencia; legalábbis egyike az azonosítási kulcsoknak. A továbbiakban a magyar mássalhangzók közül a felpattanó zöngés zárhangok, a nazálisok és a [s, z, ʃ, ʒ] percepciójával foglalkozom. A vizsgáldásokból kimaradtak a zöngétlen zárhangok, a többi réshang, a [l] és a [r] -- bár nemegyszer történik említés ezek percepcióis sajátosságairól is, a vizsgált mássalhangzók "téves" azonosításai révén.

A felpattanó zöngés zárhangok azonosítása

Számos tanulmány, amely a felpattanó zárhangok felismerésével foglalkozik, elsősorban azt igyekszik megválaszolni, melyik összetevő az, amelyiknek alapján a hallgató e mássalhangzók képzési helyét felismeri. Az eredmények szerint két tényező a meghatározó: a zárhang zörejsomójának (tehát a zárfelpattanásnak) és az átmenetnek a frekvenciaértéke (Delattre és munkatársai 1955; Fant 1969; Stevens--Blumstein 1978; Dorman--Raphael 1980; Blumstein--Stevens 1980; stb.).

Anyag és módszer

Lieberman és munkatársai kimutatták, hogy az F2 változtatásával a [b]-től a [g]-ig létrehozható a zárhangokra jellemző frekvenciaszerkezet (1963). Választ kerestem arra, hogy ugyanilyen változtatás miként hat a magyar explozívák felismerésére; valamint hogy a zárhangot követő magánhangzó minősége hogyan hat a mássalhangzó azonosítására. A kísérlet nyelvi anyagát a [b, d, g]-nek az összes magyar magánhangzóval alkotott szintetizált CV-kapcsolatai adták (a rövid--hosszú magánhangzópárok közül csak a hosszúak szerepeltek). Mi indokolta a hangkapcsolat nagyságrendet? Mindenekelőtt az az elmé-

leti és tapasztalati eredmény, amely szerint a fonémadöntés minimálisan egy magán- és egy mássalhangzó kapcsolatában történik. Másrészt pedig az a cél, hogy a felsőbb szintek működését, például a jelentés hatását kiküszöböljük.

A szintézist ugyancsak a PDP 11/34-es számítógéppel vezérelt OVE III típusú szintetizátorral végeztük. A mássalhangzót három szeletből állítottam elő: az első kettő a zárhang zöngé részének, a harmadik a zárfelpattanásnak felel meg. Ezt követi a magánhangzóhoz tartó átmenetet reprezentáló szelet, majd pedig a magánhangzó tiszta fázisát képviselő hangszeletek (a [b] akusztikai szerkezetét a 14., a [d]-ét a 15., a [g]-ét a 16., a magánhangzók hangszeleteit a 17. táblázat tartalmazza). A tesztanyagot úgy állítottam össze, hogy az F2 értékét változtattam a hangkapcsolatoknak a zárfelpattanást és az átmenetet képviselő hangszeletében. A kísérleti személyeknek (18 fő) az elhangzott hangkapcsolatot le kellett írniuk. Összesen 182 hangkapcsolatot szintetizáltam és teszteltem le (2--4 hetes eltérésekkel ugyanazon személyekkel).

Eredmények

A bá akusztikai szerkezetében a második formáns változtatása 1037 Hz-től 1270 Hz-ig 100 %-os [b]-azonosítást eredményezett. Amíg az F2 értéke nem érte el a magánhangzó második formánsának értékét, a felismerés helyes maradt. Amikor azonban megegyezett a magánhangzóéval vagy meghaladta azt, akkor a [b] felismerése csökkent: 1345 Hz-nél már 74 %-ra. Ebben a frekvenciaszerkezetben a [b] azonosítását már nem a frekvencia biztosítja mint elsődleges kulcs, hanem a bilabiális zárhangra jellemző intenzitásviszonyok. (Ennek bizonyítására a teljes időtartam harmadik harmadában a viszonylagos erősséget 10 dB-ről 2--4 dB-re csökkentettem, ami a dá intenzitásviszonyait közelíti meg. Az erre kapott adatok szerint 55 %-ra esett a [b]-azonosítás és 45 %-ban más zárhangokat ismertek fel a kísérleti alanyok.) Az F2-érték további csökkentésével arányosan csökken a [b]-érték is. Az a frekvenciasáv, amely a [d] azonosítását biztosítja, jóval szűkebb, alig 50 Hz-nyi: 1270 Hz-es F2 mellett 58 %-ra csökken a [d] azonosítása (42 %-ban [b]-t ítélnék). A [g] felismerésében nem mutatkozik elsődleges kulcsnak a második formáns. (A részletes percepció adatokat vö. Gósy 1982, 95--102.)

Az [ɔ] magánhangzós kapcsolatokban korrekt a [b] felismerése, ha az F2 értéke nem haladja meg az 1068 Hz-et, illetve ha a második formáns alulról csatlakozik a magánhangzó 1008 Hz-es F2-jéhez. A [d] felismerése az 1200--

A hang- szelet azono- sítója	Tényle- ges idő- adat /ms/	Viszony- lagos /dB/ erősség	F ₀ /Hz/	F ₁ /Hz/	F ₂ /Hz/	F ₃ /Hz/	
B1	90	1-3	116				
B2	90	3-2	116				
B	A:3	20	10	116	356-449	1200	2329-2614
	A3	16	10-14	116	400-519	1008	2614
	0:3	12	6-12	116	346	800	2397
	U:3	12	8	116	300	692	2397
	0'3	12	10	116	356-400	1425	2397
	Y:3	12	6-12	116	231	1600	2468
	E3	12	4-12	116	400	1600	2075
	E'3	15	10	116	346	1550	2770
	I:3	16	8-10	116	252	1695	2468-2691
	B	A:	39	10-12	116	449-734	1200-1307
A		40	14	116	519-599	1008	2614
0:		39	12-14	116	346-400	800	2397
U:		39	9-14	116	300	692	2397
0'		39	10-14	116	400	1425-1600	2397-2329
Y:		39	12-14	116	231	1600-1695	2468
E:		39	12-14	116	436-519	1647-1745	2075-2329
E'		39	10-12	116	346-400	1550-2016	2770
I:		52	10-14	116	252	1695-2198	2691-3020

A táblázatokban használt jelek APhI-megfelelői: A: = [a:], A = [ɔ:], 0: = [o/o:], U: = [u/u:], 0' = [ø/ø:], Y = [y/y:], E = [ɛ:], E' = [ɛ:], I = [i/i:].

15. táblázat

A hang- szelet azono- sítója	Tényle- ges idő- adat /ms/	Viszony- lagos /dB/ erősség	F ₀ /Hz/	F ₁ /Hz/	F ₂ /Hz/	F ₃ /Hz/	
D1	A:,U:,0	92	2	116			
	A:,O:,Y:, E', I:	92	3	116			
	E	46	4	116	231-218		
D2	A:,O'	24	2-1	116			
	A:,O:,Y:, E', I:	24	3-2	116			
	U:	24	2-0	116			
	E	78	4-2	116	206		
D	A:3	12	1-4	116	400-504	1345	2770
	A3	6	2	116	400	1307	2770
	O:3	21	0-6	116	400	1307	2397
	U:3	7	6	116	252	1307	2397
	O'3	12	1-3	116	356-400	1695	2691
	Y:3	6	2	116	231	1958	2329
	E3	15	0-4	116	378-436	1796	2329
	E'3	15	0-4	116	400	1796	2770
D	I:3	6	2	116	252	2016	3020
	A:	78	14-12	116	504-755	1345-1307	2614
	A	39	12-14	116	462-582	1307-1008	2770-2614
	O:	39	8-12	116	400	1307-800	2397
	U:	35	14	116	252-300	1307-692	2397
D	O'	35	14	116	400	1695-1600	2691-2468

Y:	39	6-12	116	231	1958-1695	2468
E	55	14	116	449-550	1796	2329
E'	39	14-12	116	400	1796-2016	2770
I:	12	6-12	116	252	2016-2198	3020

16. táblázat

	A hangszel- let azono- sítója	Tény- leges idő- adat /ms/	/dB/	F ₀ /Hz/ AK(dB)	F ₁ /K ₁ /Hz/	F ₂ /K ₂ /Hz/	F ₃ /Hz/
G1	A:	60	0-4	116	252-231		
	A, O:, U:, O', Y:, E, E', I:	52	0-4	116	252-231		
G2	A:	44	4-2	116	231-206		
	A, O:, O', Y:, E, E', I:	52	4-2	116	231-206		
	U:	60	4-2	116	231-206		
G	A:3	52	2-4/ 2-6	116/30	212-449/ 1385	1600/4525	2691
	A3	60	2-4/4	116/30	212-436/ 1099	1131/4525	2614
	O:3	52	2-6/4	116/30	252-300/ 1099	1068/4525	2397
	U:	52	2-6/ 0-6	116/30	206-252/ 1037	1099/4525	2397
	O'3	52	2-4/ 6-2	116/30	259-378/ 1037	1903/4525	2468
	Y:3	32	2-8/4	116/30	231/1099	1848/4525	2468
	E3	39	2-6/ 6-0	116/30	206-346/ 2075	1958/4525	2614
	E'3	12	2-6/ 6-0	116/30	252-300/ 1099	2198/4525	2770
	I:3	16	2-8/ 0-2	116/30	252/2934	2468/4525	6041
		A:	78	6-12	116	449-755	1600-1307
	A	75	6-12	116	462-582	1131-1008	2614
	O:	60	6-12	116	300-400	1068-800	2397

G	U:	90	6-14/ 6-0	116/30	300/1037	1099-692/ 4525	2397
	O'	90	4-10	116	400	1903-1600	2468-2329
	Y:	55	4-14	116	231	1849-1695	2468
	E	65	7-14	116	367-519	1958-1796	2614-2329
	E'	52	6-12	116	300-400	2198-3016	2770
	I:	55	0-14	116	252	2468-2198	6041

17. táblázat

A hangszel- let azono- sítója	Tényleges időadat (ms)	Viszonyla- gos erős- ség (dB)	F0	F1 (Hz)	F2	F3
A:1	208	12--10	116	755	1307	2614
A:2	156	10--0	116	755	1307	2614
A1	184	14	116	599	1008	2614
A2	144	14--0	116	599	1008	2614
O:1	224	14	116	400	800	2400
O:2	184	14--0	116	400	800	2400
U:1	104	14	116	300	692	2400
U:2	104	14	116	300	692	2400
U:3	104	14--0	116	300	692	2400
O'1	104	10--14	116	400	1600	2400
O'2	104	14	116	400	1600	2400
O'3	104	14--0	116	400	1600	2400
Y:1	224	14	116	231	1695	2468
Y:2	224	14--0	116	231	1695	2468
E1	88	14--12	116	550	1796	2329
E2	88	12--10	116	550	1796	2329
E3	88	10--0	116	550	1796	2329
E:1	224	12--14	116	400	2016	2770
E:2	184	14--0	116	400	2016	2770
I:1	224	14	116	252	2200	6000
I:2	224	14--0	116	252	2200	6000

1300 Hz-es, szűkebb tartományban helyes. A [g]-re jellemző akusztikai szerkezetben a [b]-nek, illetve [d]-nek megfelelő második formánsok ellenére 100 %-os a biztos azonosítás.

Az [o:]-magánhangzós hangkapcsolatokban [b]-t értenek a 600--900 Hz-es tartományban, de még 980 Hz-nél is 70 %-os volt a jó felismerés. A [d] és a [g] azonosítása az alsó nyelvállású magánhangzóval alkotott hangkapcsolatok percepciójához igen hasonló.

Az [u:] -magánhangzós kapcsolatokban korrekt a bilabiális mássalhangzó felismerése egészen a 800 Hz-es értékig. 872 Hz esetén már csak 53 %-ban azonosítanak [b]-t, 21 %-ban [d]-t és 26 %-ban [g]-t hallanak. A frekvenciaérték növelésével csökken a [b] és növekszik a [d] felismerése. 951 és 979 Hz-nél csaknem fele-fele arányban a dentális és a veláris zárhangot azonosították a kísérleti személyek. A dentális mássalhangzó felismerése viszonylag széles tartományban biztos: 1000--1300 Hz között.

Az [ø:] -s hangkapcsolatokban korrekt a [b]-értés 1500 Hz-ig, az 1600 Hz-es F2-nél 67 %-ra ugrik a [d]-azonosítás aránya. A [d]-re jellemző 1647 Hz-es értéknél a kísérleti alanyok 100 %-ban a dentális zárhangot azonosítják. Az F2 további emelésével az ítéletek a dentális és a veláris zárhang felismerése között ingadoznak: 1900 Hz-nél 42 %-ban [d]-t, 58 %-ban [g]-t azonosítanak. A [g]-felismerésre is hatással van az F2 mozgása, ellentétben a hátul képzett magánhangzókkal alkotott kapcsolattal. 1695 Hz-nél már csak 39 % a [g] és 61 % a [d]-ítéletek aránya.

Az [y:] -vel alkotott kapcsolatokban az F2=1600 Hz-es értéktől csökkentve a frekvenciát, egyértelműen jó a [b] felismerése. 1700 Hz-nél csaknem egyformán oszlik meg a kísérleti alanyok válasza három zárhang között. Mintegy 50 Hz-cel magasabb F2 esetén csaknem 80 %-os a [d] azonosítása. Az eredmények szerint az [y:] előtt akkor legbiztosabb a dentális zárhang felismerése, ha az F2 értéke 1960 Hz körül van. A második formáns frekvenciája kis mértékben befolyásolja a veláris zárhang percepcióját: az eredmények 60--95 %-os azonosítást mutatnak.

Az [e]-magánhangzós kapcsolatok F2-értékei a [b]-re jellemzőek, amíg az el nem éri az 1800 Hz-es értéket. Ekkor 42 %-ban [b]-t, 37 %-ban [d]-t és 21 %-ban [g]-t hallanak. Az F2 növelésével tovább csökken a [b] és növekszik a [d] és a [g] felismerése. 1850 Hz esetén megszűnik a bilabiális zárhang azonosítása és 61 %-os a [d]-értés. Ennek megfelelően az 1800 Hz-es értéknél alacsonyabb F2 mellett növekszik a [b] és csökken a [d] felismerése; az 1960 Hz-es frekvencia elérésével azonban ugrásszerűen megnövekszik a [g]-í-

tételek aránya. A 19. ábra a [b - d - g] azonosítás változását szemlélteti ([e] magánhangzóval alkotott kapcsolatban) az F2 függvényében.

Az [e:]-magánhangzós kapcsolatokban viszonylag szűk az a frekvenciasáv, amelyben egyértelmű a [b] azonosítása. 1650 Hz-nél már 100 %-ban [d]-t ismertek fel a kísérleti személyek. 1900 Hz-ig biztos a dentális zárhang azonosítása, ennél magasabb értékeken növekszik a [g]-felismerés aránya. A veláris zárhangot legeggyöntetűbben 1900 és 2200 Hz közötti F2 esetén azonosítják.

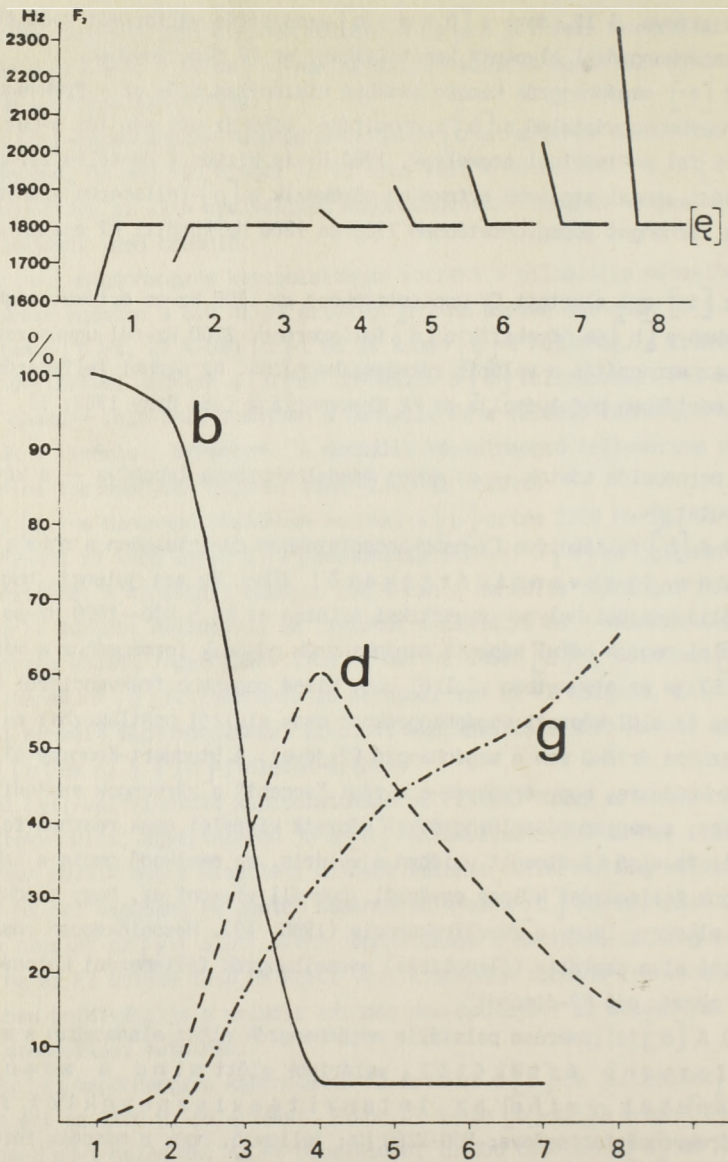
Az [i:] -vel alkotott CV-kapcsolatokban az 1700 Hz-es értéket meghaladva csökken a [b] és növekszik a [d] felismerése. 2200 Hz-nél ugrásszerűen átcsap az azonosítás a veláris zármássalhangzóba. Az utóbbi felismerését kisebb mértékben befolyásolja az F2 frekvenciája (vö. Gósy 1982: 13. táblázat).

A percepció adatok -- az egyes mássalhangzókra lebontva -- a következőket mutatják.

a) A [b] felismerése CV-hangkapcsolatokban egyértelműen a második formáns frekvenciaértékétől függ. Ez azt jelenti, hogy a bilabiális képzési helynek akusztikai szinten az F2 = 800--1800 Hz-es tartomány felel meg. A hátul képzett magánhangzók második formánsához a mássalhangzó F2-je az átmenetben alulról vagy kissé magasabb frekvenciáról csatlakozik. Az elől képzett magánhangzókhoz csak alulról csatlakozhat és nem lehet azonos értékű sem a magánhangzó F2-jével. A Studdert-Kennedy által feltett kérdésre, hogy érvényes-e a régi "recept" a zárhangok azonosítási kulcsaira, a magyar mássalhangzókkal végzett kísérlet csak részben felelhet igennel. Az első F2-átmenet valóban a veláris, az emelkedő pedig a labiális zárhangok felismerési kulcsa egyfelől, nem áll viszont az, hogy mindkét esetben alacsony lenne a zörejfrequencia (1980, 49). Hasonlóképpen részben fogadható el a dentális (alveoláris) mássalhangzók felismerési kulcsául a "magas zörej, eső F2-átmenet".

b) A [d] felismerése palatális magánhangzók előtt elsősorban a második formáns értékétől, velárisok előtt mind a második formánstól, mind az intenzitásviszonyoktól függ. Az F2 frekvenciatartománya: 900-2200 Hz; jellemző, hogy a második formáns felülről vagy alulról csatlakozik az átmenetben a magánhangzóhoz azzal a megkötéssel, hogy labiális magánhangzó esetében nem lehet azonos annak F2-jével, míg az illabiálisoknál meg is egyezhet azzal.

c) A [g] mássalhangzóra valamennyi hangkapcsolatban jellemző, hogy nem



19. ábra

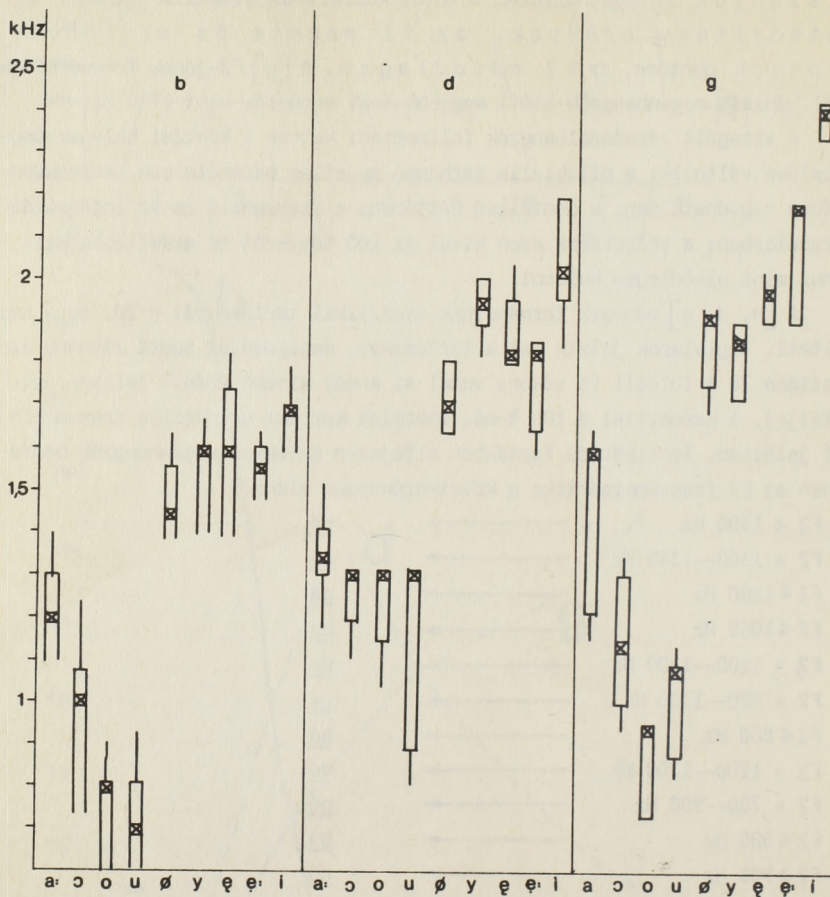
A be hangkapcsolat átmenetének F₂-értékei (1--8) és a meghangsításukra adott percepciós válaszok grafikonja

kizárólag az F2 értéke a döntő tényező a felismerésben. Az elől képzett magánhangzók előtt mind a második formáns, mind az intenzitásviszonyok, mind az F1 mozgása, sőt időviszonyok is meghatározók. A hátul képzett magánhangzók esetében az intenzitásviszonyok, az F1 menete és az időviszonyok döntőek, az F2 másodlagos. A [g] F2-jének frekvenciája elől képzett magánhangzók előtt nagyobb vagy egyenlő, mint 1740 Hz.

A vizsgált zármássalhangzók felismerési kulcsa a képzési helynek megfelelően változik: a bilabiális zárhangé egyetlen paraméterben, a frekvenciában ragadható meg; a dentálisé kettőben: a frekvencia és az intenzitás paraméterben; a veláriséba ezen kívül az idő tényezőt is számításba kell venni mint elsődleges kulcsot.

A [b, d, g] második formánsának akusztikai tartományát a 20. ábra szemlélteti. Téglalapok jelzik azt a tartományt, amelyben az adott zárhang azonosítása 70 % fölötti (a vékony vonal az annál kisebb arányú felismerést mutatja). A kereszttel a 100 %-os, abszolút korrekt azonosítás frekvenciáját jelöltem. Implikációs formákban kifejezve az egyes magánhangzók bontásában az F2 frekvenciaértéke a következőképpen alakul:

ha $F2 < 1300$ Hz	→	<u>bá</u>
ha $F2 = 1300--1345$ Hz	→	<u>dá</u>
ha $F2 \geq 1200$ Hz	→	<u>gá</u>
ha $F2 \leq 1068$ Hz	→	<u>ba</u>
ha $F2 = 1200--1300$ Hz	→	<u>da</u>
ha $F2 = 980--1300$ Hz	→	<u>ga</u>
ha $F2 \leq 800$ Hz	→	<u>bó</u>
ha $F2 = 1200--1300$ Hz	→	<u>dó</u>
ha $F2 = 700--900$ Hz	→	<u>gó</u>
ha $F2 \leq 800$ Hz	→	<u>bú</u>
ha $F2 \geq 900$ Hz	→	<u>dú</u>
ha $F2 = 850--1150$ Hz	→	<u>gú</u>
ha $F2 \leq 1550$ Hz	→	<u>bő</u>
ha $F2 = 1650--1800$ Hz	→	<u>dő</u>
ha $F2 \geq 1740$ Hz	→	<u>gő</u>
ha $F2 \leq 1600$ Hz	→	<u>bű</u>
ha $F2 = 1900--1960$ Hz	→	<u>dű</u>
ha $F2 = 1700--1900$ Hz	→	<u>gű</u>
ha $F2 < 1796$ Hz	→	<u>be</u>



20. ábra

A [b,d,g] mássalhangzók magánhangzós kapcsolataiban az átmenet második formánsának frekvenciatartománya (a 70 %-nál jobb felismerést eredményező értékeket a téglalapok, a gyengébbeket a vonalak jelzik)

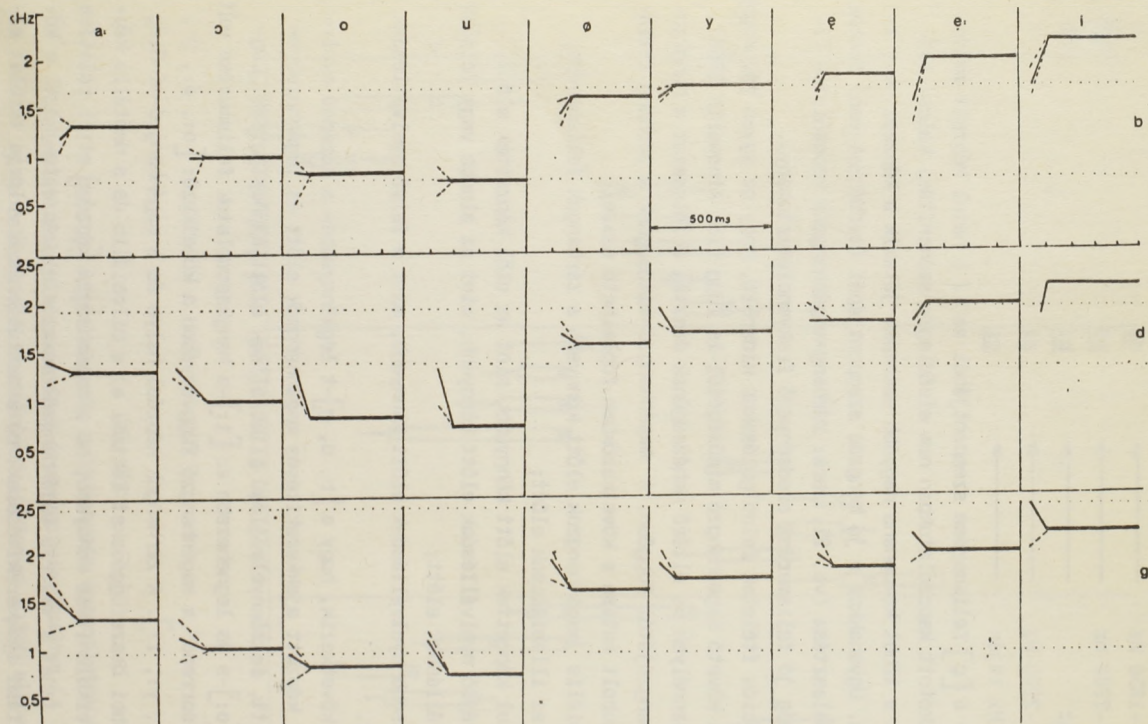
1a $1950 \text{ Hz} > F2 \geq 1796 \text{ Hz}$	—————>	<u>de</u>
1a $F2 \geq 1958 \text{ Hz}$	—————>	<u>ge</u>
1a $F2 \leq 1600 \text{ Hz}$	—————>	<u>bé</u>
1a $F2 = 1647\text{--}1850 \text{ Hz}$	—————>	<u>dé</u>
1a $F2 = 1900\text{--}2200 \text{ Hz}$	—————>	<u>gé</u>
1a $F2 \leq 1700 \text{ Hz}$	—————>	<u>bí</u>
1a $F2 = 1958\text{--}2200 \text{ Hz}$	—————>	<u>dí</u>
1a $F2 = 2450 \text{ Hz}$ táján	—————>	<u>gí</u>

Az, hogy a [g] felismerése szempontjából az F2 a hátul képzett magánhangzókkal alkotott kapcsolatokban nem elsődleges azonosítási kulcs, azt jelenti, hogy a többi komponens nagyobb szerepet játszik a képzési hely azonosításában. Ugyanakkor a jó hangzás szempontjából (természetesen) fontos az F2 jó megválasztása (vö. 21. ábra: zárhang+magánhangzós kapcsolatok F2-értéke és a még jó felismerést eredményező frekvenciatartomány).

A percepcióis ítéletek feldolgozásakor kiderült, hogy az egyes zárhangok felismerése a követő magánhangzó minőségétől is függ (vö. Ainsworth 1968). A kísérleti személyek is eltérő nehézségűnek érezték a feladatot a hangkapcsolat magánhangzójától függően. A magánhangzó-minőségnek a zárhang felismerésére gyakorolt hatása a következőkben foglalható össze:

- a) a labiális magánhangzók előtt könnyebb a zárhangok felismerése, mint az illabiálisok előtt;
- b) a hátul képzettek előtt könnyebb, mint az elöl képzettek előtt;
- c) a középső nyelvvállásúak előtt könnyebb, mint az alsóbb vagy felsőbb nyelvvállásúak előtt;
- d) az alsóbb nyelvvállásúak előtt könnyebb, mint a felső nyelvvállásúak előtt.

Mindebből az következik, hogy a [b, d, g]-t legkönnyebben a középső nyelvvállású, hátul képzett ajakkerekítéses magánhangzók előtt és legnehezebben az elöl képzett, felső nyelvvállású illabiálisok előtt azonosították. Legkönnyebb az [o:] -s és legnehezebb az [i:] -s hangkapcsolatok felismerése volt. A "nehézségi sorrend" a magánhangzó függvényében a következő: [o:, ø:, ɔ, a:, e, e:, u:, y:, i:]. A zárhangok képzési helye és a magánhangzó-minőség között az alábbi összefüggéseket láttam: a) a bilabiális és a dentális képzési terület elkülönítése könnyebb, ha a magánhangzó képzése ettől relatíve "távol" esik: tehát a veláris magánhangzók és az elöl képzettek közül a középső nyelvvállású ajakkerekítésesek; b) a dentális és a veláris terület elkülönítése könnyebb volt az ettől relatíve távoli összes középső és alsó



21. ábra

A zárhang + magánhangzós kapcsolatok átmenetének F2-értéke, valamint a még jó felismerést biztosító frekvenciatartomány

nyelvállású magánhangzó előtt. Ez feltehetően összefügg azzal is, hogy ezeknek a magánhangzóknek az azonosítása egyébként is "könnyebb".

A nazális mássalhangzók felismerése

A nazális beszédhangok akusztikai szerkezetét -- kis túlzással mondhatjuk -- a foniátriai (logopédiai) szakirodalom sokkal részletesebben tárgyalja, mint a fonetika. Az orrhangzós beszéd gyógyításához az első lépés a kórok és kórkép tisztázása, amelyhez -- mások mellett -- elengedhetetlen az ún. nazális beszéd akusztikai elemzése (Gutzman 1913; Luchsinger 1954; Fára--Sedláčková et al. 1970; Bauer 1972; Kittel 1979; stb.). Fant összegzésében a nazálist mint osztályt akusztikailag a következő három dolog jellemzi: a) egy nagyon alacsony, kb. 300 Hz-es első formáns és az ettől elkülönülő felsőbb formánsok (vö. Fujimura 1962; 1974 is); b) a formánsok intenzitásának relatíve nagy csillapítása; c) 800 és 2300 Hz között nagy energia, valamint antiformáns jelenléte (Fant 1973, 247). Az egyes kutatóknak az orrhagú mássalhangzókra kapott elemzési eredményei lényegesen nem mondanak ennek ellent, de egységes vélemény nem alakult ki (Tarnóczy 1968, 6; vö. még Trenscheł történeti áttekintését: 1977). Újabbban a hallási benyomást ítélik a legfontosabbnak (vö. Eijkman, idézi Trenscheł 1977, 64); eszerint a nazalitás kritériuma, hogy hallható legyen a nazális rezonancia. A kísérleti eredmények azt mutatják, hogy a [n] kivételével alacsony a folyamatos szövegben lévő nazális korrekt megkülönböztethetőségi aránya, s ez azt sugallja, hogy az átmenetek, mind a megelőző, mind a követő magánhangzóhoz, fontosak e mássalhangzók felismerésében (Flanagan 1965, 290--291). A percepciós kísérletek egyrészt a második formáns fontos szerepét igazolják (Lieberman--Delattre--Cooper--Gerstman 1954), másrészt azonban meg lehetőségen bizonytalan eredményeket mutatnak (House 1982). A magyar akusztikai-fonetikai szakirodalomban egyrészt a locusok értékeit találjuk; a [m]-re adott sáv: 854--974, illetve 862--960 Hz, a [n]-re: 1586--1740, illetve 1598--1776 Hz, a [ŋ]-re: 1500-2500, illetve 2000--2435 Hz (Magdics 1965, 77--81; illetve Molnár 1970, 37). Pontosabb adatokat közöl már korábban Tarnóczy; eszerint a [m]-re jellemző frekvenciaértékek: 330 és 800, a [n]-re: 260, 600 és 2300 Hz, illetve a [ŋ]-re: 260 és 2200 Hz (1948, 74--75). Bolla diagramjai elsősorban az intenzitásgörbében mutatnak eltérést a nazálisok között, semmint a frekvenciaszerkezetben (1982, 20., 24., 27. tábló).

Az első percepciós adat a nazális hangokra Vértés O. Andrásról származik, amely szerint a "hosszan hangoztatott [ŋ] és [n] középső szakaszát egyáltalán nem -- vagy csak alig -- tudjuk megkülönböztetni egymástól. Fontos, hogy a kísérleti személy ne hallja se a zár kezdetét, se a végét" (1982, 109). Megjegyzi még, hogy az egyes nazális hangok "sokkal nagyobb rokonságban" vannak egymással akusztikai síkon, mint a megfelelő orális hangok (108). Mindenesetre a rendelkezésre álló adatok szerint nehezen lehet az a kísérleti eredmény, hogy az angol [m, n] mássalhangzók felismerésében, illetőleg megkülönböztetésében -- hasonlóan a [b, d, g]-höz -- az F2-nek akusztikai felismerési kulcs szerepe van (Lieberman et al. 1954).

Két kísérletsorozatban vizsgáltam a frekvencia szerepét a nazális mássalhangzók azonosításában.

A frekvenciaszerkezet szerepe az azonosításban

A kísérlet célja, hogy meghatározzam az egyes orrhangú mássalhangzókra a felismerést biztosító frekvenciatartományokat.

Anyag és módszer

Férfi ejtésben CV és VC típusú, nazális mássalhangzót tartalmazó hangkapcsolatokat rögzítettünk magnetofonszalagra, majd 36 dB/oktáv meredekségű frekvenciaszűrővel megszürtük. Az alul áteresztő és a felül áteresztő szűrőknél a következő levágási frekvenciákat alkalmaztuk: 680, 820, 1000, 1200, 1500, 1800, 2700, 3300; illetve 470, 560, 680, 820, 1000, 1200, 1500, 1800, 2700 és 3300 Hz-et. A hanganyag összeállításának szempontjai a következők voltak: a) olyan magán- és mássalhangzót tartalmazó hangkapcsolat legyen, amely ebben a formában a magyarban jelentést nem hordoz, tehát nem szó; b) a hangkapcsolat megfordítható legyen, s az előző feltételnek ebben a formában is eleget tegyen; c) a valóságban létezzen a magyarban a létrehozott hangkapcsolat abszolút szóvégen és szó elején. Így a következő hangkapcsolatokat állítottam össze: mo, om, ém, mé, ná, án, nyö, öny, nyé, ény. Összesen 180 hangkapcsolatot hallgattattam le 30--35 fóvel. Feladatuk az azonosított hangkapcsolat leírása volt (egyszerre 30 szótagot teszteltem).

Eredmények

A percepciós ítéleteket táblázatos összefoglalásban a 18. és 19. táblázat tartalmazza. Az első oszlopban szereplő betűjelek arra utalnak, hogy a

nazális mássalhangzó ekkor még felismerhetetlen (a legtöbb egyöntetű ítéletet kapott beszédhang jelét közlöm). A táblázat számai mellett található jelek a 15 %-on felüli azonosan eltérő felismeréseket jelzik, a felismert beszédhang APhI-jelével.

A számadatokból implikációk formájában a következő összesítés adható a nazálisok felismerésére vonatkozólag (20. táblázat).

Összegezve az látható, hogy a [m] és a [ɲ] mássalhangzókat gyakran -- a fonetikai helyzettől függően -- [n]-ként, a [n]-t ugyanakkor likvida és tremuláns, azaz [l], illetve [r]-ként azonosítják. Eszerint nem egészen áll az a tétel, hogy a nazálisok "egymás között" tévesztődnek a felismerésben (House 1982). Az adatok alapján legbiztosabb a bilabiális nazális azonosítása CV és a palatális nazálisé VC helyzetben. Jellemző, hogy a [m]-t értették [n]-ként, de fordítva ez nem áll. A [n] azonosítása egyértelműen a [l], illetve a [r] fonémaosztályába csap át.

Az orrhangú mássalhangzók nagyjából ugyanazon a frekvenciatartományon belül azonosíthatók: mintegy 1500--3300 Hz között -- ez az érték azonban erősen függ a fonetikai helyzettől és a magánhangzó minőségétől is. Például a na felismerése még 1800 Hz feletti szűrés esetén is biztos, VC helyzetben azonban már 560 Hz fölött 65 %-osra csökken a helyes azonosítás. Az adatok alapján -- úgy tűnik -- kimondható, hogy a magyar nazális mássalhangzók legjellemzőbb frekvenciasávja a 800-1800 Hz, ami azt jelenti, hogy mindkétféle szűrés esetén ebben a tartományban még éppen felismerhetők.

A második formáns hatása a felismerésre

Ebben a kísérletben a megelőzőekben kapott eredményeket mintegy kiszélesítve, nagyobb anyag percepcióis elemzésével próbáltam választ kapni a nazálisok azonosítási kulcsára vonatkozólag. Célo m az F2 percepcióra gyakorolt hatásának a vizsgálata volt.

Anyag és módszer

Mindhárom mássalhangzót CV hangkapcsolatban valamennyi magánhangzóval mesterségesen előállítottuk. Néhány esetben csak a nazális kör segítségével lehetett a mássalhangzó-kapcsolatot létrehozni, ezeket a tesztelésből kihagytuk (nú, nú, ní). A [ɲ]-t az [a:] és [e] mellett generáltuk (a hangkapcsolatok nagyobb részében csak a formánsok beadásával -- nazális kör nélkül -- ezek nem állíthatók elő, vö. Olasz 1982). A nazális mássalhang-

Alul áteresztő

CV/VC	680	820	1000	1200	1500	1800	2200	2700	3300 (Hz)
mo	4	40	66	59 *	72	66	100	100	100
om	8	31	44 *	38 o	72	57 *	88	85	90
e:m	n	n	8	10 x	20 *	10 o	94	100	100
me:	4	24	9 o	24 ●	90	90	94	100	100
na:	4	27 *	29 x	72!	93	100	100	100	100 (%)
a:n	j	27	24 ●	41!!	30 o	33 x	30 *	100	100
∅	-	4	4	42	94	100	100	100	100
o∅	1	n	15!	50	81	60	75 o	100	100
∅e:	j/n	36	-	20	62 x	88	100	100	100
e:∅	-	-	n	6 xx	9 ●	12 ●	25 x	100	100

A tévesen azonosított beszédhangok és atányúk

*	25l	*	36n	*	24n	*	56n	*	24n	*	25r
o		o	50n	o	38n	o	37 l	o	76n	o	25j
x		x	30l	x	21n	x	16m	x	33l	x	62n
●		●	18n	●	52n	●	37n	●	51n		
!		!	21m	!	27l						

!! 44l /xx 40n

(%)

Felül áteresztő

CV/VC	470	560	680	820	1000	1200	1500	1800	2200	2700	3300	(Hz)
mo	100	90	100	100	83	82	59	45	15	20	20	
om	100	90	13*	29*	6*	5*	3*	31	3	-*	-*	
e:m	85	90	20 ^o	10 ^o	8 ^o	8 ^o	6 ^o	6*	6*	- ^o	- ^o	
me:	95	100	100	100	100	100	95	100	88	80	80	
na:	95	100	100	100	100	100	88	93	48	35	30	(%)
a:n	95	65	13x	37x	20x	10x	25	6	-	6	15	
n ^o	95	100	100	100	100	90	70	62	70	75	75	
óñ	100	100	70	76●	80	78	62	55	30 ^o	55	-x	
ñe:	100	100	97	100	67	87	63	62	57	70	70	
e:ñ	100	100	7●	3!	3●	3●	-x	- ^o	-x	3	-	

A tévesen azonosított beszédhangok és arányuk

*	*	*	*	*	*	*	*	*	*	*
30n	37n	18n	47n	26r	52n	52n	70n	25n		
o	o	o	o	o	o	o	o	o		
53n	77n	60n	42n	74n	34n	15j	40n	15n	(%)	
x	x	x	x	x		x		x		
47r	30r	44r	50r	46n		33n		50n		
●	●	●	●							
40n	17j	33n	50n							
	!									
	70n									

Eredeti nazális		Azonosított nazális
CV	m	m
VC	m	n
CV	n	l
VC	n	l ~ r
CV	ɲ	m
VC	ɲ	n

zók akusztikai szerkezetét a 21. táblázat foglalja össze (a magánhangzók ugyanazok, mint a felpattanó zárhangokkal végzett kísérletben). Összesen 84 hangkapcsolatot teszteltünk 20 kísérleti személlyel. Feladatuk az azonosított hangkapcsolat leírása volt.

21. táblázat

Nazális más-salhangzó szeletei	Idő (ms)	Intenzi-tás (dB)	F0	F1	B1	F2	F3
1	50	0-8	116	252	96	változ-	2614
2.	75	10	116	252	96	tatott	2614
3.	15	10	116	V-től függ	V-től függ		2614

Eredmények

A 22. táblázatban összefoglaltuk az orrhangú mássalhangzókra kapott felismerési arányokat az F2 értékének függvényében.

Az adatok jobban szórnak, mint azt más nyelvek nazálisainak vizsgálatával kapcsolatban a szakirodalom írja. Noha House is megjegyzi, hogy a svéd orrhangú zárhangoknál azonos felismerési kulcsot talált a bilabiális mássalhangzóra: F2 kb. 1100 Hz esetén, azonban míg [a:] és [i] környezetben megegyeztek, [u] környezetében eltértek az eredmények (i.m.). (Összesen a fenti három magánhangzó kapcsolatában vizsgálta a nazálisokat.)

22. táblázat

Nazális	a:	ɔ	o	u	ø	y	i	e:	e
m	800-- 1500	800-- 900	700-- 900	800-- 900	900-- 1000	800-- 1000	1000	1000-- 1100	800-- 1000
n	700, 1100	1100-- 1200	1100-- 1200	-	700	-	-	1700-- 1800	700, 1100
ɲ	600-- 800	-	-	-	-	-	-	-	700, 1300

Eredménye szerint -- az [u]-környezetet kivéve -- az alacsony F2-átmenet a [m], az 1500 Hz körüli a [n] és az ennél magasabb a veláris ([ŋ]) hang felismerését biztosítja (160). Tekintettel, hogy a kísérletben valamennyi magánhangzót szerepeltettem, még bonyolultabb eredményeket kaptam. Úgy tűnik ezek alapján, meglehetősen nehéz a kontextus nélküli akusztikai kulcsot meghatározni. Általában a 800-900 Hz-es F2-átmenet a [m] azonosításához vezet -- a magánhangzótól függően vannak kisebb-nagyobb eltérések. Amíg például az [u] mellett a [m] felismerése F2=800 Hz esetén 100 %-os, az [ø]-nél ugyanebben a helyzetben csak 50 %-os; vagy amíg az [i:] -nél 1000 Hz-es második formáns-átmenet mellett 100 %-osan felismerik a bilabiális nazálist, [ɔ] mellett ez csak 50 %-os. (A közölt esetekben és a percepciósan nem vizsgáltakban is 700, esetenként 1300 Hz-es F2-átmenet biztosítja -- egyéb paraméterekkel együtt - a [ɲ] felismerését.)

Vizsgáltam az első formáns szerepét is, amely több esetben hasonlóan mutatkozott az F2-ével. Eszerint, ha a mássalhangzó-kapcsolat akusztikai szerkezete egyébként változatlan és az F1 értékét növeljük, 400 Hz körüli frekvenciánál a felismerés már átcsap [l]-be, magas, 80--100 %-os aránnyal. Cooper és munkatársai ugyancsak [m] → [l] átcsapást tapasztaltak a felismerésben szintetizált (angol) hangkapcsolatoknál akkor, ha az F2-átmenet időtartama relatíve nagy volt (1952). Mindezek azt feltételezik, hogy a nazálisok nem határozhatók meg egyetlen elsődleges felismerési kulccsal, szimultán más paramétereknek is hasonló funkciójuk kell, hogy legyen. (Az intenzitás -- bizonyos frekvenciaszerkezet mellett -- az elsődleges kulcs szerepét mutatja, ugyanis ha az orrhangú mássalhangzó intenzitása jelentősebben csökken, akkor az azonosítás a megfelelő felpattanó zárhang: [b, d, ʃ])

irányába tolódik el.) Az intenzitás mint felismerési kulcs szerepének pontos megítélése a nazálisok és felpattanó zárhangok esetében elemzésre szorul.

Az [s, z, ʃ, ʒ] mássalhangzók felismerése

A dentális és alveoláris réshangok -- az artikuláció következtében -- meglehetősen széles frekvenciatartományban valósulnak meg (vö. Lazicius 1944; Tarnóczy 1954; Fónagy--Szendé 1969; Subosits 1978; Vértes O. 1982; Olasz 1985). Az egyes mássalhangzókra jellemző zörejek intenzitása azonban különböző, ezért a hang karakterének kialakításában a legintenzívebb komponensek vesznek elsősorban részt. Ezekkel "szokásos" a [s, z, ʃ, ʒ]-t jellemzeni. Láttuk azonban korábban, hogy vizsgálhatók a kevésbé intenzív összetevők is, amelyek jelentősége éppen a percepciósfunkciójukban van.

Ezeknek a réshangoknak a felismerését egyrészt a zörejjóc frekvenciája, másrészt az időtartam biztosítja (Gerstman 1956; Csisztovics 1965; Flanagan 1965; Cohen--t'Hart 1963; Ainsworth 1976; Vicsi 1981; Gósy 1984e; stb.). Fontos megállapítás, hogy a zöngétlen (szintetizált) mássalhangzókba betáplált F0-érték nem kelti a zöngés pár benyomását (Cohen--t'Hart 1963).

Fónagy és Szendé megállapítják, hogy a [s] és a [ʃ] közti különbségről az a hallási "benyomásunk", hogy az utóbbi "mélyebb, tompább" (1969, 287). Kísérleteink során az ezeknek a metaforáknak megfelelő fizikai értékeket próbáltuk meghatározni.

A felismerést biztosító frekvenciatartomány meghatározása

A kísérlet célja, hogy meghatározzuk milyen és mekkora frekvenciatartomány szükséges a [s, z, ʃ, ʒ] mássalhangzók biztos azonosításához.

Anyag és módszer

Az első kísérletsorozatban természetes ejtésű hangkapcsolatokban (CV és VC típusú) elemeztük a zörejfrequenciának mint felismerési kulcsnak a szerepét az egyes réshangok esetében. A hangkapcsolatok összeállításának szempontjai megegyeztek a nazálisokkal végzett ugyanilyen módszerű kísérletben leírtakkal. A hangkapcsolatokat ugyancsak alul és felül áteresztő szűrővel szűrtük, majd 31 fővel lehallgattattuk. Összesen 320 hangkapcsolat

latot teszteltünk (egyszerre átlagosan 35-t). A nyelvi anyag a következő volt: [ɔʃ, ʃɔ, is, si, ɔs, sɔ, oz, zo, uʃ, ʃu, uz, zu, iʒ, ʒi, a:ʒ, ʒa:].

Eredmények

Az észlelési eredmények összefoglalását a 23. és 24. táblázat tartalmazza. Valamennyi réshangra jellemző, hogy 1000 Hz alatti összetevőik alapján gyakorlatilag nem felismerhetők, ekkor zárhangokként, labiodentális réshangokként, [h]-ként, illetve likvidaként azonosítják őket (attól is függően, hogy az eredeti mássalhangzó zöngés volt-e vagy sem). A felismerhetőség hármas függésű: a) függ egyrészt a réshang típusától, b) a magánhangzós környezettől (nyilván általában is a környezettől) és c) a fonetikai helyzettől. Az adatokból abszolút szabályok nem vonhatók le, mivel nem szerepel valamennyi beszédhang, mint környezet, az azonban a meglévőkből is jól látszik, hogy a zöngétlen réshangok felismerése a lacsonyabb frekvenciaérték esetén "kezdődik" és rendszerint nagyobb frekvenciatartományban korrekt, mint a megfelelő zöngéseké. A VC helyzetű réshangok azonosítása kevésbé jó, mint a CV-ké. (Már a csecsemők beszédhang-differenciáló képességét is befolyásolja a fonetikai helyzet: CV helyzetben mindig jobb az azonosítás, vö. Eilers 1977; Williams--Bush 1978.) Bizonyos torzítás mellett jellemző a réshangok egymás közti helyettesítődése: a zöngéseket a megfelelő zöngétlenek "váltják föl". Előfordul az is, hogy éppen megfordítva, a zöngétlent ismerik fel zöngésnek (vö. [s] → [z]); ez feltehetően az "előre jóslás" tévedésével kapcsolatos (vö. Csisztovics 1965). Megpróbáltuk -- a rendelkezésre álló adatokból -- meghatározni az egyes réshangok felismerhetőségének frekvenciatartományát (a 70 %-on felüli azonosítások figyelembevételével; az első érték az alul; a második a felül áteresztő szűrés határértéke):

alveoláris zöngétlen réshangok:	1500--3300 Hz;
alveoláris zöngés réshangok:	1800--800 Hz;
dentális zöngétlen réshangok:	2500--3300 Hz
	(kiv.: [si:]);
dentális zöngés réshangok:	1800--3300 Hz.

Az alul és felül áteresztő szűrővel kapott eredményeket grafikon formájában ábráztuk (a középértékekkel); ez jól szemlélteti a magyar résmássalhangzók egyik elsődleges akusztikai kulcsának, a frekvenciatartománynak a kiterjesztését (22. ábra).

23. táblázat

Alul átérésztő

CV/VC (Hz)	680	820	1000	1200	1500	1800	2200	2700	3300
ʃɔ	f	5	6	30	78	100	100	100	100
ɔʃ	t	t/k	t	27	100	100	100	100	100
ʃu	f	h	78	100	100	100	100	100	100
uʃ	t	h	53	89	97	100	100	100	100
ʒa:	z	v	3	38	94	94	100	100	100
a:ʒ	l	r/l	6	27	56	82	87	100	100
ʒi	r	r	3	24	62	94	100	100	100
iʒ	-	r	t	10*	22*	27*	30*	70	100 (%)
ʒɔ	f	f	f/h	3	3	12	56	80	100
ɔʒ	-	-	t	3	3	3	62 ₀	100	100
si	h	h	6	79	94	100	100	100	100
is	t	t	t	t	t	15	25	80	100
zo	z	v	v	90	66	88	100	100	100
oz	-	l	l	l/r	6	24	100	100	100
zu	-	r	9	65	72	66	97	100	100
uz	-	l	t	t	t	3	31	80	100

Tévesen azo-
nosított be-
szédhangok
és arányuk

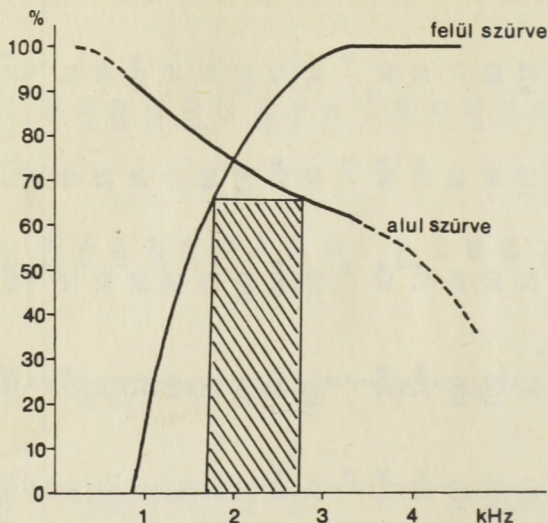
*
80f*
72f*
70f*
70f0
37 z

(%)

Felül áteresztő

CV/VC (Hz)	470	560	680	820	1000	1200	1500	1800	2200	2700	3300
ʃ	100	100	100	100	100	100	90	90	55	80	80
ʃf	100	100	100	100	100	90	76	52	27	80	80
ʃu	100	100	100	100	100	97	90	100	57*	100	100
uf	100	100	100	100	100	100	90	80	85	80	80
ʒa:	100	100	100	100	100	100	94	83	73	80	75
a:ʒ	100	100	93	86*	40*	20*	17*	30*	55	50	60
ʒi	100	100	100	100	64 _o	85 _o	80 _o	80 _o	90	95	90 (%)
iʒ	95	100	17*	20 _o	5x	-	-x	-x	- _o	-*	-*
sʒ	100	100	100	96	96	100	100	90	94	90	90
ʒs	100	100	43 _o	90	44 _o	64x	50 _o	52 _o	42	75	75
si	100	100	100	100	100	100	100	100	100	100	100
is	100	100	93	90	74	60	62	60	70	75	64
zo	100	100	100	100	80	80	80	86	76	55	50
oz	100	100	100	100	97	97	90	76	64x	70	70
zu	100	100	98	100	100	90	98	76	70	70	70
uz	100	100	73x	90	70	66 _o	70	44	67	80	80
Tévesen azo- nosított be- széhangok és arányuk			* 83f _o 57z x 27z	* 15f _o 80f	* 60f _o 20f x 95f • 56z	* 80f _o 100f x 30z • 24s	* 83f _o 15f x 95f • 30z	* 45f _o 15f x 93f • 18z	* 20z _o 85f x 15s	* 60f	* 70f

Az ábra vonalkázott területe a jó (65 %-os) felismerési arány frekvenciatartományát jelzi. Leolvasható, hogy mintegy az 1700--2800 Hz-es frekvenciaterületen a mássalhangzókra jellemző összetevők elegendőek a helyes felismeréshez.



22. ábra

A $[s, z, f, \mathfrak{z}]$ mássalhangzók felismerési tartománya frekvenciaszűréssel torzított hanganyag alapján

A zörejfrekvencia változásának hatása a felismerésre

A második kísérletsorozat célja az volt, hogy meghatározzuk azokat a frekvenciasávokat, amelyek egyfelől biztosítják a $[s, z]$, illetve $[f, \mathfrak{z}]$ biztos azonosítását, valamint egymás közötti megkülönböztetésüket.

Anyag és módszer

A tárgyalt résmássalhangzókat CV hangkapcsolatokban szintetizáltuk (a

korábban ismertetett eszközökkel és módszerrel). Magánhangzóként valamennyi magyar magánhangzóminőség szerepelt. A formánsszintetizátor két zörejjóc (K1 és K2) változtatását teszi lehetővé -- ennek megfelelően toltuk el az intenzív zörejjócokat a frekvenciatengely mentén valamennyi hangkapcsolat esetében. A kiindulást jelentő akusztikai szerkezeteket a 25. táblázat tartalmazza. Összesen 260 hangkapcsolatot szintetizáltunk. Ezt a hanganyagot (random sorrendben) 20 kísérleti személlyel hallgattattuk le. Feladatuk a felismert hangkapcsolat leírása volt.

Eredmények

Az első kísérletsorozat eredményei -- a szintetizálási tapasztalatlan-ságból adódóan -- az intenzitás mint felismerési kulcs szerepére mutattak rá (bár a kísérleti cél a frekvenciaváltozás percepciós vizsgálata volt). A zörejjócok intenzitását ekkor ugyanis a szükségesnél alacsonyabb értékre választottuk, mintegy 10--15 dB-lel. A kiindulást jelentő frekvenciaszerkezettel, ahol a két zörejjóc 5000 és 8000 Hz-en helyezkedett el, a mássalhangzók 100 %-osan felismerhetők voltak. Nem derült tehát azonnal ki, hogy az intenzitásérték megválasztása helytelen. Az eltolt zörejjócok tesztelési adataiból kiderült, hogy alacsony intenzitásérték mellett csak egy nagyon szűk frekvenciasávban (4-5000 és 8-9000 Hz) biztos a [s] hangok felismerése. A frekvenciaérték változtatásával gyakorlatilag megszűnt a réshangok felismerése, helyettük a kísérleti alanyok -- a réshang típusától függően -- [v, f, j, l, r] mássalhangzókat azonosítottak. Az intenzitás tehát mint másodlagos felismerési kulcs -- az elsődleges, a frekvencia módosulásának arányában -- jut szerephez. Ugyancsak szintetizált réshangokkal végzett kísérletek eredményei alapján jut arra a következtetésre Mártony, hogy a réshangok felismerését biztosító tényezők között harmadikként megemlíti a rés relatív intenzitását is (1963).

A percepciós adatok alapján implikációk formájában fejeztük ki az egyes réshangok felismerési kritériumait frekvenciában:

	[s]-t azonosítanak,
ha ([a:]) mellett	K1 > 3500 és K2 > 4000 Hz K1 > 2500 és K2 > 4500 Hz K1 és K2 ≥ 4000 Hz
ha ([j]) mellett	K1 és K2 ≥ 4000 Hz K1 > 3500 és K2 ≥ 4000 Hz
ha ([o]) mellett	K1 és K2 ≥ 4000 Hz

		ALVEOLÁRIS												
		MSH	IDŐ	INT.	F0	F1	B1	F2	B2	F3	AC	AK	K1	K2
Zörgés	1. szelet	40	0-3	103	252	96	1700	62-93	2614	0-14	16	1510	2540	
	2. szelet	80	4	103-97	252-346	96-60	1700-1500	93	2614	20-24	16	1510	2540	
Zörgétlen	1. szelet	120								6-22	16	1510	2540	
	2. szelet	50								28	16	1510	2540	
	3. szelet	40								24	16	1510	2540	
		DENTÁLIS												
		MSH	(ms)	(dB)	(Hz)	(Hz)	(Hz)	(Hz)	(Hz)	(Hz)	(dB)	(dB)	(Hz)	(Hz)
Zörgés	1. szelet	60	0-4	116	212	96	1600	93	2614	0-22			4032	9590
	2. szelet	40	4	116	212	96-36	1600	93-31	2614	24-20			4032	9590
Zörgétlen	1. szelet	120								0-24	22-6	6041	6041	
	2. szelet	60								22		6041	6041	
	3. szelet	60								24-16		6041	6041	
	átmenet 1.	4	0-6	116	*	60	*	0-93	2614					
	átmenet 2.	60	8	116-112	*	60	*	93-155	2614					

* = V-től függő

	$K1 \geq 3500$ és $K2 \geq 5000$ Hz
ha ([u]) mellett	$K1$ és $K2 \geq 4000$ Hz
	$K1 \geq 3500$ és $K2 > 4000$ Hz
Ha ([ø]) mellett	$K1$ és $K2 \geq 4000$ Hz
	$K1 > 3500$ és $K2 > 4000$ Hz
ha ([y]) mellett	$K1$ és $K2 \geq 4500$ Hz
	$K1 > 3500$ és $K2 > 4000$ Hz
ha ([i]) mellett	$K1$ és $K2 > 4000$ (5000) Hz
ha ([e:]) mellett	$K1$ és $K2 \geq 4500$ Hz
	$K1 \geq 3500$ és $K2 > 5000$ Hz
ha ([e]) mellett	$K1$ és $K2 > 5000$ Hz
	$K1 > 3500$ és $K2 > 5000$ Hz
[ʃ] -t azonosítanak,	
ha ([a:]) mellett	$K1$ és $K2 = 1000--3500$ Hz
	$K1 \leq 2500$ és $K2 > 4500$ Hz
ha ([ɔ]) mellett	$K1$ és $K2 = 1500--3500$ Hz
	$3000 > K1 > 1500$ és $K2 > 4500$ Hz
ha ([o]) mellett	$K1$ és $K2 = 1500--3500$ Hz
	$1500 \leq K1 < 3000$ és $K2 > 5000$ Hz
	$3500 < K1 \geq 1500$ és $K2 < 4000$ Hz
ha ([u]) mellett	$K1$ és $k2 = 1500--3500$ Hz
	$K1 = 1500--2500$ és $K2 > 4000$ Hz
ha ([ø]) mellett	$K1$ és $K2 = 1500--3500$ Hz
	$1500 \leq K1 \leq 2500$ és $K2 > 4500$ Hz
ha ([y]) mellett	$K1$ és $K2 = 2000--3500$ Hz
	$K1 = 2000--3500$ és $K2 > 4500$ Hz
ha ([i]) mellett	$K1$ és $K2 = 2000-3500$ Hz
	$3500 > K1 \geq 2000$ és $K2 > 4500$ Hz
ha ([e:]) mellett	$K1$ és $K2 = 2000--3500$ Hz
	$3500 > K1 > 2000$ és $K2 < 7000$ Hz > 4500
ha ([e]) mellett	$K1$ és $K2 = 2000--3500$ Hz
	$K1 = 2000--3000$ és $K2 > 4000$

[z]-t azonosítanak,

ha ([a:]) mellett	$K1$ és $K2 \geq 4000$ Hz
	$K1 \leq 1500$ és $K2 > 5000$ Hz
	$K1 > 3500$ és $K2 > 5000$ Hz
ha ([ɔ]) mellett	$K1$ és $K2 \geq 4000$ Hz

	$K1 \geq 3500$ és $K2 \geq 4000$ Hz
ha ([o]) mellett	$K1$ és $K2 \geq 4000$ Hz
	$K1 \geq 3500$ és $K2 \geq 5000$ Hz
ha ([u]) mellett	$K1$ és $K2 \geq 4000$ Hz
	$K1 \geq 3500$ és $K2 \geq 4000$ Hz
ha ([ø]) mellett	$K1$ és $K2 \geq 3500$ Hz
	$K1 > 3000$ és $K2 > 4000$ Hz
ha ([y]) mellett	$K1$ és $K2 \geq 3500$ Hz
	$K1 > 3000$ és $K2 > 4500$ Hz
ha ([i]) mellett	$K1$ és $K2 > 4500$ Hz
	$K1 > 3500$ és $K2 > 4500$ Hz
ha ([e:]) mellett	$K1$ és $K2 \geq 4000$ Hz
	$K1 > 3500$ és $K2 > 6000$ Hz
ha ([e]) mellett	$K1$ és $K2 > 4000$ Hz
	$K1 > 3500$ és $K2 > 4500$ Hz
	[ʒ]-t azonosítanak,
ha ([a:]) mellett	$K1$ és $K2 = 1000\text{--}3500$ Hz
	$K1 \leq 2500$ és $K2 > 4500$ Hz
ha ([ɔ]) mellett	$K1$ és $K2 = 1500\text{--}3500$ Hz
	$3000 > K1 > 1500$ és $K2 > 4500$ Hz
ha ([o]) mellett	$K1$ és $K2 = 1500\text{--}3500$ Hz
	$K1 < 3000$ és $K2 < 5000$ Hz
ha ([u]) mellett	$K1$ és $K2 = 1500\text{--}3500$ Hz
	$K1 = 1500\text{--}3000$ és $K2 \leq 3500$ Hz
ha ([ø]) mellett	$K1$ és $K2 = 1500\text{--}3000$ Hz
	$K1 = 1500\text{--}2000$ és $K2 \geq 4000$ Hz, akkor zöngétlenné lesz!
ha ([y]) mellett	$K1$ és $K2 = 1500\text{--}3500$ Hz
	$K1 = 2000\text{--}3500$ és $K2 > 4500$ Hz
ha ([i]) mellett	$K1$ és $K2 = 2000\text{--}4000$ Hz
	$K1 \leq 3500$ és $K2 > 4500$ Hz
ha ([e:]) mellett	$K1$ és $K2 = 2000\text{--}3500$ Hz
	$K1 = 2000\text{--}3500$ és $K2 \geq 4000$ Hz
ha ([e]) mellett	$K1$ és $K2 = 2000\text{--}3500$ Hz
	$K1 = 2000\text{--}3500$ és $K2 \geq 4000$ Hz

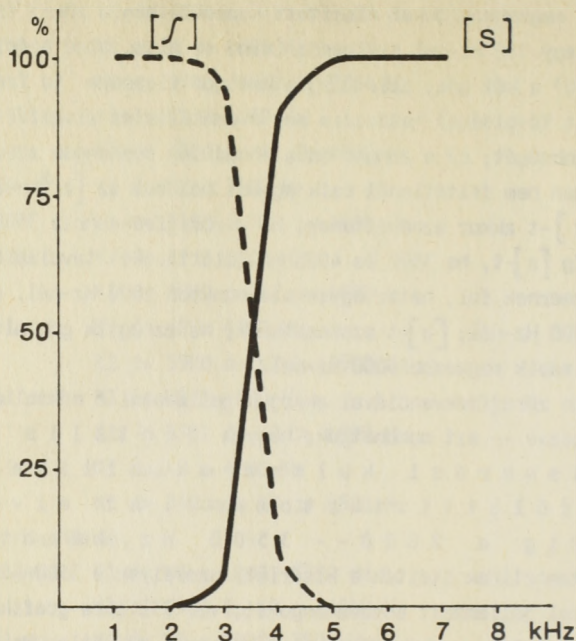
Az implikációk a következő tendenciákat mutatják: a) a következő magánhangzó akusztikai szerkezete hatással van a réshang zörejfrequenciájára.

A hátul képzett magánhangzókkal alakított kapcsolatban a zörej frekvencia-tartománya mintegy 500 Hz-cel szélesebb (mind az alsó, mind a felső érték felé tolódva); b) a két góc, illetőleg a két góc viszonya (a frekvenciategyen elfoglalt távolsága) határozza meg a dentális és alveoláris réshangok percepcióis különbségét; c) a zöngés és a zöngétlen réshangok zörejstruktúrája a percepcióban nem feltétlenül esik egybe: például az [a:] -val alkotott kapcsolatban [z] -t akkor azonosítanak, ha a zörejfrekvencia 3500 és 5000 Hz fölötti, míg [s] -t, ha 3500 és 4000 Hz fölötti. Két távolabbi zörejgóc esetén [s] -t ismernek fel, ha az egyik alacsonyabb 2500 Hz-nél, a másik pedig magasabb 4500 Hz-nél; [z] -t azonosítanak, ha az egyik góc alacsonyabb 1500 Hz-nél, a másik magasabb 5000 Hz-nél.

Ha egyetlen zörejfrekvenciával akarjuk jellemezni e mássalhangzókat, akkor -- összegezve -- azt mondhatjuk, hogy a dentális réshangok felismerési kulcsa akusztikailag a 4500 Hz fölötti tartomány; az alveolárisoké pedig a 2000 -- 3500 Hz közötti zörejsáv. Átmenetinek ítélték a kísérleti személyek a 3500--4500 Hz közötti zörejgóccal jellemzett beszédhangokat, vö. 23. ábra grafikonját. Hasonló módszerű kísérletek eredményei szerint az angol zöngétlen dentális és alveoláris réshang akusztikai határa hasonlóan 3000--4000 Hz között húzódik (Flanagan 1965, 295--296; vö. még Harris 1958). A percepcióis adatok alátámasztják a klasszikus fonetikának azt a kritikáját, hogy a zöngés és a zöngétlen réshangok csupán a zöngé meglétében, illetőleg hiányában különböznek egymástól (vö. Fónagy--Szendé 1969 is; valamint a kísérleti--fonetikai elemzések alapján: Bolla 1982).

A két zörejgóc egymástól való távolsága intenzitással kapcsolatos kérdéseket is felvet (erre azonban itt nem térünk ki). A réshangok viszonylag nagy frekvenciatartományban történt jó azonosítása nyilvánvalóan azzal is összefügg, hogy a magyarban -- nem lévén interdentális réshang -- a percepcióis mechanizmus még toleránsabb e mássalhangzókkal.

A két zörejgóc relatíve nagy távolsága esetén (1000--1500 és 9000--10000 Hz) a réshangokat [f]-ként, [d]-ként, [p]-ként, ritkábban [tʃ], illetve [tʃ] ként azonosították. (A zöngéseknél ritkábban ismernek fel más képzésű hangot, inkább felismerhetetlennek minősítik. A zöngétleneknél ez nem fordul elő.)



23. ábra

A [ʃ] és a [s] réshangok felismerése a zörejjóc
frekvenciájának függvényében

A frekvencia- és időtartamváltozás összefüggése
a [s, z, ʃ, ʒ] mássalhangzók azonosításában

A vizsgált réshangok felismerésében -- a zörejfrequenciához hasonlóan -- elsődleges akusztikai kulcs szerepet játszik az időtartam is (Gerstman 1956; Ainsworth 1976; Vicsi 1981). Az alveoláris és a dentális réshangok időtartamának csökkentése -- a többi akusztikai paraméter állandó értéke mellett -- más képzésmódú mássalhangzók felismerését eredményezi: affrikátákét, zárhangokét. Jelen kísérlet célja volt, hogy meghatározzuk azokat a konkrét időtartamértékeket, amelyek mellett a vizsgált szekvenciákban a mássalhangzók felismerése az adott típusban (pl. réshangok) korrekt.

Anyag és módszer

CV típusú hangkapcsolatokat szintetizáltunk, ahol a V-t három magán-

hangzóval: [a:, e:, ø:] -vel valósítottuk meg, a C a [s] mássalhangzó volt. Majd ezekhez a hangkapcsolatokhoz a [l] és a [r] mássalhangzók hozzákapcsolásával értelmes szavakat hoztunk létre: szár, szél, szőr. A kísérleti anyaghoz a [s] mássalhangzó eredeti, 200 ms-os időtartamát lépésekben -- 15--20 ms, illetve a hangminőség-változás határain 5--10 ms-onként -- rövidítettem. Így 16--16 hangkapcsolatot, illetőleg szót kaptam (a szavak a hangminőség megváltozása esetén is, egyetlen kivétellel, vö. cőr, értelmes szavakat eredményeztek). A random sorrendű hanganyagot, külön a hangkapcsolatokat és külön a szavakat 20 fővel hallgattattam le. Feladatuk a hallott hangjelenség leírása volt.

Eredmények

A [s] fizikai időtartamadatait és az azonosításokat a 26. táblázat tartalmazza. Mind a hangkapcsolatok, mind a szavak felismerésében bekövetkezett, hogy az időtartam bizonyos mértékű csökkenését elérve a [s] előbb [ts̃]-be, majd [t]-be csapott át a percepcióban. Az értelmes szavaknál a [s] → [ts̃] változás általában később következett be, mint a hangkapcsolatoknál, a [ts̃] → [t] pedig előbb: a jelentésnek a felismerés "rugalmasságára" gyakorolt hatásának megfelelően: például szél/cél, szőr/cőr vagy cár/tár, cél/tél. Az adatok szerint: a [s] eredeti időtartamát 104 ms-ra csökkentve még 70 %-ban szél-t hallanak, s csak 30 %-ban cél-t. Ugyanennél az időtartamnál szé-t csak 20 %-ban, cé-t már 80 %-ban azonosítanak. Az egyes mássalhangzók akusztikailag behatárolható azonosíthatósága tehát egyrészt az emlékezetben őrzött adatoknak felel meg: az anyanyelvi percepció bázis kialakulása során szerzett tapasztalatok alapján (alsóbb szintek részben automatikus működése).

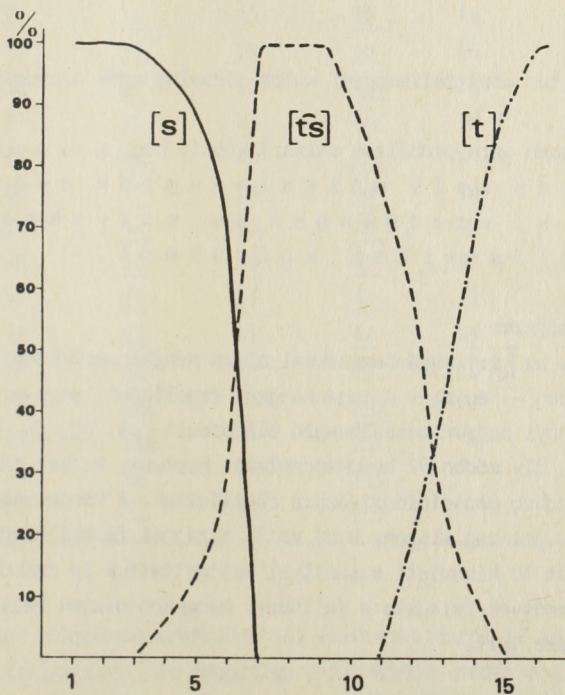
26. táblázat

Sor- szám	Hangsor eleji mássalhangzó időtartama (ms)	A V típusa	Hangkapcsolatban az azonosított			Szóban mássalhangzó		
			s	ts̃	t	s	ts̃	t
1.	200	a:	100	-	-	100	-	-
		ø:	100	-	-	100	-	-
		e:	80	20	-	100	-	-
2.	170	a:	100	-	-	100	-	-
		ø:	100	-	-	100	-	-
		e:	70	30	-	84	16	-

3.	143	a:	100	-	-	84	16	-
		ó:	90	10	-	67	33	-
		e:	67	33	-	84	16	-
4.	130	á:	90	10	-	95	5	-
		ó:	90	10	-	56	44	-
		e:	58	42	-	79	21	-
5.	117	a:	90	10	-	80	20	-
		ó:	70	30	-	40	60	-
		e:	52	48	-	75	25	-
6.	104	a:	70	30	-	50	50	-
		ó:	44	56	-	22	78	-
		e:	20	80	-	75	25	-
7.	94	a:	-	100	-	40	60	-
		ó:	33	67	-	20	80	-
		e:	-	100	-	28	72	-
8.	84	a:	-	100	-	20	80	-
		ó:	-	100	-	-	100	-
		e:	-	100	-	10	90	-
9.	75	a:	-	100	-	-	100	-
		ó:	-	100	-	-	100	-
		e:	-	100	-	-	100	-
10.	65	a:	-	100	-	-	100	-
		ó:	-	100	-	-	100	-
		e:	-	100	-	-	100	-
11.	52	a:	-	90	10	-	100	-
		ó:	-	100	-	-	100	-
		e:	-	100	-	-	100	-
12.	42	a:	-	80	20	-	72	28
		ó:	-	80	20	-	10	90
		e:	-	74	26	-	100	-
13.	32	a:	-	56	44	-	56	44
		ó:	-	47	53	-	10	90
		e:	-	63	37	-	44	56
14.	20	a:	-	20	80	-	16	84
		ó:	-	10	90	-	-	100
		e:	-	5	95	-	10	90
15.	15	a:	-	15	85	-	10	90

		ø:	-	-	100	-	-	100
		e:	-	5	95	-	10	90
16.	12	a:	-	10	90	-	-	100
		ø:	-	-	100	-	-	100
		e:	-	-	90	-	-	80

Másrészt azonban ezeket a határokat tágíthatják-szűkíthetik a szemantikai tényezők, amelyek "felülbírálják" az alsóbb szinteken hozott döntéseket. A 24. ábra grafikonja a fonémaátcsapást mutatja a hangkapcsolatok percepciójában.



24. ábra

A [s] mássalhangzó időtartam-csökkentésének hatása a felismerésre (ld. 25. táblázat adatait)

A követő magánhangzó típusa itt is hatással van a felismerésre. Az artikulációban ez úgy jelentkezik, hogy a mássalhangzó időtartama kisebb magas nyelvállású magánhangzó mellett (Kassai 1979, 48). Az artikuláció akusztikai vetületének vizsgálatai kimutatták, hogy a szótagszám növekedésével arányosan csökken a mássalhangzó időtartama. A tendencia kissé eltérő ettől a percepció adatokban. A legalsó nyelvállású [a:] mellett volt "szükség" a rövidebb időtartamú [s]-re; a [ʃs], [t] felismerését illetően nagyjából azonosak a fizikai időtartamértékek a különböző magánhangzók mellett. A percepció adatokat a 25. táblázat tartalmazza. Adataim Magdics Klára értékeit igazolják, amelyek szerint az időtartam-csökkenés sorrendje a háromféle mássalhangzót tekintve a következő: zöngétlen réshang → zöngétlen affrikáta → zöngétlen zárhang (1965).

Az intenzitás mint felismerési kulcs mássalhangzók azonosításában

A kísérletben arra próbáltam választ kapni, hogy a specifikus intenzitás mely mássalhangzók esetében jelentkezik elsődleges és melyeknél másodlagos felismerési kulcsként.

Anyag és módszer

Az [a:] és az [i:] magánhangzókkal olyan hangkapcsolatokat állítottunk össze, amelyekben -- egyszer a magánhangzót megelőzően, egyszer pedig követően -- valamennyi magyar mássalhangzó előfordult (pl. bá, áb, bí, íb, szá, ász, szí, ísz). Ily módon 92 hangkapcsolatot kaptunk, ezeket férfi ejtésben, véletlen sorrendben magnetofonszalagra rögzítettük. A hanganyagot fehér zajjal elfedtük: a jel/zaj viszony 0 dB volt. A zajjal fedett hanganyagot 2 hetes időközönként 30 kísérleti személlyel hallgattattuk le csendes teremben. A kísérleti személyek feladata a felismert hangkapcsolatok helyesírásban történő rögzítése volt.

Eredmények

A kísérleti személyek összesen 2760 választ feldolgoztuk, és táblázatban összesítettük az egyes mássalhangzókra, a helyes felismerés csökkenésének arányában (27. táblázat).

27. táblázat

Eredeti hangkapcsolat	Felismerés %	Eredeti hangkapcsolat	Felismerés %	Eredeti hangkapcsolat	Felismerés %	Eredeti hangkapcsolat	Felismerés %
<u>csá</u>	100	<u>szí</u>	95	<u>áj</u>	100	<u>isz</u>	100
<u>lá</u>	100	<u>sí</u>	95	<u>ázs</u>	95	<u>ís</u>	100
<u>bá</u>	100	<u>tí</u>	90	<u>ács</u>	94	<u>ics</u>	63
<u>szá</u>	95	<u>zsi</u>	82	<u>ás</u>	90	<u>íc</u>	60
<u>zá</u>	95	<u>ci</u>	77	<u>ám</u>	90	<u>íl</u>	60
<u>sá</u>	95	<u>ri</u>	70	<u>ál</u>	86	<u>íz</u>	40
<u>zsá</u>	95	<u>mi</u>	68	<u>ész</u>	73	<u>ízs</u>	36
<u>má</u>	95	<u>csi</u>	65	<u>ág</u>	70	<u>ít</u>	29
<u>rá</u>	94	<u>lí</u>	65	<u>ár</u>	40	<u>ír</u>	29
<u>já</u>	90	<u>ví</u>	54	<u>áp</u>	40	<u>ím</u>	29
<u>gyá</u>	88	<u>gí</u>	50	<u>ágy</u>	36	<u>ív</u>	29
<u>dá</u>	77	<u>kí</u>	47	<u>án</u>	30	<u>ín</u>	17
<u>vá</u>	77	<u>nyí</u>	40	<u>ány</u>	27	<u>íg</u>	13
<u>ká</u>	68	<u>tí</u>	23	<u>ác</u>	18	<u>íj</u>	12
<u>gá</u>	60	<u>ní</u>	23	<u>ád</u>	17	<u>íny</u>	-
<u>tyá</u>	60	<u>jí</u>	18	<u>át</u>	17	<u>íd</u>	-
<u>há</u>	50	<u>bí</u>	17	<u>ák</u>	17	<u>íb</u>	-
<u>pá</u>	36	<u>pí</u>	11	<u>áz</u>	14	<u>íp</u>	-
<u>tá</u>	36	<u>dí</u>	-	<u>áb</u>	13	<u>ík</u>	-
<u>ná</u>	30	<u>gyí</u>	-	<u>áty</u>	-	<u>így</u>	-
<u>fá</u>	23	<u>tyí</u>	-	<u>áv</u>	-	<u>íty</u>	-
<u>cá</u>	18	<u>fi</u>	-	<u>áf</u>	-	<u>íf</u>	-
<u>nyá</u>	17	<u>hí</u>	-	<u>áh</u>	-	<u>ih</u>	-

A táblázat a konkrét értékeken túl a következőkről is tájékoztat: a) a mássalhangzó felismerése nem független a szomszédos magánhangzótól és b) a felismerés nagymértékben függ a fonetikai helyzettől, jelen esetben attól, hogy megelőzi-e vagy követi a magánhangzót. A százalékos arányok azt mutatják, hogy általában jobb a mássalhangzók azonosítása a nagy specifikus intenzitású magánhangzó, az [a:] mellett, rosszabb a leggyengébb intenzitású [i:] mellett. Néhány esetben fordított eredményt is kaptunk: a [ts̃]-t például pontosabban felismerik az [i:] (77 %), mint az [a:] (18 %) mellett.

Legbizonytalanabb a felpattanó zárhangok ([b, p, d, k] stb.), valamint a [v, f, h] réshangok azonosítása.

A fehér zajjal történt elfedés a teljes frekvenciaspektrumot betölti, tehát a beszédhangokat jellemző akusztikai paraméterek közül az intenzitásban jelent elsőrendű torzulást: csökkenést. Összevetettük adatainkat a más-salhangzók -- természetes ejtés alapján mért -- specifikus intenzitásával. A specifikus intenzitást százalékban fejeztük ki, 100 %-nak a 0 dB-t (pl. a [l] intenzitása) és 0 %-nak a -27 dB-t (pl. hangsorvégi [t] intenzitása) vettük (28. táblázat). A táblázatban összegeztük a kétféle magánhangzóra kapott adatokat.

28. táblázat

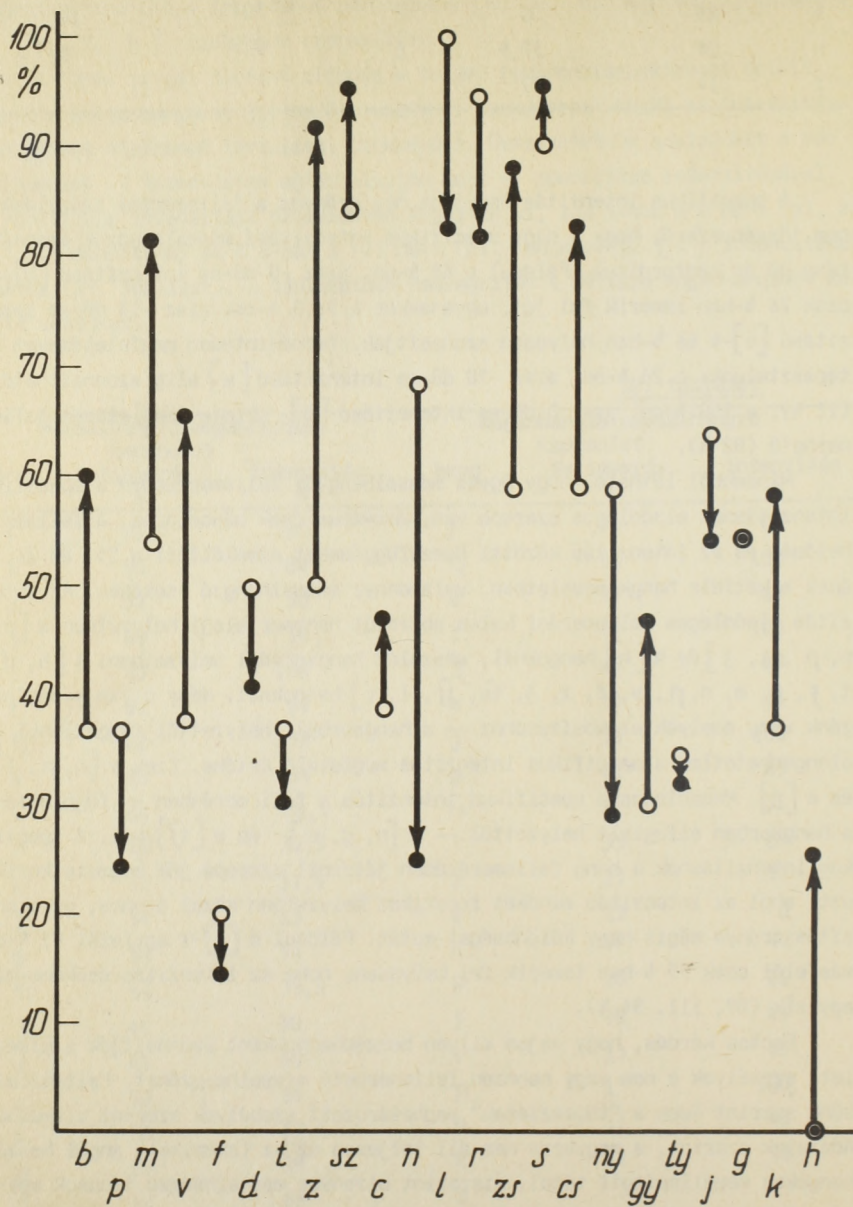
Mássalhangzó+magánhangzó kapcsolat			Magánhangzó+mássalhangzó kapcsolat		
hang	felismerés %	intenzitás %	hang	felismerés %	intenzitás %
f	95	90	f	94	82
s	95	84	t̂f	78	100
z	92	50	s	77	63
ʒ	88	59	l	72	100
l	83	100	ʒ	65	49
r	82	94	m	59	85
t̂f	82	59	j	56	100
m	80	54	g	42	45
v	66	38,5	t̂s	38	63
b	60	37,5	r	35	89
k	57	37,5	z	27	77
j	54	63	t	23	0
g	54	54	n	23	82
t̂s	47	38,5	p	20	26
ʃ	46	30	f	18	45
d	41	50	v	15	45
c	32	34	ɲ	13	100
t	30	37,5	d	11	45
ɲ	29	59	k	11	26
n	26	86	b	7	45

h	25	0	c	-	18
p	24	37,5	f	-	0
f	14	20			

A specifikus intenzitás nagysága nem arányos a felismerési százalékkal; nem törvényszerű, hogy a nagy specifikus intenzitású mássalhangzók azonosítása jó és megfordítva. Például a 68 %-os, azaz -8 dB-es intenzitású [n]-t csak 26 %-ban ismerik fel jól, ugyanakkor a 38,5 %-os, azaz -14 dB-es intenzitású [v]-t 66 %-ban helyesen azonosítják. Természetesen megfeleléseket is tapasztalunk: a 26 %-os, azaz -20 dB-es intenzitású [k] alig azonosítható (11 %), a 100 %-os, azaz 0 dB-es intenzitású [t] szinte tökéletesen felismerhető (82 %).

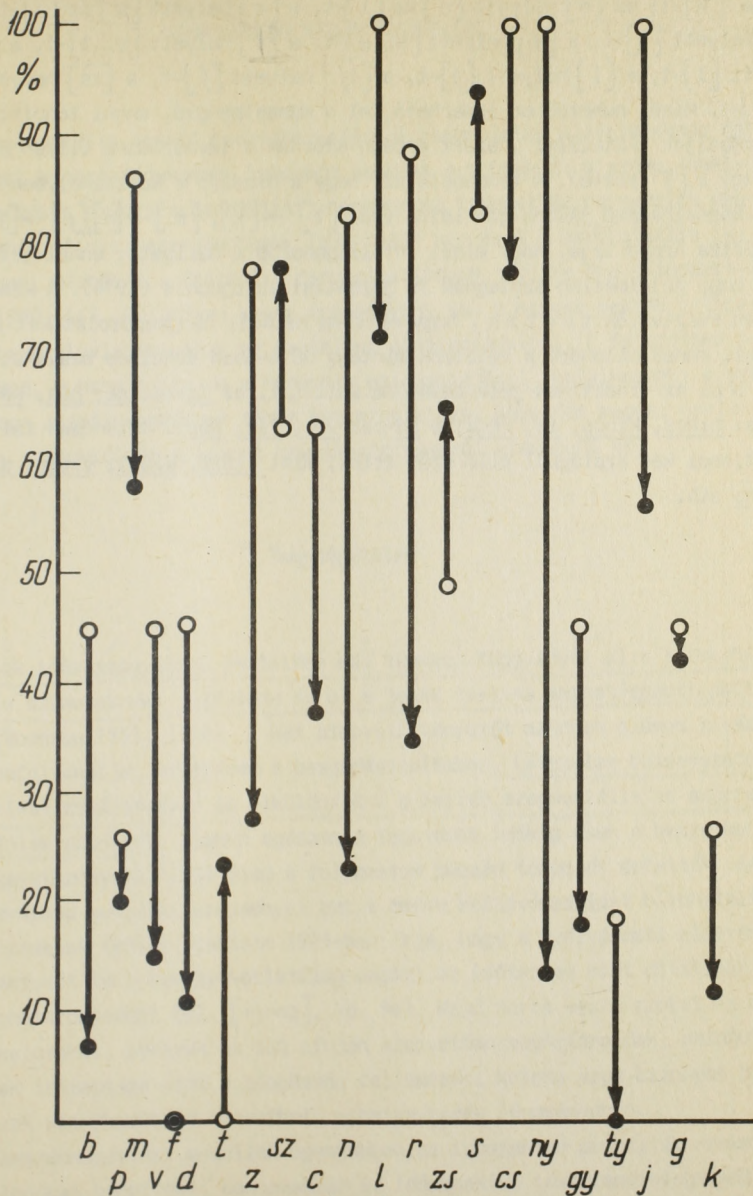
Mindebből látható, hogy egyes mássalhangzók felismerésében a specifikus intenzitásnak elsődleges szerepe van, másokéban csak másodlagos. A felismeretőség és az intenzitás közötti összefüggéseket szemlélteti a 25. és 26. ábra a kétféle hangkapcsolatban, valamennyi mássalhangzó esetében. Az intenzitás elsődleges felismerési kulcs abszolút hangsor eleji helyzetben a [p, n, ŋ, l, j] és a [h] hangoknál; abszolút hangsorvégi helyzetben: a [b, d, t, ʃ, k, m, n, ŋ, v, f, z, j, ts, tʃ, l, r] hangoknál. Négy olyan mássalhangzónk van, amelyek azonosításához -- a hangsorbéli helyzettől függetlenül -- elengedhetetlen a specifikus intenzitás megfelelő értéke. Ezek a [n, ŋ, l] és a [j]. Másodlagos a specifikus intenzitás a felismerésben -- függetlenül a hangsorban elfoglalt helyzettől -- a [c, g, s] és a [tʃ]-nél. A specifikus intenzitásnak a hang felismerésében játszott szerepe jól megmutatkozik ott, ahol az intenzitás mindkét fonetikai helyzetben közel azonos, az azonosítás aránya mégis nagy különbséget mutat. Például a [r]-t egyfelől 82 %-ban, másfelől csak 35 %-ban ismerik fel helyesen, noha az intenzitás csaknem megegyezik (89, ill. 94 %).

Fontos kérdés, hogy vajon milyen beszédhangokként azonosítják a kísérleti személyek a nem vagy nehezen felismerhető mássalhangzókat. Feltételezésünk szerint ezek a "tévesztések" meghatározott szabályok szerint alakulnak. Adataink szerint, a magyarra nem áll teljesen az az (amerikai) angol beszédhangokra megállapított tétel, miszerint bizonyos mássalhangzó-típusok egymás között cserélődnek a felismerésben (Flanagan 1965). A "szabályosnak" tekinthető hibázások a következők: a [b] helyett [d, g, v]-t azonosítanak, a [p] helyett [t]-t, a [t] helyett [k]-t, a [ʃ] helyett [j]-t, a [c] helyett



25. ábra

Az intenzitás (o) és a felismerés (●) arányának viszonya a magyar mássalhangzóknál (hangsor eleji helyzetben)



26. ábra

Az intenzitás (o) és a felismerés (●) arányának viszonya a magyar mássalhangzóknál (hangsorvégi helyzetben)

[f] és [k]-t; az [m] helyett [n] és [l]-t, a [n] helyett [m] és [l]-t, a [ŋ] helyett [j]-t; a [v] helyett [b, g]-t, a [f] helyett [p, t]-t, a [r] helyett [l]-t, a [l] helyett [j]-t, a [tʃ] helyett [ʃ]-t, a [tʃ̂] helyett [s] és [t]-t. Minél nehezebben ismerhető fel a mássalhangzó, annál többféleképpen próbálják "kitalálni", ezért erősen szórnak a tévesztések (ilyen pl. a [d] vagy a [f] esete). A tendencia az, hogy a hibázás a képzéshelyben hátrábbi mássalhangzó javára jön létre (vö. [t] → [k], [m] → [n], [ŋ] → [j] stb.).

Götze Árpád írja, hogy minél intelligensebb a hallgató, annál jellemzőbb, hogy értelmetlen szótagnak is jelentést tulajdonít (1974). A kísérletünkben résztvevők tudták, hogy értelem nélküli hangkapcsolatokat fognak hallani, ennek ellenére a válaszok mintegy 30 %-ában értelmes szavakat írtak ott, ahol az eredetinek jelentése nem volt (pl. az ís-re: és, hús, kés; az ícs-re: nincs, kincs, is, Pécs; a bá-ra: vár, fáz, bár). Különösen feltűnő ez ott, ahol két szótagból álló szót írtak, mint kiáll, kiált, virág, diák, kilincs stb.

SZUPRASZEGMENTUMOK ÉSZLELÉSE

A szegmentális szintű beszédészleléssel kapcsolatos kísérletekhez és modellekhez képest lényegesen kevesebb azoknak a munkáknak a száma, amelyekben a szerzők a szupraszegmentális jelenségeket tárgyalják percepciós megközelítésben. A szupraszegmentumok funkcióját egyrészt a perceptuális integritásban, másrészt a frazeális csoportosításban jelölik meg (Studdert-Kennedy 1980, 46). Különösen kevés eredményünk van a magyar beszéd intonációs és hangsúlyviszonyainak az észlelési sajátosságairól. E fejezetben a beszédtempó, a hanglejtés és a hangsúly percepciójával foglalkozunk: a szakirodalomban leírt eredmények tömör összefoglalását követően magyar nyelvi kísérleti adatokról és a levont következtetésekről számolunk be.

Tempóészlelés

Jelek időviszonyainak észlelése két tényező függvénye: a) a jelek hiányának (a szüneteknek) jelenléte és b) a jelek tartama meghatározott periódusban (Woodrow 1951, 1224). E két alapvető tényezőt azonban számos további faktor befolyásolja, különösen a beszédészlelésben. Látszólag célravezető az észlelésvizsgálatokban is elkülöníteni a beszéd szegmentális és szupraszegmentális szintjét: ezeken egészen a legutóbbi időkig csak a beszédhangok időtartam-viszonyait, illetőleg a folyamatos beszéd tempóját értették. Az utóbbi évtized kutatási eredményei ezt a merev kettéválasztást bizonytalanná, majd kérdésessé tették. Lehiste 1984-ben írja, hogy a nyelvészeti elemzés az idő aspektusát tekintve gyakorlatilag megáll az időtartam mint disztinktív jegy meghatározásánál (pl. [+long], vö. 96). Majd sorra veszi azokat az észlelési kulcsokat, amelyek az idő síkján alapvetően meghatározóak, például az átmenetek időtartama mint a zöngésség felismerési kulcsa vagy bizonyos beszédrészek viselkedése a mondatbeli elhelyezkedés függvényében.

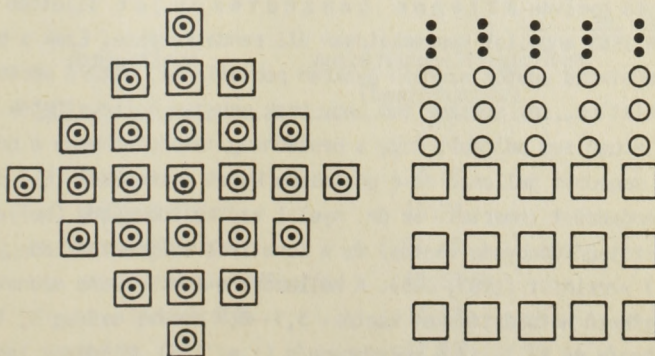
A beszédmegértési modellek ugyanebben az időszakban próbálták ezeket az idői kulcsokat beépíteni, pontosabban az időaspektus invarianciafogalmát meghatározni, s az egyes nyelvekre lebontva egyértelműen jellemezni. Az alapkérdés itt az, hogy miként képes a hallgató észlelni és interpretálni az

időzítési kulcsokat, amikor minden egyes beszédsegmentum változhat az idő síkján, s ezt számos különféle tényező előidézheti. Máig a legátfogóbb választ erre Klatt modellje kínálja (Klatt 1982). Eszerint a hallgatónak mindössze 5 szabályt kell alkalmaznia az észlelési folyamatban, s ezek segítségével megoldhatók az észlelés időzítési problémái. Az 5 szabály sorrendben a következő: 1. Az egyes akusztikai események időviszonyai megegyeznek az elvárásokkal, vagyis a szöfelismerés hipotézise megvalósul; 2. Ha az összes időtartam rövidebb, mint az elvárt (feltételezett), akkor valószínű, hogy a szó egy összetétel első tagja; 3. Ha valamennyi időtartam megegyezik a feltételezettel, s a szó beazonosítható lenne, csupán az utolsó magánhangzó, illetve az utolsó magánhangzó és a mássalhangzó(k) hosszabbak a vártnál, akkor nyilvánvaló, hogy a szó egy szintaktikai egység utolsó tagja; 4. Ha minden időtartam megfelel az elvárásoknak, de egy szótag hosszabb a többinél (és ez nem az utolsó), akkor valószínűleg hangsúlyos szóról történik döntés; 5. Bármi egyéb jelenség következik be, egyfajta lexikális (szöfelismerési) stratégiát kell alkalmaznunk (i.m. 84--85). Nagy előnye Klatt modelljének, hogy a szabályok felállítása és az azok alkalmazása közötti paradoxont egy egymásra épülő véges számú szabályrendszerrel igyekszik feloldani. Mi ez a paradoxon? Egy nyelvi jel időtartamát igen sok alternatív szabály rövidítheti vagy hosszabbíthatja, honnan tudhatja tehát a hallgató, hogy ezek közül melyiket válassza az adott feldolgozási folyamatban? Az 5 szabályból álló modellnek nagy érdeme, hogy megpróbálja összekapcsolni a szegmentális és a szupraszegmentális észlelési stratégiát (egyetlen folyamatba integrálva, ahogyan az a valóságban lejátszódik), s nem egy pontján igen jól alkalmazható bármely nyelvre. Szembetűnő hátránya, hogy lényegében mégiscsak a szószintig vagy szókapcsolat-szintig jut el (vö. 3. szabály), valamint hogy az utolsó szabály egy "joker", amely bármikor elővehető és alkalmazható, ha a felsoroltak egyike sem válik be. Ezáltal a szabályrendszer ugyan rugalmassá válik, de lényegében bizonytalanná is, hiszen nem tudjuk, hogy a "joker" szabály pontosan mit tartalmaz, s használatának mik az időbeli feltételei.

Az időzítés problémakörének csupán az említett két területe is óriási kutatómunkát igényel, ha valamennyi kérdésre nyelvspecifikusan, kísérletileg igazolt eredményekkel akarunk választ adni. A jelen tanulmány mindezek közül egy részterületet próbál áttekinteni a szakirodalomban, illetőleg kísérletileg vizsgálni magyar anyagon, magyar anyanyelvű kísérleti személyekkel. Az eredmények alapján a felvázolt két problémakörre visszatérve igyekszik részleges választ, illetőleg megoldási javaslatot adni. A vizsgálatok tárgya a

beszédtempó észlelési sajátosságainak, valamint a tempóészlelés és a beszédmegértés összefüggéseinek elemzése.

A szakirodalom jelen eredményei alapján megállapítható, hogy az időzítés tekintetében viszonylag sokat tudunk: a) az egyes beszédhangok időtartamának az észlelésre gyakorolt hatásáról (a magyarra vö. Gósy 1986), b) az átmenetek, illetőleg a beszédhangok részeinek az időészlelést befolyásoló szerepéről és c) a folyamatos beszéd tempóészlelési sajátosságairól. A folyamatos beszéd időzítésének vizsgálata erősen hangsúlyozza a viszonyítás szerepét és tényét. Tudjuk, hogy egy hangszínkép és a cochlea neurális kimenete között igen sok a különbség, különösen a temporális viszonyokat illetően. Szemléletesen mutatja ezt Scott gyémánt-hasonlata (1980, 52): a komplex akusztikai hullámforma olyan, akár a gyémánt szerkezete. Amit erről a szerkezetről tudunk, az annak elemei, részletei, s azok jellemzői (27. ábra). Jóformán azonban semmit nem tudunk ezeknek az elemeknek az összekapcsolódási mechanizmusáról, egymásra hatásukról, az elemek együtteséről: a jobb oldali kép már nem adja a gyémántszerkezetet. Ez természetesen nem azt jelenti, hogy az elemek, részek vizsgálata, elemzése felesleges, csupán azt kell tisztázni, hogy az elemek beható ismerete még nem egyenértékű a komplex viselkedésük ismeretével.



27. ábra

A gyémántszerkezet-hasonlat (a teljes akusztikai jel /bal oldalon/, az elemzett paraméterek /jobb oldalon/)

A beszédtempó észlelésével kapcsolatos eredmények két nagy csoportra oszthatók: 1. a tempóészlelést meghatározó tényezők vizsgálata és 2. a beszédtempó és a beszédmegértés kapcsolatának vizsgálata. Tanulmányok sorozata megegyezik abban, hogy a tempóészlelést alapvetően három tényező határozza meg: az artikulációs sebesség (a beszédhang-sorozatok képzésére fordított idő és a létrehozott beszédhangok vagy szótagok száma), a szünetek száma és a szünetek időtartama (vö. Feldstein--Bond 1981, 390). A három tényezőt együttesen számba véve beszélünk a beszéd tempójáról vagy a beszéd sebességéről. Shapley adatai szerint a kettő a következőképpen viszonylik egymáshoz: férfi ejtésben a beszédtempó 3,9--4,2 szótag/másodperc, női ejtésben ugyanez 3,9--4,4 szótag/másodperc; az artikulációs sebesség ugyanakkor férfi ejtésben 4,2--5,7 szótag/másodperc, ugyanez női ejtésben: 4,7--6,5 szótag másodpercenként (1987, 74). Megjegyezzük, hogy a beszédsebesség számításakor a szerző csak azokat a csendes szüneteket vette figyelembe, amelyek a 0,3 s-nyi (azaz 300 ms-nyi) időtartamot elérték. Ez azonnal jelzi a beszédtempó mérésének egyik legnagyobb nehézségét, a szünet meghatározása (definíciója) és kiszámítása körüli problémát (vö. Szende 1976; észlelésire: Kassai 1988). A tempóészlelési vizsgálatokat tovább nehezíti a különféle más faktorok hatásának felismerése, kimutatása, fontossági sorrendjük megítélése, mint például az alaphangmagasság-változás, az átlagos intenzitás, a szógyakoriság, a szótagszerkezet, a szintaktikai sajátságok szerepe (vö. Cutting--Pisoni 1978).

Az egyes nyelvek átlagos beszédtempóját illetően a legtöbb adat az amerikai angollal kapcsolatosan áll rendelkezésre. Ezek a beszédsebességre vonatkozó adatok azonban gyakran pontatlanok, eltérő mérések adataiból születtek, így sokszor nem vagy csak nehezen összevethetők. Nincs egyetértés abban sem például, hogy a beszédhang/másodperc vagy a szótag/másodpercben megadott jellemző ad-e pontosabb képet a tempóról. Első den Os háromféle tempómérést ismerttet: az ún. nyelvi szótag/időegység (helyesírás), a fonetikai szótag/időegység (ejtés) és a fonetikai szegmentum/időegység (beszédhangok) szerintit (1985, 125). A hollandra a fenti három módnak megfelelően a következő határértékeket kapták: 5,5--8,9 nyelvi szótag/s, 5,1--7,4 ejtett szótag/s és 13,1--19,4 beszédhang/s (i.m. 126). Mindezek ismeretében közöljük a 29. táblázatban néhány nyelvre megadott beszédtempó-értékeket. Nem szerepeltetjük ott Goldman-Eisler már klasszikusnak számító adatát az angolra vonatkozólag (1961, 171), amely 4,4--5,9 szótag másodpercenként (ez saját durva átszámításunkban kb. 21,4 beszédhang/másodpercet jelent). A táblázatban közölt adatokat Vaane (1982), den Os (1985), illetve Osser--Peng

(1964) munkáiból vettük az ott szereplő eredeti alapján. Az adatok egy része szótag/s, a másik része pedig beszédhang/s formájában volt megadva.

29. táblázat

Nyelv	Beszédtempó	
	szótag/s	beszédhang/s
holland	5,5--9,3	13,1--19,4
francia	4,7--6,8	
spanyol	4,6--7,0	
arab	4,6--7,0	
olasz	5,3--8,9	12,1--19,4
angol	4,5--8,03	
amerikai angol		9,92
japán		9,54

A magyar nyelvre Fónagy és Magdics mérési adata: átlagosan 11,35 beszédhang másodpercenként (1960, 185); a 30. táblázatban az általuk mért értékeket foglaltuk össze (i.m. 186). (A magyar beszéd tempójára vonatkozóan korábban Tarnóczy Tamás végzett méréseket, hasonló eredménnyel.)

30. táblázat

Szövegtípus	Beszédtempó-átlagértékek
	(beszédhang/s)
vers	9,4
mese	9,57
prózaolvasás	10,73
gyermekbeszéd*	11,68
párbeszéd (beszélgetés)	12,89
sportközvetítés	13,83
átlag	11,35

* Életkor: 5--6 év,
nagycsoportos óvodások

Ezekhez az adatokhoz annyit kell hozzáfűznünk, hogy vizsgálataink szerint a

beszédtempó az elmúlt 28 év alatt jelentősen felgyorsult, s az átlag magasabb érték a Fónagyék által közöltekénél.

A továbbiakban a tempóészlelésben szerepet játszó egyéb akusztikai paraméterek vizsgálatával kapott eredményeket ismertetjük röviden. Különböző módszerű kísérletekkel próbáltak választ kapni arra, hogy az átlag hallgató képes-e az elhangzó beszéd sebességének közel-objektív megítélésére (amikor szemantikai sajátságok az ítéletet nem befolyásolják). Közismert ugyanis az a pszichoakusztikai benyomás, hogy az anyanyelvi beszédtempóhoz képest az idegen nyelvi mindig gyorsabbnak tűnik, a hallgató idegen nyelvi kompetenciájától függően. Ezek szerint az eredmények szerint az átlag beszélő/hallgató képes az idegen nyelvi beszéd sebességét más idegen nyelvvel, illetőleg a saját anyanyelvével összevetni, s az objektív adatokhoz nagyon közeli, sokszor azokkal megegyező észlelési ítéleteket hozni (den Os 1985; Vaane 1982). Vaane megállapítása az eredményei alapján mégis az, hogy a tempóészlelésben nem a lexikai tudás a fő azonosítási kulcs. Ez valószínűleg nem teljesen igaz, csupán arra mutat, hogy a lexikai tudás éppúgy integrálva van a feldolgozási folyamatba, mint bármelyik más egyéb faktor, s az adott feladat (kommunikációs helyzet) határozza meg, hogy meghatározóvá válik-e vagy sem.

Feltételezhető, hogy az intenzitás és a frekvenciaváltozás hatással van a tempóészlelésre. Szubjektív döntésekkel (skála használata, ld. később) mutatták ki ezek együttes és külön-külön is ható szerepét (Feldstein--Bond 1981, 390). Kísérletükben háromféleképpen változtattak egy 30 s-nyi, spontán beszédből kivágott szövegrészletet (a változtatások az alapfrekvenciát és az átlagos intenzitást érintették). A megértési tényezőt úgy zárták ki, hogy az eredeti, illetőleg a változtatott beszédrészleteket visszafelé játszották le a kísérleti személyeknek. A kapott eredmények egyik lehetséges magyarázatául az ún. háló-effektus érvényesülését jelölték meg (391). Ennek lényege, hogy ha valaki adaptálódott egy dologhoz, akkor a többihez is adaptálódik; s ha egy faktor megváltozik, akkor ez maga után vonja a többi ugyanazon irányban történő változását -- hasonlóan a pókhálóhoz.

Az alaphangmagasságnak a tempóészlelésre gyakorolt hatását vizsgálva, den Os csaknem teljesen eltérő eredményekre jutott (1985). Adatai szerint az F0 nem játszik szerepet a tempóítéletekben, ugyanakkor megjegyzi, hogy ha az anyanyelvi szövegből kiiktatták a dallamot, ez a tempóészlelést megnehezítette. Ez az eredmény -- nagy valószínűséggel -- egy módszertani tévedésnek tudható be: a tanulmány szerint ugyanis a kísérleti anyagot szűréssel "monotonizálták", ami természetesen lehetetlen (hiszen bármilyen szűrés ellenére

a beszédjel a frekvenciaváltozást mindig tartalmazza); következésképpen a kísérletben nem azt az összetevőt vizsgálták, amelyre a magyarázatot adták. Ezt az is megerősíti, hogy ugyanebben a tanulmányban olvasható, hogy a monoton (?) mondatokat a kísérleti személyek gyorsabbnak hallották, mint a nem monotonokat (i.m. 173).

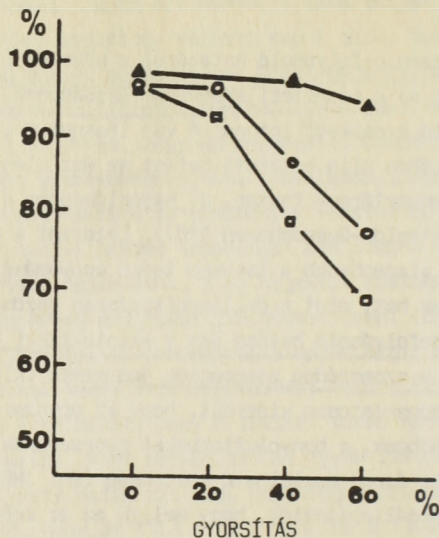
Az alaphangmagasság befolyásoló hatásáról a következő megállapítások állnak rendelkezésre: a) a kísérleti személyek lassabbnak ítélik azokat a mondatokat, amelyekben ereszkedő intonáció van (Hoequist 1983), ugyanakkor b) a hanglejtés általában alig mutatott hatást az észlelésre, c) az átlagos hangmagasság erősen meghatározó faktor, d) hasonlóképpen a hangmagasság-változás mértéke is (Rietveld--Gussenhoven 1987). Eszerint a magasabb dallamszint a gyorsabb, az alacsonyabb a lassabb tempó benyomását kelti; a monoton beszéd gyorsabbnak hat, mint a dallamváltozásban gazdag.

Tempóészlelést befolyásoló hatása van a szintaktikai határok helyének, a szótaghangsúlynak, a szomszédos szegmensek jegyeinek (Klatt 1980, 248). A szótagszerkezettel kapcsolatosan kiderült, hogy az egyszerűbb szerkezetűeket a hallgatók lassabbnak, a bonyolultabbakat gyorsabbnak ítélték (Hoequist--Kohler 1986, 21). Az éppen észlelhető különbség (ún. JND) meghatározását célzó kísérletekkel megállapították, hogy melyik az az időérték, amely ahhoz szükséges, hogy a kísérleti személy a szótag időtartamában bekövetkező változást észlelje: ezt az értéket kb. 20 ms-ban jelölték meg (Osser--Peng 1964).

A ritmikai szerkezet is hat a tempóészlelésre, ezt mind logatomokkal, mind szavakkal és mondatokkal végzett kísérletek igazolták német anyaggal (Hoequist--Kohler 1986; Kohler 1986). E két munkában a tempóészleléssel kapcsolatosan az alábbi eredményeket kapták: a) legfontosabb kulcs az átlagos időtartam, b) az átlagos hangmagasságszint fontosabb kulcs, mint a hangmagasság-változás, c) az eső dallam dominánsabb, mint az emelkedő (vö. Kohler 1986, 168). Felmerül a kérdés, hogy mindezen eredményekben melyek azok a tényezők, amelyek az általános emberi észlelési-megértési mechanizmus működéséből adódnak, és melyek azok, amelyek nyelvspecifikusak. Az a tény, hogy a beszédsebesség egy sokdimenziós jelenség (Kohler 1987, 170), megnehezíti a különválasztást, az együtt- és/vagy egymástól függetlenül ható faktorok fontossági sorrendjének megállapítását.

Ismeretes, hogy az életkor változásával együttjárnak bizonyos frekvenciafüggő változások a hallási rendszerben (Price--Simon 1983, 129). Mind a beszédfejlődés (anyanyelv-elsajátítás) folyamatában, mind egy bizonyos életkor elérésével, illetőleg meghaladásával (öregedés) a legszembetűnőbb válto-

zások a beszéd idői feldolgozásának síkján jelentkeznek. A 28. ábra szemléletesen mutatja a beszédtempó változásának (gyorsulásának) hatását a megértésre különböző életkorokban 18 évtől 70 éven túlig (Beasley--Maki 1976, 428).



28. ábra

A beszédtempó gyorsulásának hatása a beszédmegértésre különböző életkorokban (▲ = 18-27 évesek, ○ = 60-69 évesek, □ = 70 éven felüliek)

Jól látható, hogy a "fiatalok" (18--27 év) észlelése alig mutat változást, míg a 60 év körülieké már lényeges, s az "öregké" még feltűnőbb csökkenését jelzi a helyes válaszoknak a gyorsulás hatására. A gyermek tempóészlelési sajátosságaival igen kevés tanulmány foglalkozik. Amerikai angol anyanyelvű gyermekekkel végzett kísérletsorozat eredményeit összegzi a 31. táblázat (vö. Konkle--Beasley--Bess 1974). Magyar anyanyelvű gyermekekkel végzett vizsgálatok is igazolták a fokozatos fejlődést a beszéd időviszonyainak észlelésében (Gósy--Laczkó 1987). Hasonló eredményeket kapott a gyermekekkel végzett tempóazonosítási kísérleteiben Beasley--Beasley (1973) és King és Weston (1974); bár az összehasonlítás az eltérő nyelvi anyag és az eltérő módszerek következtében meglehetősen nehéz.

31. táblázat

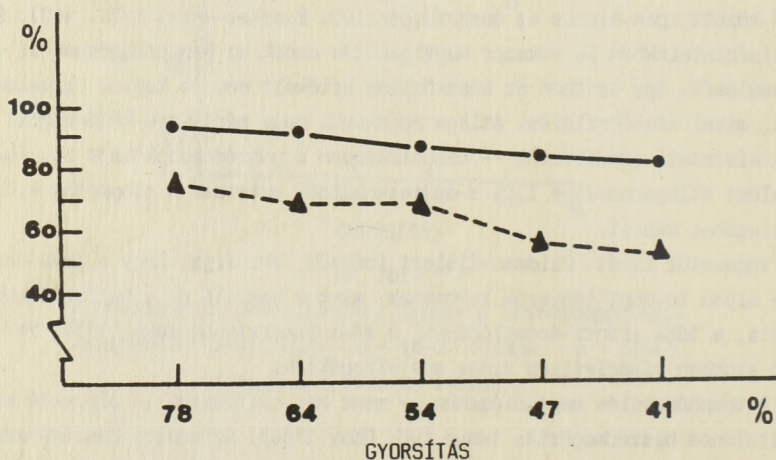
Gyermekcsoportok	Helyes felismerés %-ban		
	eredeti	Tempógyorsulás 30 %-kal gyorsabb	60 %-kal gyorsabb
4 évesek	88	82	68
6 évesek	94	92	88
8 évesek	100	98	88

Kísérletekkel még kevésbé igazolták, annál több azonban a tapasztalat és a megfigyelés arra vonatkozólag, hogy a tempóészlelés és a saját beszédtempó között igen szoros az összefüggés (vö. Beasley--Maki 1976, 442). Saját kísérleteinkben is sokszor megfigyeltük ennek az összefüggésnek az érvényesülését; egy esetben az adatközlőnk értékeit nem is tudtuk figyelembe venni, mivel tempóészlelési átlaga rendkívül nagy mértékben különbözött a többi kísérleti személyétől -- hasonlóképpen a beszédprodukcója is. Tempóészlelési átlageredménye 1,56 s-os reakcióidőt mutatott a csoportja 0,69 s-os átlagához képest.

Végezetül ismét Goldman-Eislert idézzük, aki írja, hogy a tempóészlelésre olyan további tényezők is hatnak, mint a beszélő és a hallgató kapcsolata, a téma iránti érdeklődésük, a téma ismeretének foka (1954, 94); ezeket azonban kísérletileg azóta sem vizsgálták.

A tempóészlelés megítélésében -- mint már említettük -- alapvető kulcs az általános beszédmegértés ténye (vö. Gósy 1988). Az emberi beszédészlelő rendszer egyfelől lehetővé tesz olyan feldolgozásokat, amelynek alapján szubjektíven eldönthető az adott beszédszakasz sebessége. Erre az ítéletre akkor is képesek vagyunk -- amint azt számos szakirodalmi adat bizonyítja --, ha az elhangzott üzenet tartalma bizonytalan vagy teljesen érthetetlen volt számunkra (pl. egy ismeretlen idegen nyelvet hallgatva vagy erősen beszédhibás ember beszédét figyelve). Ez a fajta feldolgozás a beszédmegértés fejlődésében (anyanyelv-elsajátítás folyamatában) különös jelentőséggel bír, és mindig megelőzi a tulajdonképpeni megértés kialakulását (Gósy 1984c). A legáltalánosabb beszédmegértési modellt szemlélve, ezek a feldolgozások még a fonetikai szinten lejátszódó működéseket (rendszerint) megelőzőleg mennek végbe. A szakirodalomban közölt (kevés számú) kísérleti eredmény arra enged következtetni, hogy a beszédtempó bizonyos mértékű fokozásával elérhető egy

olyan állapot, amikor az üzenet megértése lehetetlenné válik, de bizonyos észlelési működések még tetten érhetők, vagyis a fonetikai szint "még" dolgozik. Ez tulajdonképpen fordítottja annak, amikor ismert, de nem eléggé jól tudott idegen nyelv megértésében vannak ugyan nehézségeink, de az észlelési rendszer már kellőképpen jól működik: az üzenet ismétlése gyakorlatilag végrehajtható. A tempóészlelési kísérletek azt mutatják, hogy "észlelés még van, amikor megértés már nincs", illetőleg hogy a kettő között numerikusan kimutathatóak a különbségek. A 29. ábra fonetikailag kiegyenlített szósorok észlelésének és megértésének különbségét mutatja olyan beszédtempó-értékek mellett, ahol a beszédmegértés már sehol sem éri el a 100 %-ot (Foulke--Sticht 1967, 95).



29. ábra

A beszédészlelés (●---●) és a beszédmegértés
(▲---▲) különbsége

Amint az ábrából leolvasható, a különbség átlagosan 20 %-os, és a tempó gyorsulásával arányos növekedést mutat. Az idézett munkában nem találhatók pontos értékek a tempót illetően, így feltételezésekre vagyunk utalva az átszámításban (pl. beszédhang/s-ra). Az időkompresszióra vonatkozó százalékok irányadók e tekintetben.

Eltérő beszédsebesség eseteiben vizsgálták a mondatérthetőséget (Licklider--Miller 1951, 1066), az eredményeket a 32. táblázat mutatja.

32. táblázat

Tempó	Korrekt mondatérthetőség
115 szó/perc	közel 80 %
150 szó/perc	közel 74 %
170 szó/perc	közel 62 %

Noha a két mérési adat között 16 év különbség van, némileg tájékoztatnak a szó/percben kifejezett beszéd sebességéről. Az adatok értelmezéséhez a 33. táblázat további fogódzókát nyújt; férfi hangon ejtett mondatok gyorsítására kapott megértési eredményeket tartalmaz Beasley--Bratt--Rintelmann adatai alapján (1980, 728). A kísérletben részt vevő személyek felnőttek, egyetemi hallgatók voltak.

33. táblázat

	0 %	40 %	60 %	70 %
Időkompresszió				
helyes megértés	98,9 %	97,5 %	94,6 %	66,8 %

Megint más adat szerint (Tallal 1978, 85) az amerikai angol beszéd kb. 30 fonéma/szekundum (?) sebesség mellett még érthető: ez a tempó a szerző állítása szerint kb. 400 szót jelent percenként.

Befejezésként egy olyan aspektust említünk, amellyel ugyan nem célunk foglalkozni, de a tempóészlelési kutatások eredményeinek gyakorlati alkalmazása szempontjából igen fontos. Úgy tűnik, hogy az észlelésben és a beszéd-megértésben jelentkező nehézségek gyakran mindennél korábban az időzítés síkján érhetők tetten; a feldolgozási problémák legelőször itt mutathatók ki. Nagyothallókkal, siketekkel és afáziásokkal végzett kísérleti eredmények és tapasztalatok szerint az időzítés devianciái az észlelésben szignifikánsak már akkor, amikor más jellegű nehézségek még nem jelentkeznek. Gyakran éppen ezek hívják fel a figyelmet a később másutt is előtérbe kerülő dekódolási


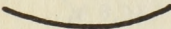
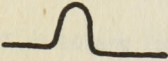


problémákra (vö. Stickt--Gray 1969).

Mindezeknek az eredményeknek és ismereteknek a figyelembevételével kísérletsorozatot terveztünk magyar anyanyelvű kísérleti személyekkel. Célja a tempóváltozás és zlelésre gyakorolt hatásának a vizsgálata volt.

Anyag és módszer

A kísérlet anyagául mesterségesen előállított mondatot választottunk: Maritól tegnap kérték a feladatokat. Ezt a közlést MEA 8000-es szintetizátor segítségével és Commodore 64-es személyi számítógép vezérlésével szintetizáltuk férfi hangú ejtésben. A mondat teljes időtartama a kiinduló változatban: 3,38 s, az "artikulációs sebesség" 9,2 beszédhang/s, azaz 3,8 szótag/s volt. A mondatot ötféleképpen generáltuk az F0, vagyis az alaphangmagasság változtatásával (minden egyéb paraméter változatlan maradt), ld. 34. táblázat.

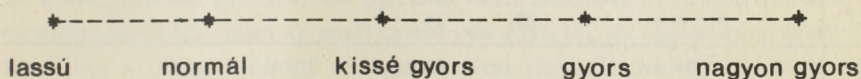
34. táblázat

Az F0 változásának sémája	Az alaphangmagasság változása időtartam (ms)	frekvencia (Hz)
	3380	100 Hz-en lebeg végig
	2450 80	100 -- 70 70 -- 110
	1500 100 100 1680	100 -- 80 80 -- 100 100 -- 70 70 -- 60
	2960 70 220	100 -- 70 70 -- 110 110 -- 80
	3380	100 -- 65 fokozatosan

Az ötféle dallamú mondatot Varyspeech segítségével (időkompresszió elektro-
mos úton) kétféleképpen meggyorsítottuk: a) az eredeti tempó 1,2-szeresére

és b) az eredeti tempó 1,6-szorosára; ennek megfelelően a mondat sebessége az első esetben 11 beszédhang/s (4,6 szótag/s), a másodikban 15 beszédhang/s-ra (6,2 szótag/s) változott. Mind a mesterséges beszédanyag, mind a Vary-speech alkalmazása biztosította, hogy valóban csak a beszédsebesség módosulása legyen az egyedüli eltérés a különböző dallamú mondatok között. Az így módon létrehozott 15-féle mondatot véletlenszerű sorrendben magnetofonszalagra rögzítettük.

Magyar anyanyelvű kísérleti személyek vettek részt a lehallgatási kísérletekben, összesen 25 fő (12 férfi, 13 nő, átlag életkoruk 35 év). A kísérletet egyenként folytattuk le csendesített teremben Uher magnetofont használva. A kísérleti személyek feladata az volt, hogy megítéljék az elhangzó mondatok tempóját egy előre kézbe adott tesztlap segítségével (a tesztlap részletét ld. 30. ábra). A kísérlet megkezdése előtt megadtuk a kitöltési útmutatót; azt kértük, hogy a kísérleti személyek a magyar beszédtempóról kialakított véleményükhöz hasonlítsák az elhangzó mondatok sebességét.



30. ábra

A tempóészlelési kísérlet tesztlaprészlete















A döntéseiket karikázással kellett jelölniük. A kísérletvezető előre közölte, hogy valamennyi mondat ugyanolyan tartalmú lesz, ez tehát nem okozott a lehallgatóknak váratlan meglepetést.

Eredmények

A kísérleti személyek eredményeit százalékos megoszlásban a 35. táblázat tartalmazza. Ezek az adatok több ponton alátámasztják a szakirodalomból már ismert tempóészlelési sajátosságokat; néhány ponton azonban eltérést tapasztaltunk. A kísérleti személyek képesek közel-objektíven megítélni az elhangzott közlés (itt: mondat) sebességét. Az adatok nem mutatnak a vártnál nagyobb szórást; ezt az egyéni különbségekkel magyarázhatjuk. Legnagyobb az észlelési egyöntetűség az eredeti tempójú mondatoknál, ezeket 70, illetve

annál nagyobb százalékban észlelték lassúnak. Jobban megoszlanak a vélemények az 1,2-szeresére gyorsított mondatok esetében, különösen a végig emelkedő és a szólamvégi csúcsot tartalmazó mondatoknál: itt -- bárha kismértékben is -- négy-négy fokozatban (a lassútól a gyorsig) fordultak elő ítéletek. Legnagyobb mértékben az 1,6-szorosára gyorsított (15 beszédhang/s) mondatokról akotott döntéseikben különböznek a kísérleti személyek, ezeket kissé gyorsnak, gyorsnak és nagyon gyorsnak egyaránt észlelik. A legtöbb ítélet szerint ezek a mondatok "gyorsak".

35. táblázat

Tempó	Dallamforma	Beszédtempó-ítéletek (%)				
		lassú	normál	kissé gyors	gyors	nagyon gyors
9,2 beszédhang/s 3,8 szótag/s		75	15	10		
		75	25			
		70	30			
		70	30			
		85	15			
11 beszédhang/s 4,6 szótag/s			80	20		
		10	30	50	10	
			70	30		
		10	55	30	5	
15 beszédhang/s 6,2 szótag/s			5	20	40	35
				15	60	25
				15	60	25
				15	60	25
				5	75	20

Megnéztük, hogy vajon az egyes dallamformák hatnak-e, s ha igen, miként, az észlelésre. Adatink szerint az eredeti tempó megítélésében gyakorlatilag alig van hatása a dallamformáknak: a kísérleti személyek csupán a szólemeleji csúcsot tartalmazó, valamint a végig emelkedő dallamformáknál jelöltek 30--30 százalékban gyorsabb tempót. Ebben a csoportban legnagyobb arányban a monoton dallamú mondatot találták lassúnak. Az 1,2-szeresére gyorsí-

tott mondatok esetében a dallam ugyancsak a szólamvégi csúcst tartalmazó, valamint a végig emelkedő mondatnál osztotta meg jobban a résztvevők ítéletét: a többihez viszonyítva ezeket a mondatokat gyorsabbnak találták. Eltérően az előző, objektíve lassúbb mondatok észlelésétől: itt a monoton dallamú közlést feltűnően gyorsnak ítélték, a 45 % normál tempó mellett 35 %-ban kissé gyorsat és 20 %-ban gyorsat jelöltek! Ez az adat egybevág a szakirodalomban talált, s a bevezetésben részletezett kísérleti eredményekkel. Az 1,6-szorosára gyorsított mondatok közül gyorsnak, illetve nagyon gyorsnak találták ismételten a csúcst tartalmazó közléseket, valamint a végig emelkedő intonációjúakat. Leggyorsabbnak ebben a csoportban is a monoton dallamút ítélték; ugyanakkor viszonylag lassúnak hallották az ereszkedő dallammenetű mondatot (ugyanazt mutatták az előző csoport adatai is).

Mindezek alapján eredményeink az alábbiakban foglalhatók össze: a) a kísérleti személyek tempóészlelése arányaiban megegyezik az objektív mérési eredményekkel: a három gyorsítási fokozatnak megfelelő három ítéletsorozat, b) a normálként közölt mondatok lassúnak, az enyhébb gyorsításúak normálnak, az erőteljesebben gyorsítottak gyorsnak hallatszottak, c) kimondható az -- a beszédprodukcióna ugyan szignifikánsan nem adatolt, de szubjektíve érzékelhető tény --, hogy a Fónagy és Magdics által közel 30 éve megállapított tempóadat revízióra szorul: az észlelés szintjén legalábbis bizonyítható beszédünk felgyorsulása, d) a beszéddallam, vagyis az alapfrekvencia változása a magyarban is befolyással van a tempó megítélésére: a mondatban található frekvencianövekedés és -csökkenés (dallamcsúcs) a gyorsabb tempó benyomását kelti, hasonlóképpen az emelkedő típusú dallamminta, függetlenül az aktuális beszédtempótól, e) a monoton dallamú mondatoknak a tempóészlelésre gyakorolt hatása -- adataink szerint -- nem független az objektív tempótól: egy bizonyos átlagérték elérése után jelentkezik csak az észlelésben a "gyorsító" hatás. Feltevésünk szerint erre az lehet a magyarázat, hogy egy bizonyos sebességérték elérésével az intonációs észlelési kulcsok felismerése és azonosítása nehezebbé válik, s ekkor a monotonitás, mint egy egybefolyó akusztikai jelsor, azt a benyomást kelti, hogy az elhangzott közlés gyorsabb, hiszen semmilyen frekvenciában tettenérhető változást nem tartalmaz. Lassúbb elhangzáskor világosan érzékelhető az intonáció hiánya, s rájövünk, hogy az intonációs kulcsok megtalálásának a nehézségét az okozza, hogy gyakorlatilag nincsenek: ezáltal a közlés "unalmassá", a dallammintát tekintve feldolgozhatatlanná válik, vagyis a dekódoló mechanizmus nem kap feladatot: a

tempóítélet szerint a közlés tehát lassú.

Összefoglalva a legfontosabbnak ítélt eredmények a következők: 1. a szubjektív tempóészlelési döntések megerősítik azt a tapasztalatot, hogy a magyar beszéd tempója az elmúlt évtizedekben meggyorsult; 2. a dallamstruktúra hatással van a tempóítéletekre, ez azonban nem független az akтуális (objektív) beszédtempótól; 3. a dallamcsúcsot, illetőleg emelkedő dallammintát tartalmazó közlések gyorsabbnak hatnak, mint az ereszkedő típusúak; 4. a tempó egy meghatározott értékéig a monoton dallamot a magyar anyanyelvűek lassabbnak, ezen értéken túl (tehát gyorsabb elhangzásban) gyorsabbnak ítélik az összes többinél.

A beszéddallam észlelése

Hadding és Studdert-Kennedy sokak által idézett szintetizált dallamokkal végzett kísérlete a frekvencia és az idő változásával arra keresett választ: mely tényezők dominálnak abban, hogy a kísérleti személyek az elhangzott szöveget kérdésnek vagy állításnak (kijelentésnek) érzékeljék (1972). Úgy találták, hogy a kezdő csúcs, a töréspont és a végpont szerint variált dallamokat annál nagyobb százalékban jelölték kérdésnek, minél magasabb frekvenciájú volt a végpont. Malley és Peterson a dallam és nyomaték kölcsönviszonyát vizsgálta szűrőzéssel eljárással 15 kétszótagú szó oly módon történő változtatásával, hogy a hanglejtés hol kérdő, hol állító volt, a hangsúly pedig hol az első, hol a második szótagon helyezkedett el. A kísérletben először mindkét paraméter (frekvencia és dinamika) részt vett a szólamban, majd a dallamváltozást kiiktatták, csak a dinamika maradt (1966). Eredményeik szerint az állító, illetve kérdő dallamok meghatározása 80-90%-ban tökéletes, ha mind az alaphfrekvencia, mind az intenzitáskomponensek szerepelnek; ha az alaphangváltozást eliminálták, a hangsúly önmagában csak 60%-os korrekt eredményt adott. A hangsúlyok észlelésében az alaphangnak jelentős a szerepe. Következtetésük: már rég nincs fonémaazonosítás, amikor prozódiai észlelés még mindig van.

A magyar dallamformák, illetőleg a dallam és a nyomaték percepció vizsgálatára két kísérletsorozatot végeztünk el.

Anyag és módszer

32 különböző dallamformájú magyar mondatot (női ejtésben) imitáltunk a beszéd-alaphang- és intenzitásimitátorral. A percepciók kísérlethez megfelelően előkészített anyagot (két részben) 10-10 főiskolai hallgatónak játszottuk le. Kísérletünkkel választ próbáltunk kapni arra, hogy a) hogyan észlelik a kísérleti személyek a dallamformák időviszonyait, a hangmagasságot és változásait, illetve a hangfekvést, hangterjedelmet, hangközöket; b) mi az akusztikai sajátosságok és a behelyettesített tartalom, illetőleg a nyelvi funkció viszonya.

A kísérleti személyek feladata az volt, hogy a hallott dallamokat az előre elkészített, három hangfekvést jelző vonalrendszerben ábrázolják a tesztalapon, és hogy e hanglejtéseknek adekvát nyelvi tartalmat helyettesítsenek be. A kitöltött tesztlapokat a 32 dallamminta alaphangmagasság- és intenzitásmérővel korábban kielemezett és mingográfval kiíratott görbéivel vetettük össze.

A nyelvi anyag mondatainak többsége értelmi (27 db), kisebb része (5 db) érzelmi jellegű volt. Tartalmilag 11 eldöntendő kérdést, 6 kiegészítendő kérdést, 5 felkiáltást, 4 befejezetlen kérdést, 3 felszólítást és 2 kijelentést tartalmazott (vö. 31. ábra).

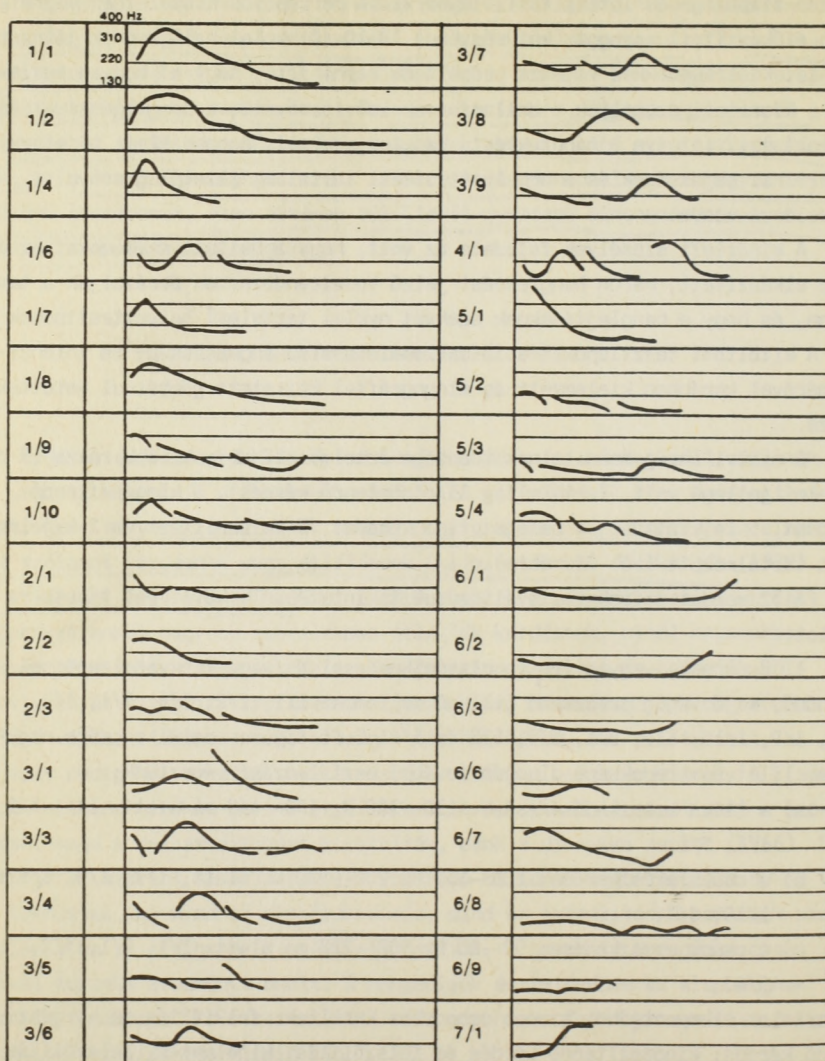
A 32 mondat a frekvenciaváltozások tekintetében a következő típusokra oszlott:

1. Csúccsal, vagyis frekvenciaemelkedéssel és/vagy frekvenciaeséssel kezdődő, majd végig ereszkedő jellegű dallamok: 1/1, 1/2, 1/4, 5/1, 1/7, 1/8, 1/9, 1/10, 6/7, 2/1, 2/2, 5/2 és 5/4. E fő típuson belül a csúcs nagysága, illetve meredeksége alapján további osztályozást végeztünk:

- a) a csúcs ereszkedése/esése: 130--180 Hz 150--430 ms alatt; 1/1, 1/2, 1/4, 5/1,
- b) a csúcs ereszkedése: 100--130 Hz 200--520 ms alatt; 1/7, 1/8, 1/9, 1/10, 6/7,
- c) a csúcs ereszkedése: 50--80 Hz 800--220 ms alatt; 2/1, 2/1, 5/2, 5/4.

A tartalom szempontjából azonos csoportba kerültek: felkiáltás és kiegészítendő kérdés; kiegészítendő kérdés és felszólítás; kijelentés, felszólítás és felkiáltás.

2. E csoport dallamai szintén kis csúccsal, frekvenciaemelkedéssel és -eséssel jellemezhetők, azzal a különbséggel, hogy ezt minden esetben 100--160 ms idejű ereszkedő dallamrész előzi meg. A csúcscok akusztikai paraméterei az



31. ábra

A 32 dallamminta sematikus ábrája

ereszkedő ágat figyelembe véve: 390 ms alatt 50 Hz, 100 ms alatt 70 Hz, 120 ms alatt 50 Hz és 220 ms alatt 110 Hz ereszkedés, illetve esés; 3/3, 3/4, 1/6, 2/3. Tartalmuk: kijelentés, felkiáltás, kérdés.

3. A dallam második felében helyezkedik el csúcs, ezek ereszkedő ágának paraméterei: 260 ms alatt 80 Hz, 200 ms alatt 180 Hz, 205 ms alatt 80 Hz, 250 ms alatt 70 Hz, 170 ms alatt 110 Hz és 100 ms alatt 60 Hz ereszkedés; 3/1, 3/5, 3/7, 3/9, 4/1, 5/3. Tartalmuk: eldöntendő kérdések.

4. Rövid: 360--550 ms időtartamú dallamok, sikló jellegűek (siklónak azt a dallamívet nevezzük, amikor valamilyen mértékű emelkedés vagy szökés után a dallam mintegy 20--60 ms ideig ereszkedik vagy lebeg). A dallam első 150--230 ms-os tartományában a lejtésment enyhén ereszkedő: 40, 50, 20 Hz különbséggel (3/6, 3/8, 6/), illetve lebegő (6/9, 7/1). Ezt követően emelkedik, majd ereszkedik a dallammenet: 100 ms alatt 80 Hz emelkedés, majd 220 ms alatt 90 Hz ereszkedés (3/6), 150 ms alatt 140 Hz, illetve 130 ms alatt 100 Hz (3/8), 80 ms alatt 20 Hz, illetve 50 ms alatt 20 Hz (6/6), 140 ms alatt 30 Hz, 30 ms alatt 20 Hz (6/9), végül 80 ms alatt 90 Hz szökés és 50 ms alatti lebegéssel végződő forma (7/1). Tartalmuk: eldöntendő, illetve befejezetlen kérdések.

5. Az utolsó csoportba azon dallamformákat soroltuk, amelyek lebegő jellegűek és a befejező 80--130 ms-on emelkednek: 20--80 Hz-et (6/1, 6/2, 6/3, 6/8). Tartalmuk: befejezetlen kérdés, eldöntendő kérdés.

Eredmények

Vizsgáltuk az időviszonyok alakulását, a hangfekvés, a hangmagasság változásainak észlelését. Az objektív időtartamok szubjektív alakulását az ábrák (rajzok) hosszúsága, valamint a behelyettesített dallamok nyelvi adatai alapján elemeztük. A műszeres mérések szerint a dallamok hossza: 360--3400 ms, a felső érték egyetlen esetben jelentkezik (1/2), ezt leszámítva a felső határ átlagosan 1800 ms. A rajzok tanúsága szerint a kísérleti személyek jól érzékelik az egyes dallamok időben jelentkező különbségeit; például a leghosszabb, 1/2-os dallamnál valamennyien kihasználják a tesztalapon rendelkezésre álló teljes hosszat. A rajzok cm-ben mért "időtartamai" is jól korrelálnak a behelyettesített tartalmak "hosszával".

Megnéztük, hogy a csúcsok időtartama hogyan tükröződik a rajzokban. Szemben a szólam teljes időtartamával, a csúcsok egymáshoz viszonyított időtartamai nem közelítik az objektív adatokat. Nem mindig arányos a csúcs és a teljes időtartam viszonya; a frekvenciaváltozásban azonos, csak időtartam-

ban és csúcshangfrekvencia-értékben különböző dallamok esetében (1/2 és 1/7) a rajzok a kettő között elsősorban frekvenciaváltozásban, hangmenetben jelentkező különbséget mutatnak. Időben azonban egyformának tüntetik fel ezeket.

A 2. típusba tartozó dallamoknál minden kísérleti személy érzékelte a csúcst megelőző 100-160 ms idejű ereszkedő dallamrészt; a rajzokon ezt vagy lebegésnek, vagy a csúcs hangmenetbeli változásának: elnyújtott emelkedésnek jelölik.

A hangmagasság és a hangfekvés észlelése. -- A dallamformák objektív hangterjedelme: 150--380 Hz az 1/1-től a 3/9-ig és 180--320 Hz a 4/1-től a 7/1-ig (az 5/1-nél 380 Hz a felső frekvenciaérték). Azokat a dallamokat, amelyek az adott frekvenciatartományon belül mindhárom hangsávot kihasználják, a kísérleti személyek (mindkét csoportban) szintén a rendelkezésükre álló teljes terjedelemben jelölték. Az objektív mérési eredményeket reprezentálják az 1, illetve 1,5 hangsávot kihasználó dallamábrázolások. Kísérleti személyeink megkülönböztették a hangmagasságban és a hangfekvésben jelentkező változásokat, a kettőt nem cserélik föl. Jelölik a csúcshangfrekvenciák különbözőségét is, a rendelkezésekre álló adatok szerint 50 Hz-nyi eltérés 90 %-ban, az ennél kisebbek 30--60 %-ban jelentkeznek a rajzokon. Legkevesbé megbízhatóan a szólamvégi frekvenciaemelkedést jelölték, noha ez az adott dallamok esetében 50--80 Hz-nyi eltérést jelentett.

A hangmagasságváltozás: a hangmenetek észlelése. -- A hangmenetek egyértelmű elemzéséhez az alábbi önkényes behatárolást választottuk: emelkedőnek akkor tekintjük a változást, ha az a 100 ms alatt 100 Hz frekvenciakülönbségnél nem nagyobb; ellenkező esetben, az ezt meghaladó frekvenciaváltozás esetén szökről, illetve eséről beszélünk. A rajzok alapján első megállapításunk az, hogy az objektíve azonos struktúrájú dallamoknak szubjektíve is megegyezők az ábrái. Sajátos a szólamleji csúcshang észlelése: az emelkedő (szökő) és ereszkedő (eső) ágat egyaránt tartalmazó esetében a jelölések 50--50 %-ban megegyeznek: az eredetinek megfelelő, illetve azon ábrázolás között, amely csak az ereszkedő (eső) ágat érzékelte--jelölte. Azt a szólamleji csúcst, amely "felszálló" dallamívet nem tartalmazott (5/1), a kísérleti személyek 90 %-ban helyesen, az objektív paramétereknek megfelelően érzékelték. A szólamleji csúcshang emelkedő (szökő) ágát annál jobban érzékelik, minél meredekebb. A 90 ms alatt 130 Hz-t szökő ágat 70 %-ban jelölik (1/4). Szemben például az ugyancsak 90 ms idő alatt, de csak 40 Hz-et emelkedő szólamleji csúcshang (1/10), ahol csak 40 %-ban érzékelték ezt a dallamívet. A dallam középső--hátsó részeiben jelentkező csúcshangokat pontosan

észlelik. Megvizsgáltuk a szólamvégek ábrázolását: az ereszkedést, az emelkedést, a siklást és a lebegést.

Az ereszkedő dallamvégeket (1/4, 1/6) 80 %-ban ereszkedőnek és 20 %-ban lebegőnek észlelték. Szólamvégi emelkedés öt dallamnál fordult elő: 1/9, 6/1, 6/2, 6/3, 6/7. A pszichoakusztikai eredmények szerint ezt a fajta dallamzárást a kísérleti személyek 48 %-a emelkedésnek (némelyikük meglehetősen erőteljes, meredek vonallal), 21 %-ban ereszkedésnek, 16 %-ban lebegésnek és 15 %-ban siklásnak érzékelték. Más képet kapunk, ha az egyes dallamokra lebontva nézzük meg az adatokat. A 6/1, 6/2, 6/3 és 6/7-es dallamok szólamvégi frekvenciaemelkedését 80--90 %-ban jelölték, míg az 1/9-nél, valószínűleg a szólemeleji csúcs következtében, csak 10 %-ban. A sikló dallamformákat 70 %-ban emelkedésnek, 13 %-ban ereszkedésnek, 10 %-ban siklásnak és 7 %-ban lebegésnek érzékelték.

A lebegő dallamot (akár szólamban, akár szólamvégen szerepel) vagy lebegőnek, tehát vízszintes vonallal, vagy ereszkedőnek jelölik. A lebegő szólamvéget tartalmazó dallamokat (2/1, 2/2, 2/3, 5/2, 5/1 és 1/8) 52 %-ban lebegőnek és 48 %-ban ereszkedőnek jelölték. A vizsgált dallamvégek akusztikai adatai: 70--220 ms időtartam alatt 170,180 Hz, egyetlen esetben 200 Hz. A lebegés érzékelésében két tényezőnek van szerepe: az időtartamnak és a teljes struktúrának, abban az értelemben, hogy milyen a szólemeleji csúcs és a dallamvég frekvenciaértékének viszonya. A nagy esésű kezdő csúccsal rendelkező dallamformák lebegő végeit inkább ereszkedésnek, 50--60 %-ban, a kis esésű kezdő csúcsot tartalmazó szólamok lebegő befejezéseit 70 %-ban lebegésnek érzik.

A zöngétlen hangok, amelyek az imitálásban szünetekként jelentkeznek, nem befolyásolták az észlelést, a kísérleti személyek minden dallamformát folyamatos vonallal ábrázoltak.

Az akusztikai paraméterek és a nyelvi funkció. -- Elemzésünket a forma oldaláról kezdjük, tehát az azonos struktúrára adott válaszokat vizsgáljuk. Az azonos típusú dallamok eredeti tartalmait összevetjük a behelyettesítésekkel.

1. A szólemeleji csúccsal rendelkező dallamokat elsősorban kiegészítendő kérdésnek és kijelentésnek ítélik, de felkiáltó és felszólító tartalmúak is akadnak. Az első alcsoport szólamaira kapott válaszok 43 %-ban kiegészítendő kérdések és 5 %-ban kijelentések (vö. 1/1, 1/2, 1/4, 5/1). Az eredeti tartalmak: felkiáltás és kiegészítendő kérdés. A felkiáltások behelyettesítései: 50 %-ban kiegészítendő kérdés, 40 %-ban felkiáltás és 10 %-

ban felszólítás. A kiegészítendő kérdéseké: 40 %-ban az eredetinek megfelelő, 24 %-ban felszólítás és 12--12--12 %-ban felkiáltás, kijelentés és eldöntendő kérdés.

A kisebb frekvenciaesésű szólemeleji csúcsot tartalmazó alcsoportban (1/7, 1/8, 1/9, 1/10, 6/7) a kapott válaszok megoszlása: 64 %-ban kiegészítendő kérdés, 16 %-ban felszólítás, 14 %-ban kijelentés és 6 %-ban felkiáltás. Az eredeti tartalmak: kiegészítendő kérdés és felszólítás. A kiegészítendő kérdések behelyettesítései 62 %-ban az eredetivel azonosak, 20 %-ban felszólítások, 12 %-ban kijelentések és 6 %-ban felkiáltások. A felszólításra adott válaszok 40 %-ban kiegészítendő kérdések, 30 %-ban kijelentések, 20 %-ban felszólítás és 10 %-ban felkiáltás.

Ezek az adatok -- mint egymáshoz intonációs típus tekintetében közel álló struktúrák nyelvi tartalmai -- azt mutatják, hogy a magas frekvenciájú és meredek esésű szólemeleji csúccsal rendelkező dallamok a kérdés, valamint a felszólítás/felkiáltás nyelvi jegyeivel rendelkeznek (150--200 ms határértékek között 130--180 Hz frekvenciaváltozás). Ha a felkiáltást emocionális jellegűnek tekintjük, akkor azt látjuk, hogy a 300 Hz, illetve az azt meghaladó frekvenciájú szólamok a percepcióban expresszív benyomást keltenek (vö. 1/1: 80 % felkiáltás). Természetesen az abszolút frekvenciaérték csak az egyik tényező, ugyanolyan lényeges a frekvenciaváltozás időbeli lefutása; például az abszolút frekvenciaértékben megegyező (380 Hz), de időbeli lefutásában eltérő formák észlelése különböző: a 150 ms alatt 150 Hz-et eső csúcsú dallamot 60 %-ban kiegészítendő kérdésnek és 40 %-ban felkiáltásnak ítélték. A szintén 380 Hz-ről -- 280 ms alatt 130 Hz-et -- ereszkedő csúcsú dallamformát már 80 %-ban felkiáltásnak ítélték. Közel azonos a kétféle struktúrára adott felszólító ítélet.

A harmadik alcsoportra jellemző szólemeleji csúcsok akusztikai adatai: 80--220 ms alatti 50--80 Hz frekvenciaesés, jól korrelálnak a kapott válaszokkal: 66 %-ban kijelentések, 20 %-ban felkiáltások, 7-7 %-ban felszólítás és kiegészítendő kérdés. A felkiáltás viszonylag magas aránya a következőkkel magyarázható. Ha önkényesen ugyan, de feltesszük, hogy a második alcsoportba sorolt dallamformák intonációs jegyei képviselik az általános, emóció nélküli nyelvi tartalmakat, akkor az ettől "feljebb" vagy "lejjebb", tehát magasabb vagy alacsonyabb hangfekvésben megvalósuló struktúrákban kell keresnünk az emocionális tartalmakat. Ezt igazolja az is, hogy míg a második alcsoportba tartozó dallamok hangterjedelme: 170--330 Hz, addig az első alcsoporté: 150--380 Hz, a harmadiké pedig: 170--270 Hz. A hangközök: 160 Hz,

illetve 230 Hz és 100 Hz. Az eredeti tartalmak: kijelentés, felszólítás és felkiáltás. A kijelentés dallamára kapott válaszok: 70 %-ban kijelentés, 20 %-ban kiegészítendő kérdés és 10 %-ban felkiáltás. A felszólítás dallamára: 70 %-ban kijelentés, 15 %-ban felkiáltás, 10 %-ban felszólítás és 5 %-ban kiegészítendő kérdés. A felkiáltás dallamára: 50 %-ban kijelentés, 40 %-ban felkiáltás és 10 %-ban felszólítás.

2. Az intonációs jegyeik alapján a második nagy csoportot alkotó dallamformák (1/6, 3/3, 3/4, 2/3) behelyettesített nyelvi tartalmak: 55 %-ban kijelentés, 15 %-ban kiegészítendő kérdés, 12 %-ban felszólítás, 8 %-ban felkiáltás, 5--5 %-ban pedig eldöntendő kérdés, illetve olyan kérdés, amely módosítószóval kezdődik. E struktúrák 100--160 ms elteltével kis csúcsot tartalmaznak, majd végig ereszkednek. Eredeti tartalmuk: csak, ugye módosítószóval kezdődő kérdés, felkiáltás és kijelentés. A felkiáltás behelyettesített tartalmak: 70 %-ban kijelentés és 30 %-ban felszólítás; a kijelentésé: 80 %-ban kijelentés, 10--10 %-ban felszólítás és kiegészítendő kérdés. A módosítószóval kezdődő tartalmak dallamaira adott válaszok: 35 %-ban kijelentések, 25 %-ban kiegészítendő kérdések, 15 %-ban felkiáltások, 10--10 %-ban eldöntendő kérdések, illetve módosítószóval kezdődő kérdések, 5 %-ban pedig felszólítások. Nem véletlen, hogy az eredetileg felkiáltást -- dallamstruktúrája alapján -- felszólításnak érzékelték (a 70 %-os kijelentés) mellett, mivel elsősorban hangköze, de hangterjedelme alapján is inkább rokonítható a felszólító (eredeti) tartalmú dallamokkal, mint a felkiáltásokéival.

3. A harmadik csoportba azokat a dallamformákat soroltuk, amelyeknek a második felében különböző nagyságú (frekvenciaértékű) csúcsok vannak. Ez az egyetlen dallamtípus, amelyet 100 %-ban azonosnak, s az eredetiekkel teljesen megegyezőnek jelöltek: eldöntendő kérdésnek.

4. A szólamvégi emelkedést tartalmazó dallamformáknál (1/9, 6/1, 6/2, 6/3 és 6/7), amelyek mind szólemeleji csúccsal, mind szólamvégi emelkedéssel rendelkeznek, a hallgatókban azt a benyomást keltik, hogy kérdést hallottak, mégpedig a (főleg Pesten) olyan elterjedt "felkapással" a végén (Kontra--Gósy 1987). Ez magyarázza, hogy 80--100 %-ban ítélték kiegészítendő kérdéseknek.

Jobban megoszlanak a vélemények a többi, szólamvégi emelkedést tartalmazó dallamformáról (annak ellenére, hogy struktúrájuk hasonló: a szólam végéig enyhén ereszkedő, helyenként lebegő dallamív). Hangterjedelmük: 180--

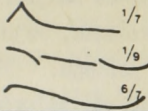
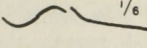
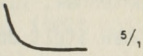
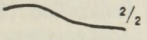
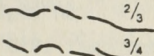
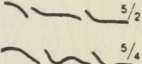
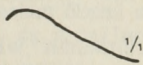
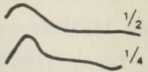
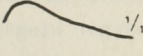
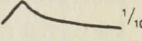
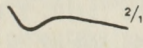
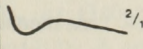
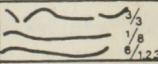
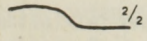
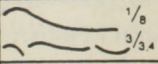
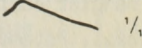
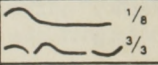
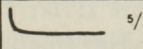
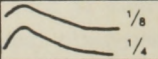
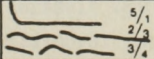
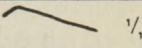
260 Hz, a hangköz: 80 Hz. Teljes időtartamuk: 960--1620 ms. Eredeti tartalmuk ún. befejezetlen kérdés, mint És erre a tanár? A kísérleti személyek válaszai eléggé megoszlóak: 28 %-uk idegennek, nem magyarnak minősítette a hallott dallamot. Ebben az ítéletben két akusztikai paraméter együttesen játszott szerepet: egyrészt a dallam szűk hangköze, másrészt az ehhez párosuló szólamvégi emelkedés. Önmagában egyik akusztikai jellemző sem kelti a nem magyar dallam benyomását, együttes megjelenésük azonban igen. 24 %-ban kiegészítendő kérdést írtak, amelyet az igen kicsi: 100 ms alatt 40--50 Hz-et eső szólameleji csúcs magyaráz mindhárom dallam esetében. 20 %-ban kaptunk eldöntendő kérdést válaszként, 10--10 %-ban kijelentésnek, illetve olyan mondatnak érezték, amelynek a végén egyszótagú kérdőszó van. Végül 9--9 %-ban felkiáltást és befejezetlen kérdést írtak. (Az eredeti tartalommal teljesen megegyező behelyettesítést nem, csak hozzá hasonlót kaptunk.)

5. Az ún. sikló dallamformák eredeti tartalma eldöntendő kérdés, illetve befejezetlen kérdés, mint Lakhelye? A behelyettesített tartalmak 92 %-ban egy és két szótagból álló kérdések vagy eldöntendő kérdések, 8 %-ban pedig felkiáltások (szintén egy és két szótagúak). A hosszabb eldöntendő kérdéseket az 1340 ms-os (6/8-as) dallamformára adták.

Az eltérő struktúrájú, de azonos eredeti tartalmú dallamok behelyettesítéseinek vizsgálata. -- Az összes kiegészítendő kérdést figyelembe véve azt látjuk, hogy az adott válaszok 52 %-ban kiegészítendő kérdések, 20 %-ban felszólítások, 12 %-ban kijelentések és 8--8 %-ban eldöntendő kérdések, illetve felkiáltások. Az eredetileg kijelentést tartalmazó dallamformák behelyettesítései: 75 %-ban kijelentések, 15 %-ban kiegészítendő kérdések, 5--5 %-ban felszólítások és felkiáltások. Az eredeti felszólítások hanglejtéseinek nyelvi tartalmai 70 %-ban kijelentések, 15 %-ban felkiáltások, 10 %-ban felszólítások és 5 %-ban kiegészítendő kérdések. A felkiáltások dallamaira kapott behelyettesítések: 35 %-ban kiegészítendő kérdések, 33 %-ban felkiáltások, 20 %-ban kijelentések és 12 %-ban felszólítások.

A kísérleti személyek válaszai alapján megnéztük azokat a struktúrákat, amelyek az észlelésben dominánsan ugyanannak a nyelvi tartalomnak a formai megoldását sugallták. A 32. ábrán a 20 %-nál nagyobb részesedésű ítéletek nyelvi tartalmai és az objektív akusztikai paraméterek alapján ábrázolt dallamformák láthatók.

80--100 %-ban kiegészítendő kérdésnek érzékelik a szólameleji csúccsal rendelkező dallamformákat ereszkedő vagy emelkedő dallamvéggel.

Ítélt tartalom		Behelyettesített tartalom			
		kiegészítendő kérdés	kijelentés	felkialtás	felszólítás
80-100 %	kieg. kérdés				
	kijelentés				
60-79 %	kieg. kérdés				
	kijelentés				
	felkialtás				
40-59 %	kieg. kérdés				
	kijelentés				
	felkialtás				
20-39 %	kieg. kérdés				
	kijelentés				
	felkialtás				
	felszólítás				

32. ábra

A dallamformák eredeti és ítélt tartalmai a
válaszok arányában

Teljes időtartamuk: 860--1220 ms. Hangterjedelmük: 170--320 Hz, a hangköz 150 Hz. A szólemeleji csúcs: 220-280 ms alatt 100--120 Hz-et ereszkedik (1/7, 1/9, 6/7). 80--100 %-ban kijelentésként azonosították az 1/6-os dallamot; paramétereit: 1500 ms időtartam, 170--250 Hz-es hangterjedelem, 80 Hz-es hangköz; 50 ms alatti 40 Hz-es ereszkedéssel kezdődik, majd ugyanennyi idő alatt 70 Hz-et emelkedik, majd végig ereszkedő jellegű.

60--80 %-ban azonosítottak kijelentő dallamformát (2/2), ennek teljes ideje: 820 ms volt, hangterjedelme: 180--260 Hz, hangköze: 80 Hz; enyhe csúccsal indul, majd végig ereszkedik, a csúcának ereszkedése: 220 ms alatt 80 Hz. Ugyanilyen százalékos arányban ítélték az eredetileg felszólító tartalmúakat kijelentésnek, illetve a felkiáltókat kijelentésnek és felkiáltásnak. A kijelentőnek ítélt dallamok paramétereit: teljes időtartamuk: 940--1600 ms, hangterjedelmük: 170--280 Hz, a hangköz: 110 Hz; a dallamstruktúra a kezdeti kis csúcsok után végig ereszkedő (2/3, 3/4, 5/2, 5/4). A felkiáltásnak ítélt dallamforma: 1850 ms időtartamú, hangterjedelme: 150--380 Hz, hangköze: 230 Hz, a kezdő csúcsot 120 ms alatt 130 Hz-es szökés, majd 280 ms alatti 130 Hz-es ereszkedés jellemez (1/1). Kiegészítendő kérdésnek ítélték az eredetileg felkiáltó tartalmú 5/1-t, szólemeleji csúcsa meredeken eső.

40--60 %-ban kiegészítendő kérdésnek ítélték az eredetileg is kiegészítendő kérdéseket (1/2, 1/4), valamint a felszólító és felkiáltó tartalmúak közül kettőt (1/10, 1/1). Jellemzőik: 590--3400 ms-os időtartamúak, hangterjedelmük: 150-380 Hz, a hangköz 230 Hz; a szólemeleji csúcsok 90 ms alatt 40 és 130 Hz-et emelkednek (szöknek), illetve 120 és 200 ms alatt 130 és 180 Hz-et szöknek (emelkednek). Majd 190, 200 ms alatt 170, 130 Hz-et ereszkednek, illetve 280, 430 ms alatt 130, 180 Hz-et ereszkednek. Egy dallamot (2/1) mind kijelentésnek, mind felkiáltásnak jelöltek: 500 ms időtartamú, hangterjedelme 170--250 Hz, a hangköz: 80 Hz; eső--ereszkedő dallamív jellemzi. Paramétereit alapján a magas százalékban kijelentésként érzékelt dallamstruktúrák közé illik; a kisebb százalékarány nyilván a szólemeleji 80 ms alatti 80 Hz-es esés következménye.

Hangsúlyészlelés

A hangsúly vizsgálatok arra keresünk választ, hogy milyen akusztikai tényező(k) következtében lép föl az a hatás a percepcióban, amely

hangsúlyélményként tudatosodik a hallgatóban. A beszédakusztikai vizsgálatokból előre jósolható volt, hogy az egyes nyelvekben más-más akusztikai paraméterek fognak hangsúlyélményt előidézni. Egyértelműen az észlelési kísérletekkel határozhatjuk meg ezeknek a paramétereknek a részvételét, valamint hangsúlykeltő szerepük kritériumait.

A magyar hangsúly elemzésekor Fónagy Iván kísérletileg igazolta, hogy a nyomatéki hangsúly, azaz a hangerőnövekedés (az intenzitás szint emelkedése) mellett más akusztikai jelenségek is velejárái a hangsúlyélmények (1958, 16). Egy összehasonlító percepciós kísérlet tanúsága szerint az orosz anyanyelvűek hangsúlyként elsősorban az időtartamot és a hangmagasságot észlelik, míg a magyar anyanyelvűek számára a hangmagasságnak és a hangnyomásnak, illetőleg a hangmagasságnak és/vagy a főhangsúlyosnak ítélt szót megelőző szünetnek a szerepe volt döntő (Gósy 1980, 104). Angol anyanyelvűekkel végzett percepciós vizsgálatok eredményei szerint a hangsúly érzékelésében a hangerőnek, az időtartamnak és a hangmagasságnak volt szerepe (Lehiste 1970, 126--139).

A hangsúlyészlelés, illetőleg a hangsúly- és a dallamészlelés összefüggéseinek vizsgálatára kísérletsorozatot terveztünk.

Anyag és módszer

Férfi bemondóval 6 mondatot rögzítettünk normál tempóban magnetofonszalagra. Ezek egyetlen mondat dallamformában és nyomatékeloszlásban különböző variációi voltak. Az első három mondatban mindig más-más szón jelent meg a főhangsúly; a második három mondat a kijelentés, az eldöntendő kérdés és a befejezetlen kérdés hangletjtésformáját reprezentálta. A mondatok a következők voltak (a hangsúlyt a megfelelő szótag aláhúzásával jelöltük):

1. A városban harangoznak délben. 2. A városban délben harangoznak. 3. A városban délben harangoznak. 4. A városban délben harangoznak. 5. A városban délben harangoznak? 6. A városban délben harangoznak...

E 6 mondatot elektroglottográf alkalmazásával egyidejűleg a természetes elhangzás mellett csak az alaphangmagasság változásait magnetofonszalagra rögzítettük. A normál elhangzású mondatokat imitátorral imitáltuk. Ily módon 12 dallamot kaptunk, amelyek közül a 2-2 azonos dallaformájú között a különbség az intenzitásviszonyok meglétében vagy hiányában volt. A glottográfál készült dallamok csak az alaphangmagasság változásait, az imitátorral készültek pedig mind az alaphangmagasság, mind az intenzitás változásait tartalmazták. A teszteléshez összeállított anyagot 10 (nyelvész) kísérleti sze-

mélynek játszottuk le. Feladatuk: a) a hallott dallamok ábrázolása, b) a fő- és mellékhangsúlyok bejelölése volt az általuk rajzolt dallamvonal megfelelő helyének aláhúzásával és c) a természetes hangzású mondatokban a hangsúlyhelyek bejelölése volt. A kapott dallamformákat alaphangmagasság- és intenzitásmérővel elemeztük.

Arra kerestünk választ: mi a percepciós viszony a dallam és a nyomaték érzékelésében az intenzitás paraméterének megléte, illetve hiánya esetén. Hogyan korrelál mindez a természetes mondatokról alkotott ítélettel?

Eredmények

A dallam ábrázolást az előző kísérlethez hasonlóan négyvonalas rendszerben kértük, amely három hangfekvést határozott meg. A vonalak objektív frekvenciaértékei: 200, 250, 300, 350 Hz. A kísérleti személyek rajzai mind időtartam, mind hangterjedelem, hangfekvés és hangmagasság tekintetében megközelítik az objektív paramétereket. Az intenzitáskomponensektől függetlenül 95 %-ban jól észlelték a hallott dallamstruktúrát. 87 %-ban különbségeket láttunk a glottográfval, illetve az imitátorral készült dallamok rajzában. (A továbbiakban ezt a 87 %-ot tekintjük 100 %-nak.) Kísérleti személyeink azokat a dallamokat, ahol intenzitáskomponensek is szerepeltek 72 %-ban -- az objektív értékekhez viszonyítva -- pontosabban és magasabb tartományban ábrázolták. A hangmenet tekintetében egyformának, de az imitált dallamokat 13 %-ban magasabb tartományban hallották. A természetes beszédhez közelebb álló -- intenzitáskomponenseket is tartalmazó -- dallamokat tehát a kísérleti személyek pontosabban képesek ábrázolni. (A hangfekvésben jelentkező különbséget okozhatja az is, hogy az imitált dallamok felvételi szintje 2-10 dB-el erősebb volt, s ez kelthette a magassági eltérés benyomását.)

A hangsúlyok jelölése lényegében ugyanazt mutatja, mint amit a dallamok rajzában tapasztaltunk. Az imitált dallamformákra sokkal egyöntetűbben jelölik a hangsúlyokat, mint a csak alaphangmagasságot tartalmazókba. Nagyobb mértékben jelölnek az előbbibe hangsúlyhelyeket, százalékosan kifejezve: a csak alaphangmagasság-változást tartalmazó dallamformák esetében 70 %-ban jelölnek főhangsúlyt, 30 %-ban semmit vagy csak mellékhangsúlyt. Az imitált dallamoknál 85 %-ban főhangsúlyt, 12 %-ban mellékhangsúlyt és 3 %-ban nem jelölnek hangsúlyt.

A főhangsúly helyének meghatározása mindkét módszerrel készült dallamformánál egybeesik, de a glottográfus dallamok esetében lényegesen kisebb a

százalékos részesedés (vö. 36. táblázat). Ez azt jelenti, hogy -- nem függetlenül ugyan a mondatstruktúrától --, de kevesebb az egyöntetű hangsúly-ítélet a glottográfus szólamoknál.

36. táblázat

A főhangsúly észlelése ugyanazon helyen

Dallamtípus sorszáma	Glottográfus (csak dallam)	Imitált (dallam+intenzitás)
1.	80 %	90 %
2.	70 %	80 %
3.	30 %	20 %
4.	20 %	70 %
5.	60 %	80 %
6.	20 %	50 %
Átlagosan:	46,6 %	65 %

Az objektív adatok birtokában megnéztük, hogy milyen fizikai paramétert érzékelnek a kísérleti személyek hangsúlyként. A glottográfákkal készült dallamoknál természetesen a frekvenciacsúcsokat. Ennek megfelelően az 1. dallam csúcsát -- amely 250 Hz-ről 80 ms alatt 340 Hz-re szökik, majd innen 130 ms alatt 80 Hz-et ereszkedik -- 80 %-ban; a 2. dallamét -- 250 Hz-ről 320 Hz-re szökik 70 ms alatt, majd innen 90 Hz-et ereszkedik 190 ms alatt -- 70 %-ban; az 5. dallamét -- 270 Hz-ről 340 Hz-re szökik 70 ms alatt, majd 110 ms alatt innen 250 Hz-re ereszkedik vissza -- 60 %-ban; a 3. dallamét -- 250 Hz-ről 320 Hz-re emelkedik 100 ms alatt, majd 200 ms alatt ereszkedik vissza -- 30 %-ban; a 4. dallamét -- 250 Hz-ről 300 Hz-re emelkedik 130 ms alatt, majd 40 Hz-et ereszkedik 190 ms alatt -- 20 %-ban; a 6. dallamét -- 230 Hz-ről 320 Hz-re emelkedik 240 ms alatt, majd 150 ms alatt ereszkedik 260 Hz-re -- 10 %-ban érzékelték főhangsúlyosnak.

Az imitált dallamoknál a magasabb frekvenciaérték és az intenzitás együttesen kelt hangsúlyélményt, de az is előfordul, hogy csak a csúcs, tehát a frekvenciaemelkedés dominál; 20 %-ban eredményezett a nagyobb intenzitásérték hangsúlyélményt. Az intenzitás fokozott szerepe a hangsúlyélmény kialakításában jól lemérhető a 3. és 4. dallamokra adott válaszok összehasonlításában. A 3. dallamnál -- lévén, hogy két kis csúcs is található a struktúrában -- mind a glottográfus, mind az imitált dallamra adott vála-

szok bizonytalanok. A csak alaphangmagasságot tartalmazó dallam esetében 30 %-ban a magasabb (második) csúcst ítélik a főhangsúly helyéül. Az imitált dallamban két főhangsúlyt jelölnek szintén 30 %-ban: az egyik helye az első kisebb csúcs, amelynek 32 dB az intenzitása, a második a dallam közepe táján található, 260 Hz frekvenciájú és 28 dB intenzitású. A 4. dallam esetében a glottográfusra adott hangsúlyítéletek ugyancsak bizonytalanok -- az igen kicsi, csak 50 Hz-nyit emelkedő csúcs miatt. Ez önmagában csak 20 %-ban érzékelődik hangsúlyként. Ugyanez a forma az imitált esetében már 70 %-ban ad hangsúlyélményt ugyanazon a helyen, mivel az 50 Hz-nyi dallamemelkedést 29-28 dB-es intenzitásérték is "támogatja". A 6. dallamforma esetében a glottográfus dallamban 20 %-ban a szólamvégi emelkedést ítélték hangsúlyosnak, az imitált formában 50 %-ban viszont a szólemeleji csúcst, mivel ennek intenzitásértéke: 30 dB, míg a szólamvégi emelkedése csak 24 dB. A normál elhangzásban bejelölt hangsúlyokkal az összevetést a 37. táblázat tartalmazza.

37. táblázat

Mondat sorszám	A főhangsúly-ítélet ugyanazon helyen	
	természetes elhangzás	imitált dallam
1.	100 %	90 %
2.	100 %	80 %
3.	100 %	20 %
4.	80 %	70 %
5.	60 %	80 %
6.	70 %	50 %
Átlagosan:	85 %	65 %

A glottográfus, az imitált és a normál elhangzású dallamokra adott hangsúlyítéletek az 1. és 2. mondat esetében azonos helyre kerültek, csak százalékarányban van köztük különbség. A 3.-nál normál elhangzásban 80 %-ban nem jelölnek főhangsúlyt, 10--10 %-ban pedig a szólam közepét, illetve végét érzik hangsúlyosnak. Az 5. mondat esetében a glottográfusnál és az imitáltnál a szólam végét jelölik hangsúlyosnak, normál elhangzásban azonban 60 %-ban a szólam közepét, 20 %-ban a végét, 10--10 %-ban egyéb részét.

Ennek az a feltételezett oka, hogy normál elhangzásban a mondat logikai főhangsúlya nem esett egybe e hanglejtéstípus kötelezően az utolsó előtti szótagon jelentkező frekvenciaemelkedésével. A 6. mondat esetében a normál elhangzásra adott ítéletek hasonlóak ahhoz, amit a csak alaphangmagasságot tartalmazóra adtak a kísérleti személyek: 70 %-ban a szólamvéget tekintették hangsúlyosnak.

Összefoglalás. -- Megállapítottuk, hogy a magyar hanglejtések vizsgálata ezzel a módszerrel értékes eredményeket ad. A kísérleti személyek képesek a hallott dallamforma teljes egészének, illetve bizonyos részek időarányainak helyes érzékelésére: a szubjektív ítéletek arányosak az objektív, műszeres mérések adataival. A hangmenetben azonos, de időtartamban és frekvenciaértékben különböző csúcsok ábrázolásakor szubjektíve a hangmagasságbeli változásoknak tulajdonítottak jelentőséget. A dallamok hangfekvései szintén az objektív adatokat közelítik, a hangmagasság és a hangfekvés jellemzőit nem cserélték fel. Nem kaptunk pontos felvilágosítást arra vonatkozóan, mi lehet az oka a szólamvégi frekvenciaemelkedések bizonytalan és pontatlan jelölésének, noha ezek gyakran 50--80 Hz-nyi változást mutattak. A hangmenetek percepciójáról is általános megállapításunk, hogy az objektíve hasonló struktúrákat hasonlóképpen ábrázolták. A szólameleji csúcsok emelkedő-szökő ágát 50 %-ban észlelték; fontos tényező a frekvenciaváltozás mértéke: a nagyobb frekvenciaértékű és meredekebb dallamívet nagyobb százalékban észlelik. Az ereszkedő, sikló és lebegő hangmenetek jelölése 50--80 %-ban az eredetinek megfelelő, nagyobb mértékben az ereszkedő és lebegő, kisebb mértékben a sikló dallamívet azonosították pontosan. Utóbbit 70 %-uk emelkedésnek minősítette: nyelvi értelmezésben tehát az emelkedésnek (szökésnek) van relevanciája.

A forma és a tartalom kapcsolatát, tehát a dallamstruktúra és a behelyettesített tartalom, valamint az eredeti mondat viszonyát több oldalról is megnéztük. Úgy találtuk, hogy a szólameleji csúccsal rendelkező dallamformákat elsősorban kiegészítendő kérdésnek és kijelentésnek ítélik -- a frekvenciaesés mértékének függvényében. Ha a szólameleji csúcs frekvenciacsökkenése 50--80 Hz körül mozog 80--220 ms alatt: nagymértékben kijelentésnek érzékelik; ha ez a változás 130--180 Hz 150--520 ms alatti, akkor a dallamforma már a kérdés és/vagy felkiáltás/felszólítás nyelvi jegyeit viseli magán. A magas frekvenciaérték és a nagy időbeli lefutás (pl. 380 Hz-ről történő ereszkedés és 280 ms) expresszív hatást kelt, a kísérleti személyek felkiáltást azonosítottak. A rövidebb időtartam az értelmi jellegű behelyettesíté-

sek arányát növeli, jelen esetben a kiegészítendő kérdését.

Az akusztikai adatok és a válaszok összehasonlításából kiderült, hogy a 230 Hz hangközű struktúrák nyelvi értelmezései megoszlanak, nagyobb mértékben érzelmi jellegűek, mint a kisebb: 100 és 160 Hz-es hangközű dallamoké. A hanglejtésformák második részében elhelyezkedő csúcspot minden kísérleti személy érzékelte, a behelyettesítésekben kivétel nélkül eldöntendő kérdés szerepelt. A szólamvégi hangmagasságemelkedést tartalmazó dallamok nyelvi értelmezései kisebb-nagyobb bizonytalanságot mutattak. A kísérleti személyek 28 %-a idegennek minősítette azokat, amelyeknél a szólamvégi frekvenciaemelkedést semmilyen szólameleji csúcs nem "egészítette" ki. (Mindkét kérdésfajta előfordult, de kijelentés, felkiáltás és befejezetlen mondat is.) Az eldöntendő kérdésekhez hasonlóan egyöntetű -- s az eredeti tartalommal is szoros kapcsolatú -- az ún. sikló dallamok megítélése. 92 %-ban egy- és kétszótagú kérdéseknek minősítették ezeket.

A hangsúlyok jelölése egyértelműen az imitált dallamoknál pontosabb és egyöntetűbb. A csak alaphangmagasság-változást tartalmazó dallamokban átlagosan 46,6 %-ban, az intenzitást is tartalmazókban 65 %-ban jelölnek főhangsúlyt azonos helyen. A kísérleti személyek hangsúlyként érzékelték a frekvenciaemelkedést, illetve az imitált dallamoknál a frekvenciaemelkedést és az intenzitást. Mindössze 20 %-ban jelöltek főhangsúlyt csak az intenzitásérték alapján. A természetes elhangzású mondatokban a hangsúlyjelölések még egyöntetűbbek.

A BESZÉDÉSZLELÉS FEJLŐDÉSE

Bevezetés

A gyermek beszédelsajátítása folyamán szoros kapcsolatban fejlődik az ún. artikulációs és a percepciós bázis. Amikor a beszédészlelés kialakulásának folyamatát vizsgáljuk, lényegében két egymással összefüggő működést elemzünk: az egyik a hallási rendszer, a másik a beszéd tulajdonképpeni azonosítása, felismerése. A gyermekkor, különösen a korai időszak beszédészlelési sajátosságairól viszonylag keveset tudunk. Könnyebb ugyanis megfigyelni és leírni azt, hogy mikor mit és hogyan mond a gyermek, mint pontosan meghatározni azt, hogy mikor mit és hogyan észlel, azonosít, dekódol. Ez utóbbi tekintetében sokszor hipotézisekre, a gyermek beszédprodukciója alapján vizs-
szakövetkeztetésekre és nem mindig hiteles kísérleti eredményekre vagyunk utalva.

A hallás általában és a beszédhallás, valamint ennek nyomán a beszéd megértése nem azonos folyamatok. A megszületés pillanatában a hallási rendszer alkalmas hangjelenségek felfogására, hiszen az újszülött hallószerve morfológiailag kialakult (Selnes--Whitaker 1976). Az újszülöttnak a felnőttével közel azonos méretű "tympanum"-ja és jólformált cochleája van (Fourcin 1978, 48). A hallási rendszernek a beszédmegértéshez szükséges működése azonban relatíve hosszú folyamat, tanulás, fejlődés eredménye. Ezalatt a különböző ingerek feldolgozása egyre pontosabbá, részletezőbbé válik, alkalmassá a beszédhangjelenségek elemzésére és szintézisére. A folyamat az anyanyelv akusztikai szerkezetének dekódolásával megy végbe.

A gyermek hallásérzékenysége kb. a 14 éves kor körül éri el azt a szintet, amely majd a felnőtt hallásképességét jellemzi. A kutatási eredmények alapján feltételezik, hogy a megfelelő idegpályák és agyi centrumok fejlődése, a bevéődés például, az életkorral nagyjából lineárisan halad előre (Selnes--Whitaker 1976). Saját beszédaudiometriás eredményeink is alátámasztották, hogy a 3--8 éves gyermekek beszédfelismerése életkoruktól függően változik (Farkas--Gósy--Hirschberg 1983). 2175 normál hallású 5--14 éves amerikai gyermek mindkét fülének hallásérzékenységét vizsgálták szinuszos hanggal négy különböző decibelértéken és hét frekvencián (250, 500, 1000, 2000, 4000, 6000 és 8000 Hz-en). Az eredmények a következők:

- a) a jobb és a bal fül nagyjából egyformán érzékeli a szinuszos hangot a különböző decibel- és frekvenciaértékeken,
- b) a frekvenciaérték növekedésével csökken a hallásérzékenység: minél magasabb a hang, annál erősebben kell hangoznia, hogy észlelhető legyen,
- c) elsősorban 6000 Hz-nél, kisebb mértékben 4000 Hz-nél kevésbé érzékeny a fül a szinuszos hang észlelésére,
- d) a hallásérzékenység az életkor függvénye: 5 éves kortól egyre jobban az eredmény, főként a 250--4000 Hz-es tartományban. 12--13 éves korban kiemelkedően jók az eredmények, ekkor a legérzékenyebb a hallási rendszer; ehhez képest némi visszaesés tapasztalható 14 éves korban. A 12--13 évesek hallásképességének különös érzékenységére egyelőre nincs tudományosan megalapozott magyarázat. A hallási mechanizmusnak a felsorolt sajátosságai a beszédhallás kutatása számára két ponton különösen fontosak:

1. a hallásképesség az életkor függvénye, az életkor növekedésével párhuzamosan fejlődik;
2. a hallásképesség fejlődése legnagyobb mértékben azt a frekvenciatartományt érinti, amelyben a beszéd akusztikailag döntően megvalósul, vagyis a 100--4000 Hz-es tartományt.

A percepció bázis kialakulása, fejlődési menete az egyes gyermekek között nagy eltérésekkel valósulhat meg, egy-egy életkori szakaszban a beszédmegértés komoly különbségeket mutathat. Közismert, hogy az a gyermek, akihez sokat beszélnek, akit sokat beszéltetnek, figyelemmel hallgatnak, az jobban képes kifejezni magát, nagyobb a szókincse, előbb tud grammatikailag bonyolultabb mondatokat formálni, artikulációja rendszerint pontosabb. Kevésbé köztudott -- hiszen csak közvetve tapasztalható --, hogy ez a gyermek a beszédmegértés tekintetében is jóval fejlettebb.

Az elmúlt 10 évben a gyermekek beszédmegértésének tanulmányozása nagy gyorsasággal fejlődött, amit számos kutatási eredmény fémjeléz (elsősorban angol, amerikai nyelvterületen). A kutatások Élő összegzésében a következő témákra oszlanak: a beszédmegértés fejlődésében bekövetkező változás folyamata; a nyelvi tapasztalat szerepe a fejlődésben; a gyermek percepció "képességének" eredete; e "képességek" filogenetikai jelentése; a csecsemő-kori beszédpercepció kapcsolata a későbbi gyermekkor beszédmegértésével (1980, 23). Hogyan alakul ki és hogyan fejlődik tehát a beszéd megértése, amely a hallás és ezen belül a beszédhallás fejlődésére épül rá? Az általános törvényszerűségek a következők:

A) Verbális ingerek megléte nélkül a beszédmegértés folyamata nem indulhat meg (függetlenül attól, hogy emogott miféle ok rejtőzik, pl. inger-szegény környezet, hallásproblémák stb.).

B) A beszédmegértés és a beszédprodukciónak kölcsönhatásban fejlődik, bármelyik zavara hatással van a másik folyamatra.

C) A beszéd megértése -- a fejlődés korai szakaszaiban -- nagymértékben, később kisebb mértékben, de felülmúlja a beszédprodukciónak.

D) A beszédmegértést hosszú időn keresztül extra- és paralingvisztikai tényezők segítik; ezek funkcióját fokozatosan nyelvi formák veszik át.

Megválaszolandó kérdés, hogy a beszédfelismerés minek alapján történik:

a) az elsajátított azonosság vagy b) az elsajátított megkülönböztetés alapján (Hörmann 1971, 63). Az első esetben a gyermeknek meg kell tanulnia azt, hogy ne vegye figyelembe azokat a különbségeket, amelyek a kommunikáció szempontjából nem fontosak. (Például felismerje az [a:] vagy a [s] hangot akár férfi, akár nő mondja, akár öreg, akár fiatal, bármilyen érzelmi állapotban stb.) A második esetben meg kell tanulnia élesebben megkülönböztetni azt, ami a beszédértés szempontjából fontos. (Ha például az adott nyelv megkülönbözteti a zöngés és zöngétlen ajakzárhangot, vagy az interdentális és a dentális réshangokat.

A beszédészlelés kezdetei

A beszédészlelés (-megértés) kezdetei a születéshez nyúlnak vissza. Noha az első év beszédpercepciójáról szóló ismereteink, az elvégzett kísérletek igen fontosak, elsősorban pszichoakusztikai jellegű adatokat közölnek, ezek pedig áttételesen kapcsolódnak a beszéd tulajdonképpeni megértéséhez.

A születéstől a 18 hónapos korig terjedő időszak beszédpercepció vizsgálatában három módszer ismert általánosan: a) a szopási reflex felhasználása, b) a szívdobogás ütemének felhasználása és c) a vizuálisan megerősített válaszreakció (Eilers 1980, 23-24). A csecsemő diszkriminációs képességének vizsgálatában a három módszert az életkortól függően alkalmazzák: az elsőt az első 3--4 hónapban, a másodikat az 1 és 6--8 hónapos kor, a harmadikat a 6 és 18 hónapos kor között. A módszerek alkalmasak arra, hogy megtudjuk, az adott életkori szakaszban miként, mennyire képes a gyermek a különféle beszédrészeket megkülönböztetni (például a zöngékezdesi idő alapján a zárhangokat, a különböző nyelvállású magánhangzókat stb.).

Az újszülött két hetes korában már képes az emberi hangot más

hangoktól -- például dobpergés, csengő, egyéb zajok --, illetve az emberi beszédet más "emberi" hangmegnyilvánulásoktól elkülöníteni (vö. Clark--Clark 1977, 376). (Közismert, hogy a női, elsősorban az édesanya hangja megnyugtató hatással van a síró újszülöttre.) Általánosan elfogadott nézet, hogy a csecsemő a szupraszegmentális jelenségekre (dallam, hangsúly, tempó, hangerő, ritmus, hangszínezet) reagál először; ez teszi lehetővé számára például a megszokott (hallott) anyanyelvtől eltérő idegen nyelvi hangzás felismerését. Kéthónapos csecsemők már az emberi hang emocionális változásait is érzékelik (Clark--Clark i.m. 378).

Az újabb kísérleti eredmények szerint a szupraszegmentális jelenségek felismerése a beszédprodukciónak fejlődésében nem elsődleges. Mind a szegmentális, mind a szupraszegmentális szint percepciója egyformán korai (Clark--Clark i.m. 380--4). Más nézet szerint nem is az a kérdés, hogy mi az elsődleges a felismerésben, hanem az, hogy milyen viszony van a szegmentális és a szupraszegmentális információ között a csecsemő, illetve a kisgyermek számára (Eilers 1980, 35). Feltehetően a csecsemő mindkét szint -- korlátozott - megkülönböztetésére és azonosítására egyidejűleg képes; tény azonban, hogy a beszédmegértés fejlődésében a szupraszegmentumok felismerése megelőzi a szegmentumokét, tehát a beszédhangokét, hangkapcsolatokét. Ez részben a csecsemő összetett hallási feldolgozásának eredménye; valamint annak, hogy a gyermek számára szükséges "megértés" elsősorban a szupraszegmentumokban jelentkező információ felismerését teszi szükségessé. Másrészt a szupraszegmentális tényezők elemszáma a szegmentumokénál jóval kisebb, sokkal kevésbé változnak az időkben, s ez lehetővé teszi, hogy a csecsemő a mondhatni "durvább", nagyobb különbségeket könnyebben felismerje. A dallam, nyomaték, hangerő, stb. megkülönböztetése mind pszichoakusztikai folyamatában, mind a ráépülő nyelvi értelmezésben egyszerűbb feladat. Egy- és kéthónapos csecsemőkkel végzett kísérletek szerint a gyermekek képesek egészen finom hangmegkülönböztetésekre, mint amilyen a zöngés és a zöngétlen hangok elkülönítése vagy a mássalhangzó képzési helyének megkülönböztetése. Amerikai kutatók azt találták, hogy a fiatal csecsemők a ba/pa szekvenciákat jól el tudják különíteni a zárhangok zöngékezdési idejétől (VOT) függően (Eimas--Siqueland--Jusczyk--Vigorito 1971, 303--6; Clark--Clark 1977, 377; Villiers--Villiers 1979, 25). A különböző nyelvekben meglévő bilabiális felpattanó zárhang-fonémáknak -- amelyek egymástól a zöngesség jegeiben térnek csak el -- a hangrealizációi akusztikailag különfélék. Ez a különbözőség főként a zöngékezdési időtartamban ragadható meg. Az angol [b]

a zéró és a mínusz értékű, a [p] pedig a plusz 25--40 ms értékű zöngেকেzdés-sel jellemezhető. A libanoni arabban a -70 ms a zöngés és a +5/15 ms a zöngétlen hang zöngেকেzdési (VOT) ideje (vö. McNeill 1979, 138). Egy kenyai bantu nyelvben, a kikuyuban az egyetlen ajakmássalhangzó megfelelő időtartama - 60 ms (Pauka 1980, 119). Az idézett mássalhangzók zöngেকেzdési idejének nyelvenkénti különbözőségétől függetlenül képesek a csecsemők bizonyos differenciálásra, de nem minden zöngés/zöngétlen zárhang megkülönböztetését tudják elvégezni. Eimas kísérletében a kis csecsemők az angol [b] és [p] zöngেকেzdési idejét elkülönítették, de a spanyol explozívákét például nem (Eimas et al. 1971). Ezek a kísérleti eredmények ahhoz a következtetéshez vezettek, hogy a csecsemő beszédpercepciója veleszületett, ösztönös képesség. A hallási rendszer eredendően alkalmas bizonyos különbségek észlelésére; ez az érzékenység a zöngেকেzdési időt tekintve feltehetően éppen azt a tartományt érinti, amelyben az angol nyelv ajakzárhangjai megvalósulnak. A megkülönböztetési képesség bizonyos auditorikus tapasztalaton is alapszik, de a nyelv i tapasztalattól teljesen független neurális rendszer működési eredménye. (Meg kell továbbá azt is jegyeznünk, hogy nem minden vizsgálat zárult pozitív, illetve értékelhető eredménnyel az újszülöttek és a csecsemők hangmegkülönböztetési képességének sajátosságait illetően.)

Természetes beszéddel, CV és VC szekvenciákkal vizsgálták 30 egészséges 9--19 hetes csecsemő beszédhang-differenciáló teljesítményét (Eilers 1977). Arra kerestek választ, hogy az adott fonetikai helyzetben mi ad(hat)ja a "kulcsot" a zöngéesség tekintetében különböző dentális zárhangok és résmássalhangzók ([d]/[t], illetve [s]/[z]). Az eredmények szerint a gyermek a felpattanó zárhangokat legnagyobb mértékben a magánhangzó időtartamának függvényében, kisebb mértékben a fonetikai helyzettől függően ismerték fel (ez utóbbi különösen a CV szekvenciákban volt döntő). Ugyanakkor a résmássalhangzók zöngéességének, illetve zöngétlenségének a megítélése jóval több tényező függvénye, úgymint: a magánhangzó időtartama, a r és időtartama, a formánsátmenet és a zöngéesség "foka".

Meglepőek azok az eredmények, amelyek szerint a beszédhangok felismerésében a két agyfélteke már csecsemőkorban különböző feladatot lát el (vö. 29. oldal).

Sok kísérlet arra irányult, miként ismerik fel a csecsemők a mássalhangzó artikulációs helyét (Clark--Clark 1977, 378). Williams és Bush például 6--12 hetes csecsemőkkel végzett vizsgálatának eredményei szerint, a gyerme-

kek képesek voltak a formánsátmenet alapján megkülönböztetni a [da] szekvenciát a [ga]-tól (1978). Eredményeiket más kísérletek is alátámasztják, bár a döntés alapja hol a formánsátmenet, hol a zörejjóc helye volt. Érdekes viszont, hogy a fonetikai pozíció döntően befolyásolja a percepció eredményt: a csecsemők CV-hangkapcsolatban jól oldották meg a feladatot, míg VC helyzetben nem, intervokális helyzetben pedig bizonytalanul azonosították a kérdéses mássalhangzókat.

A magánhangzók felismerésével kapcsolatban kevesebb adat áll a rendelkezésre; az eddigi eredmények szerint azonosításuk jó (4--17 hetes csecsemőkkel végeztek kísérleteket). Más eredmények szerint a 6--8 hónapos csecsemők magánhangzó-felismerése az anyanyelvi környezettől teljesen független (Eilers 1980, 33--4).

A diszkriminációs képesség fejlődése a nagyobb eltérés felismerésétől a kisebb felé halad (Eilers--Wilson--Moore 1977). Különböző életkorú csecsemőkkel CV "kontrasztokat" teszteltek, mint [sa/va], [sa/za], [fa/θa] stb., illetőleg VC-ket, mint: [as/a:z], [at/a:t] stb. Az 1-3, 6-8, illetve 12--14 hónapos csecsemők azonosítási képessége életkoruknak megfelelően finomodik: [as/a:z], [a:s/a:z], [a:t/a:d], végül [a:t/at]. Egy másik kísérlet, amelyet különböző anyanyelvű gyermekekkel végeztek, arról győz meg, hogy "a tapasztalat központi szerepet játszik a gyermek beszédmegértésében" (Eilers--Gavin--Wilson 1980, 113); a hallási gyakorlat növekedésének fontosságát másik is hangsúlyozzák (pl. Fourcin 1978, 66).

Számos impozáns kísérlet történt egy-egy akusztikai komponens észlelésének vizsgálatára a csecsemőkorbán. Az eredmények szerint, a gyermekek az első életévben 4000 Hz alatt érzékenyebbek a szinuszos hangok változására, mint a fölött (Menyuk 1971, 39--40); az 500 Hz alatti frekvenciákra pedig különösen érzékenyek (Fourcin 1978). Ugyancsak szinuszos hangokkal vizsgálták az életkorral együtt változó észlelési érzékenységet. Erősségük 40--50 dB volt. A négyhónaposak a 800 Hz-es hangtól csak a 4000 Hz-éset, az öthónaposak a 800 Hz-estől a 3000 Hz-éset, a hathónaposak pedig már a 2000 Hz-éset is el tudták különíteni (Pauka 1980, 120). Az emelkedő és az eső dallam közti különbséget már kéthónaposok is észlelik (a kérdéses dallamot a [a] szekvencia hordozta, vö. Pauka i.m.); más kísérletek szerint viszont ez a képeség csak nyolchónapos korra alakul ki (Menyuk 1971, 59).

A szegmentális és a szupraszegmentális hangszerkezet percepciójának megismeréséhez adott újabb eredményt az a kísérlet, amelyben 1 és 4 hónapos csecsemőkkel vizsgálták az emelkedő-eső és a monoton dallamminta elkülöníté-

sét. A gyermekek jól végezték a megkülönböztetést, ha a kétféle dallamot ugyanaz a magánhangzó hordozta; tévedtek viszont, ha a magánhangzó-minőség változott (Kuhl 1976). Ez a kísérlet éppen azt mutatja, hogy az adott "szegmentális" hangszerkezetben meglévő információ fontosabb volt a gyermek számára, mint a "szupraszegmentális"-ban lévő. Más kísérlet is zárult olyan eredménnyel, hogy -- a feltételezéssel ellentétben -- a gyermekeknek nehézséget okozott bizonyos prozódiai jegyek megkülönböztetése: például kétszótagú hangsorban a hangsúlyos, illetőleg hangsúlytalan szótag felismerése. A kísérlet eredményeiből arra következtettek, hogy az intenzitáskülönbség -- feltehetően -- önmagában nem volt elegendő a hangsúlyosság megítéléséhez (Eilers 1980, 36).

Mindezek alapján legvalószínűbb, hogy a csecsemő már a kezdetektől a maga komplexitásában érzékeli az emberi beszédet. Hangmegkülönböztetésekre hallási mechanizmusa alkalmassá teszi. A már igen korán kiépülni kezdő nyelvi percepció azonban bizonyos értelemben egyszerűsíti a hallási működést: abban az értelemben, hogy az életkornak megfelelő beszédmegértés korlátozza a hallási feldolgozás részletességét a nyelvi értelmezés érdekében. Az értéshez a gyermek nem veheti figyelembe az abszolút fizikai értéket, hanem a viszonyokat, vagyis a hasonlóságokat kell megjegyeznie (Fourcin 1978, 54). Ha hitelesnek fogadjuk el a csecsemők hangdifferenciálói képességét célzó kísérletek eredményét, akkor azt mondhatjuk, hogy "hiába" képes a gyermek megkülönböztetésekre, mivel ezek a nyelvi funkció nélküli képességek mégsem teszik lehetővé számára az elhangzó beszéd megértését. Más szavakkal: nyelvi kompetenciájának éretlensége korlátozza hallási diszkriminációs képességét.

A különböző nyelveket anyanyelvként elsajátító csecsemők vokalizációját elemezve azt találták, hogy nagyjából és körülbelül hathónapos korra tehető bizonyos anyanyelvi sajátosságok jelentkezése (Gósy 1983). Ez egyértelműen azzal függ össze, hogy ebben az életkorban már beszélhetünk a szegmentális hangszerkezet bizonyos fokú percepciójáról, mégpedig nyelvi tartalmat hordozó megértéséről. Ezt akár a magyar szakirodalom korán, 7 hónaposan beszélni kezdő megfigyeltjei (Kenyeres 1926), s egy saját kísérletem is alátámasztja. Hathónapos fiam a hinta-palinta kifejezés dallammal összekapcsolt elhangzására minden esetben négykézlábra állt, és előre-hátra ringatta magát. Megpróbáltam csak a dallamot zümmögni, majd az ismerttől egészen eltérő hangsorokat mondtam a megfelelő dallammal. A négykézlábra állás és előre-hátra mozgás reakciója kivétel nélkül mindig bekövetkezett. Egy alka-

lommal teljesen monotonon mondtam a hinta-palinta hangsort; legnagyobb meg-
 lepetésemre a gyermek pontosan ugyanúgy reagált, mint az előzőekben. A reak-
 ció csupán egyetlen esetben nem következett be: amikor monoton hangon külön-
 böző szavakat mondtam, amelyek csak szótagszámban egyeztek meg az eredeti
 hangsorral. Ebben az életkorban a gyermek már "bonyolultabb" felismerésekre
 is képes, pl. az édesanyja "hello"-ját elkülöníti idegen férfi és nő "hello"-
 jától (Menyuk 1971, 43). Saját gyermekeimről hathónapos koruk után számos
 feljegyzésem van, amelyek mindegyike a szegmentális szerkezet dekódolásáról
 tanúskodik. Az azonban, hogy egy adott beszédhelyzetben melyiknek van első-
 sége a megértés számára, ti. a szegmentális vagy szuprasegmentális ténye-
 zőknek, az ekkor, de még jóval később is, elsősorban a beszédhelyzettől függ.
 Kissé eltér ettől Fourcin véleménye, aki szerint a tonális nyelveket elsajá-
 tító csecsemő az ún. tonémák alapján végzi a megkülönböztetéseket, míg a nem
 tonális nyelveket elsajátító az eloszlásukban különböző szegmentumok alapján
 (1978, 55).

A tulajdonképpeni beszédmegértés akkor kezdődik, ami-
kor a gyermek bizonyos jeleknek és ezen jelek legegyszerűbb viszonyának kon-
ventionális tartalmát elsajátította. Ez az anyanyelvi percepciós bázis kia-
lakulásának kezdete. Fontos kérdés a produkciós és a percepciós rendszer kap-
 csolata az egyes életkori (fejlődési) szakaszokban. Döntő ui. a beszédmozgá-
 sok során szerzett tapasztalat és ezek akusztikai következményeinek egyidejű
 megfigyelése a beszédpercepció számára. Pszichológiai megfigyelések szerint
 a fogalom jeltartalmának megértése előtt kialakul egy "beleérzés" a helyzet
 összértelmébe és hangulati tartalmába. Ez ősbibb, mint a beszédmegértés. Eb-
 ből alakul ki az ún. globális percepció, amellyel egyéves
kor táján, tehát nagyjából a holofrázisos időszakában jellemezhető a gyer-
meki beszédmegértés Lényege, hogy a hallási információ felfogása a maga e-
 gészében történik, amelybe a szegmentális, szuprasegmentális hangszerkeze-
 ten kívül az adott szituációra jellemző egyéb tényezők: az extra- és para-
 lingvisztikaiak tartoznak bele. Ez a komplex megértés azonban minőségileg
 (és mennyiségileg is) más, mint az idősebb csecsemők percepciója. Amikor a
Ne nyúlj a késhez! vagy a Vigyázz, mert le fogsz esni! közlések elhangzanak,
 a gyermek ebben csak a veszélyhelyzetet ismeri fel, s messze nem a mondat-
 struktúrát elemzi ki. A beszédmegértés finomodását, vagyis a szegmentális és
 szuprasegmentális szintek akusztikai paramétereinek összekapcsolódását a
 nyelvi funkcióval, nyelvi tartalommal, a gyermek beszédprodukciója egyre
 jobban tükrözi. Például, amikor motoros korlátainál fogva a hallott szót

ugyanúgy realizálni nem is tudja még, de a hallotthoz leghasonlóbbat kívánja ejteni, például a szótagszám egyeztetésére törekszik. Így lesz a cipő korábbi pő alakjából pőpő (vö. MacNeilage 1980, 19).

Az ún. telegrafikus beszéd időszakában, tehát kb. két éves kor körül ez a globális percepció tovább finomodik; mind kisebb szerepet kapnak a nem nyelvi tényezők, s egyre kisebb kontrasztokkal valósul meg a gyermeki beszéd felismerés. A gyermek tökéletesen elkülöníti anyanyelvét idegen nyelvektől; egyre jobban próbálkozik a hangsorok elemzésével. Hosszabb szekvenciából is tudatosan azonosítani igyekszik a számára ismertet. Nem beszédhangok megkülönböztetését és azonosítását végzi ugyan el, hanem a globális percepción belül stratégiákat alkalmaz a megértéshez. Nevezhetnénk kulcsszó-stratégiának is. A felismert és azonosított egy vagy két szemantikai egység alapján (amely a gyermek számára egyetlen komplex egység még akkor is, ha grammatikai értelemben két vagy több szóról van is szó, pl. nem szabad vagy menj innen) következteti ki az egész közlés tartalmát. Mindebben már a gondolkodásnak, az asszociációs mezők működésének nagy a szerepük, noha a beszédet kísérő beszéden kívüli jelenségek még gyakran meghatározóak maradnak. Saját kísérletem ezzel kapcsolatban a következő volt. Kétéves fiamnak hosszabb közlést mondtam angolul, amelyben egyetlen szó, a kulcsszó volt számára ismert, magyar szó: Give me the könyv, please! Az adott szituációban a gyermek a kérést végrehajtotta (a könyvet odaadta). Hasonló közléseket hallott német és orosz nyelven, a feladatot minden esetben teljesítette. A beszéd megértés mértéke pontosan követhető volt kisebbik fiam válaszaiból két éves kora táján. Ha nem értette a hozzá intézett közlést -- akár kérdés, akár felszólítás volt -- kivétel nélkül mindig a mondat utolsó részletét (rendszerint az utolsó szót) ismételte meg. Például: A Hímezzek neked egy terítőt? kérdésre a válasz: terítőt; a Megtanítsalak síelni? kérdésre a válasz: síelni. Idegen nyelven feltett, dallammenetében, hangsúlyozásában a magyar kérdésekre emlékeztető közlések esetében a reakció ugyanaz volt. Abban az esetben, ha a gyermek a közlés lényegét képes volt megérteni, az ismételt kifejezés mindig a közlés értelméhez igazodott. Például: Vegyem ki Petit az ágyból? Válasz: vegyem Petit; vagy: Petya, apa hol van? Válasz: apa (és a feljegyzések szerint keresgélni kezdett a szobában); stb.

A két év körüli gyermekek beszéd megértése három fokozatban fejlődik. Az elsőben -- azt mondhatjuk -- a gyermekek külön tárolják saját és külön a felnőttek ejtését, s mindig a megfelelőt hívják elő. Ez magyarázza azt, hogy pontosan tudják, hogy az ő ejtésükben realizált efelán (elefánt) vagy lula

(ruha) ugyanaz a szemantikai tartalom. Fiam a zokni szót sócsinak ejtette hosszabb időn keresztül. A kétféle tárolásra vonatkozólag többször a következő kísérletet végeztem el vele. A legkülönbözőbb beszédhelyzetekben saját közléseimben hol a normál hangsort, hol a gyermek "elferdített" sócsi hangsorát mondtam. A gyermek kivétel nélkül azonosan reagált, vagyis mindkét hangsort ugyanannak a szemantikai tartalomnak feleltette meg. A különbség csupán annyi volt, hogy néhány hónap elteltével csodálkozva nézett rám, amikor a saját ejtését hallotta, mintegy furcsálva a "gyermeknyelvi" szót a felnőtt közlésében. A kísérletet majd egy fél éven keresztül kb. 15 szóval folytattam, amelyek hangalakja kisebb vagy nagyobb mértékben eltért a felnőtt ejtéstől. Az eredmények mindig a leírtaknak feleltek meg. Gyermekek kiváló példát nyújtott arra, hogy ez a kétféle reprezentáció milyen erős ebben az életkori szakaszban. A kisfiú a televízió szónak a környezetében megszokott tévé megnevezését jól ismerte, saját produkciójában ez péhé-ként jelentkezett. Miután a gyermek beszédhangjai között (más szavakban) a tévé szó mindkét más-salhangzója megjelent már, megpróbáltam a gyermeknek a "felnőtt nyelvi" hangsort megtanítani. Nem túl hosszú játékos gyakorlás után (táhá, pává, tává, tóvó, tővő, tévé stb. hangsorok ismételtetésével, tudatosításával) a gyermek tökéletesen kimondta a tévé hangsort. Ugyanakkor a televízió megnevezése saját beszédprodukciójában még hosszú időn keresztül péhé maradt, noha az el-sajátíttatástól kezdve a tévé hangsort felszólításra bármikor fonetikailag hibátlanul visszhangozta. A gyermek tehát egyszersmind ragaszkodik is a kétféle reprezentációhoz, hiszen ha ez nem lenne, akkor a megértés érdekében a saját beszédprodukcióját a felnőttéhez hasonlóra kellene alakítania -- ami motoros korlátainál fogva nem is sikerülne neki.

A következő fokon ugyancsak kétféle reprezentáció létezik, a gyermek ejtése szerinti és a felnőtté, vagyis a környezet ejtésének megfelelő. A kettő viszonya azonban változáson megy át a korábbi fokozathoz képest. A gyermek minden esetben a felnőtt ejtése alapján végzi el a feldolgozást, a sajátját csak akkor hívja elő, ha az nem áll szemben a felnőttével. Más megfogalmazásban: a kétféle reprezentációt csak akkor felelteti meg a gyermek ugyanannak a szemantikai tartalomnak, ha a megfelelés egyértelmű, s a felnőtt nyelvi készletben más szemantikai tartalomnak saját ejtése nem feleltethető meg. Angol anyanyelvű gyermektől megkérdezték, hogy mi az a [sórt] (shirt 'esernyő'). A gyermek magától értetődően rámutatott a kérdéses tárgyra, noha ennek helyes ejtése: [sórt], ámde [sórt] hangalakú jelentés az angol nyelvben nincsen, tehát a kétféle reprezentáció csak egyetlen szemantikai tartal-

mat idézhet fel. Ugyanakkor, amikor ettől a kisgyerektől megkérdezték, hogy mi az a [sip] (sip 'kortyol'), csak az igét volt hajlandó felismerni és semmiképpen sem a hajó tárgyat, amelynek angol neve pedig, ti. [ʃip] ship az ő ejtésében az igéével azonosan, [sip]-nek hangzott (Clark--Clark 1977, 386). A jelenség jól ismert: a gyermek képes még méregbe is gurulni, s bizonygatja, hogy az "nem jigó, hanem jigó!" Természetesen a kontextus és a beszédhelyzet a legtöbb esetben egyértelművé teszi a hangsor és a szemantikai tartalom kérdését, de kísérletesen saját gyermekemmel is ki tudtam mutatni a felnőtt ejtés alapján történő megértést. A mássalhangzó időtartamának nagyjából hasonló ejtésével a száll/sál, illetve a magánhangzó időtartamának bizonytalan ejtésével az orrom/ollóm szavakat mondtam, s a gyermek mindig a felnőtt ejtés alapján ismerte fel a kérdéses jelentést.

A legfejlettebb fokon a gyermek közli, hogy az adott szót ő úgy nem tudja ejteni, sőt a "gyermeki" hangsor hallatán rendszerint meg sem kísérli a felismerést.

Az elmondottakból látható, hogy a globális percepció finomodása szűkülést is jelent, amely azonban nélkülözhetetlen a megértés további fejlődéséhez. Az első fokozatról a harmadikra történő fejlődés durván a két- és hároméves kor között zajlik, a hároméves gyermek anyanyelvi percepciósi bázisa alapfokon kialakult. A globális percepció jellemzője ekkor, hogy a beszédmegértés továbbra sem az egyes hangok megkülönböztetésén és azonosításán alapszik, hanem jelentések, szemantikai tartalmak különbözőségén. Természetesen a szegmentális hangszerkezet hordozza azt a jegyet, amelynek alapján lehetőségessé válik az elkülönítés. Például a kép és a gép szavak különbsége a gyermek számára nem a /k/ és a /g/ fonémák eltéréseiben jelentkezik, hanem a két alakzat a maga egészében, (elsődlegesen) a jelentés alapján különül el egymástól. A gyermek előbb képessé válik a hangok megkülönböztetésére, de ekkor még képtelen az azonosítására, ami a további fejlődés eredménye lesz (Clark--Clark 1977, 374). Orosz anyanyelvű gyermekekkel végzett kísérletek szerint a gyermek a fonémák hangrealizációinak különbözőségét csak akkor sajátítja el a megértés számára, ha ahhoz konkrét jelentésbeli különbség társul (Clark--Clark i.m.). A kísérletben részt vevő 2--3 évesek bizonyos zárhangokat csupán azután különböztettek meg a megértésben, miután saját beszédjükben megjelentek azok a hangsorok, amelyek úgy feleltek meg különböző szemantikai tartalmaknak, hogy egy-egy disztinktív jegyben különböztek egymástól. A kísérletben különböző játékgépekkel játszottak, amelyeknek a neve mindössze egyetlen disztinktív

jegyben tért el egymástól, például: tis és dis. A kísérlet eredményei szerint a gyermek valóban mindig a testesebb szegmentumokhoz rendeli a jelentést, és megfordítva: a kis akusztikai különbségeket csak akkor képes azonosítani, ha annak már -- saját nyelvi értelmezésében -- nyelvi szerepe, funkciója van (ld. erre még: Barton 1976, 232).

A beszédmegértés sajátosságairól más aspektusból nyújt információt a logatomok felismerése és utánmondása. 3;5--6 éves gyermekekkel végzett kísérletek szerint -- ellentétben a korábbi feltételezésekkel -- a fiatal gyermek is képes az értelem nélküli hangsorokat visszamondani. Normál hangerő esetén közel hasonlóan az értelmes szavakhoz. A hangerő csökkenésével az egy- és kétszótagú logatomokat a gyermekek lényegesen rosszabbul mondták vissza, mint az értelmes szavakat. A logatomok visszamondása azt jelzi mindenekelőtt, hogy a gyermek ismételni tudja beszédhangok sorozatát, vagyis képessé vált az anyanyelvére jellemző akusztikum felismerésére és azonosítására. Noha a globális percepció fejlettebb fokon 3--6 éves korban is működik, a gyermek beszédhallása már kellően fejletté vált a beszédhangok szegmentálására, pontosabban hangsorok disztinktív jegyeinek adott egymásutániségben való rögzítésére a rövid idejű memóriában.

A születésüktől fogva csökkent hallású gyermekek a logatomot csaknem ugyanúgy vagy teljesen ugyanolyan mértékben képesek ismételni, mint az értelmes szavakat, aminek elsősorban az a magyarázata, hogy beszédmegértésük és beszédprodukciónak is más szabályok szerint fejlődik, mint az éphallásúaké. Számukra különös jelentősége van annak, hogy a hallott jelsorozat akusztikai tartalmát feldolgozzák az első szinteken, s csak ezután következik a feldolgozás eredményének összekapcsolása a szemantikai tartalommal. Éphallóknál ez a folyamat relatíve "gyorsabb", s normál körülmények között nem szükséges a hallási információk részletekbe menő feldolgozása: a globális elemzés "egyszerre" foglalja magában az akusztikai és szemantikai jegyek feldolgozását. Éppen ezért mondják vissza a jelentés nélküli hangsorokat rosszabbul, mint az értelmes szavakat.

Körülbelül hat éves korára válik képessé a gyermek hangsorok szegmentálására, vagyis a hangsort alkotó kisebb szerveződési egységek, például magánhangzók, egyes mássalhangzók elkülönítésére. Jól ismert, hogy a Mondj egy állatot, amelyiknek a neve bé-vel kezdődik! feladatra a gyermek mindig a békát fogja mondani és sohasem bárányt vagy bikát vagy bolhát vagy baglyot. Egyrészt az életkor, másrészt az iskola, az írni, olvasni tanulás ténye okozza azt, hogy a gyermek beszédmegértése oly módon tudatosodik, hogy a

globális percepció az eddigi értelemben megszűnik (feltételezve természetesen, hogy a gyermek anyanyelvi szókincsének megfelelő hányadát, valamint a grammatikai szabályszerűségeket az életkorának megfelelő fokon már ismeri), a gyermek az elhangzó közléseket részleteiben is érti, alapjaiban kialakult a felnőttre jellemző beszédmegértési folyamat (Gósy 1984c). Az eltérés a felnőttétől természetesen még jelentős, sokkal nagyobb mértékű ekkor a beszédmegértési folyamat alsóbb szintjeinek részvétele a megértésben, mint például az asszociációs mezőké.

Ennek igazolására kísérletet terveztem 10 fő 10--13 éves gyermekkel, amelyben azonos zaj/jel szint mellett értelmes szavakat, illetőleg logatomokat kellett visszamondaniuk. Elemeztem a reakcióidejüket (az elhangzás és a visszamondás megkezdése között eltelt idő), valamint azt, hogy mit mondtak vissza. A felnőttek eredményeivel összevetve a következőket tapasztaltam: a) a gyermek reakcióideje sokkal rövidebb -- ennek nyilván pszichológiai vonatkozásai is vannak; b) nincs olyan lényeges eltérés a reakcióidőkben a szavak és a logatomok visszamondása között, mint a felnőttek esetében; c) a szavak visszamondása megegyezik a felnőttekével, egyik-másik gyermek esetében kissé bizonytalanabb; d) logatomokat a gyermekek legnagyobb részét logatomként mondták vissza, ellentétben a felnőttekkel, akik a logatomok elhangzásakor is értelmes szavakat igyekeztek visszamondani és e) a gyermekek visszamondásában nem volt különbség a logatomok között, míg a felnőttek esetében döntő szempont volt a logatom fonotaktikai struktúrája, hogy ti. megfelel-e az és milyen mértékben a magyar nyelv hangsorépítési szabályszerűségeinek (Gósy 1984b).

Az anyanyelvi percepció bázis kialakulása alighanem nyelvenként változó. Angol anyanyelvűek esetében szinte a 12 éves korig tolódik ki; a hangszólás és a dallam felismerése 2--3 éves korban még igen rossz, a fonotaktikai elrendeződések azonosítása kb. 4 éves korra alakul ki (Clark--Clark 1977). Menyuk szerint a beszédelsajátítás egyrészt 4--5 éves kor körül, másrészt a pubertás előtt "fejződik be" (1971, 48). Magyar anyanyelvű gyermekekkel erre vonatkozó kísérleteink nincsenek; a szakirodalom empirikus adatai arra engednek következtetni, hogy a 6--10 éveseknek viszonylag szilárd a percepció bázisuk.

A percepció bázis kialakulása -- az általános tendenciák mellett -- nagy egyéni különbségekkel járhat: egy-egy életkori szakaszban az egyébként ép hallású, jó intelligenciaszintű gyermekek beszédmegértése feltűnően eltérő lehet. Két azonos korú, a szinuszos hanggal történt hallásvizsgálattal

azonosan ép hallásúnak nyilvánított gyermek beszédmegértése nem feltétlenül azonos. Ezek a különbségek pedig sokféle következményt vonnak maguk után. Kísérleti adataim szerint (Gósy 1984d) mindenekelőtt az írás, olvasás elsajátítására, a későbbiekben pedig kimutathatóan az idegennyelv-elsajátításra vonatkozólag. 6 és 16 évesekkel végzett kísérletben csökkenő hangerősségű egyszótagú szavakat (amelyeket jobb fülbe, fejhallgatón keresztül adtam) kellett a gyermekeknek visszamondaniuk. Az eredmények a következők: hallás, néha beszédhallás még mindig létezik, amikor beszédmegértés már nincsen; a beszédmegértés szoros összefüggést mutat az írás-olvasás, idegennyelv-tanulás készségével: minél magasabb szintű a beszédmegértés, annál könnyebb az írás, olvasás, illetve az idegen nyelv elsajátítása; a beszédmegértés -- feltehetően elsősorban 6 éves kor előtt --fejleszthető.

Szerte a világon keresik a választ arra, hogy miért képes a gyermek a pubertás kor előtt akcentusmentesen elsajátítani egy második, azaz idegen nyelvet, és miért sikertelen ez később. Az egyik feltevés szerint az ok a lateralizáció időpontjában van, amelyet a hetedik évre és még későbbi életkorokra tesznek. Számos kísérlet bizonyítja azonban, hogy nyelvi ingerre már csecsemőkorban jellemző a jobbfül-dominancia. A válasz -- nagy valószínűséggel az --, hogy az észlelési és a produktív rendszer a pubertásig marad rugalmas annyira, hogy az anyanyelvtől különböző jelek anyanyelvi szintű feldolgozására (és létrehozására) alkalmas.

A beszédészlelés fejlődése óvodáskortól

Három különböző kísérletet végeztünk. a) Az első szótalálási feladat volt; itt arra kerestünk választ, hogy vajon mekkora jelcsoport szükséges az egyes életkorokban a szóazonosításhoz. b) A másodikban a szófelismerés fejlődését vizsgáltuk. c) A harmadikban a mondatazonosítási teljesítményt elemeztük.

A szótalálási kísérletben a kísérletvezető meghatározott hosszúságú részleteket mondott hangosan előre összeállított szólistából. A résztvevők feladata az volt, hogy a jelcsoport elhangzása után próbálják azokat értelmes szavakká kiegészíteni (pl. ru --> ruha vagy sá --> sál, sár, sátor). A szófelismerés nyelvi anyagát tíz egy és két szótagú szó alkotta: dob, autó, szél, garázs, hal, gyűrű, vár, szőnyeg, óvoda, mese. Ezeket 6 mp-es szünetekkel férfi ejtésben magnetofonszalagra rögzítettük, majd ún. fehér zajjal

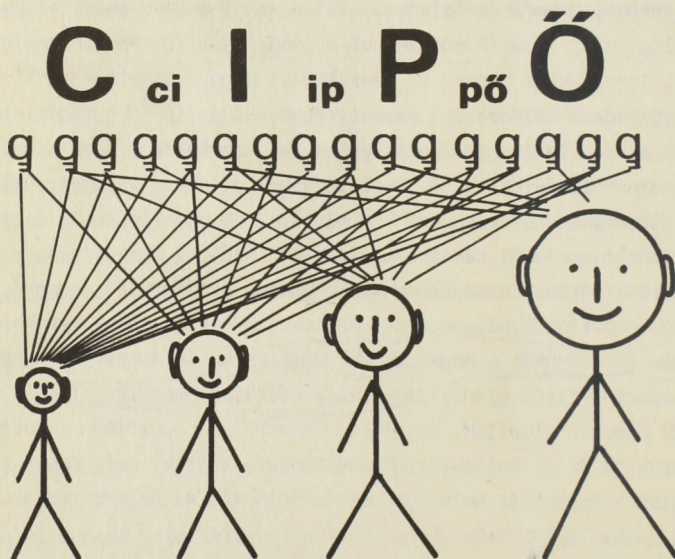
elfedtük, vagyis a megértést erősen megnehezítettük. A zaj/jel viszony 0 dB, a kísérletben résztvevők feladata a felismert szavak azonnali hangos ismétlése volt. A harmadik kísérlet nyelvi anyagában tíz rövid mondat szerepelt: A várakból romok lettek. A nőket rögtön bezárták. Mondjad, miért sírdogálsz! Alig mertek rálépni. A sebed helyén hegek vannak. Ettől a sörtől megrészegegett. E táncokat a múltban ismerték. Az utak hatalmasabbak az egekénél. A cárok késekkel dobálóznak. Milliók élnek az ügyért. A mondatazonosítási teljesítményt két módszer segítségével vizsgáltuk. Az egyik azonos volt a szöfelismerésben alkalmazottal: fehér zajjal elfedtük az ugyancsak férfi bemondóval 10 mp-es szünetekkel magnetofonszalagra rögzített mondatokat. A kísérleti személyek feladata itt is az elhangzás utáni azonnali hangos ismétlés volt. A másik módszerrel a mondatokat Varyspeech segítségével az eredeti tempójukhoz képest 25 százalékkal meggyorsítottuk; a kísérleti személyek feladata változatlan maradt. Mind a szó-, mind a mondatazonosítási feladatban a kísérleti személyeket egyénileg teszteltük: Sony típusú kazettás magnetofonról, fülhallgatón keresztül hallották az anyagot, átlagosan 65 dB-es intenzitáson, csendes helyiségben. Egy-egy kísérleti személy csak egyféle kísérletben vett részt (kivéve a szótalálási feladatot).

Az első kísérleti csoportban kiscsoportos óvodások (3 évesek), a másodikban nagycsoportos óvodások (5 évesek) vettek részt. A harmadikban másodikosztályosok (7 évesek), a negyedikben ötödikesek (10 évesek) szerepeltek. Az ötödik csoportot elsős gimnazisták (14 évesek), a hatodikát főiskolások (átlagosan 20 évesek) alkották. Az utolsó (kontroll) csoportban felnőttek (átlagéletkoruk 40 év, iskolázottságuk különböző) voltak. Valamennyi csoport szociológiai szempontból heterogénnek tekinthető. Az összes résztvevő száma: 214 fő.

Eredmények

a) A felnőtt beszédmegértésére jellemző, hogy szavak azonosításakor az akusztikai-fonetikai szinteken lejátszóó folyamatok szinte nyomban kiváltják a szószint aktivizálódását. Egy relatíve kevés információt hordozó beszédjel-csoport is elegendő ahhoz, hogy számtalan (a szókészlet és az egyéni szókincs korlátain belül) feltevésünk legyen a (jelen esetben) folytatást, azaz a teljes szót illetően. A szer hangcsoportról eszünkbe juthat a szerzám, a szerda, a szerelem, a szerep, a szerencse, a szeret, a szerelem stb. szó. Ennek a folyamatnak az az alapja, hogy az említett alsóbb szintek csak részlegesen vesznek részt a megértési

folyamatban. Ez azt jelenti, hogy a feldolgozáshoz nincs szükség az adott beszédjel valamennyi sajátosságára, csak a legfontosabbakra. Ezzel szemben a kisgyermeknél már az első szintektől kezdve valamennyinek a teljes értékű részvétele szükséges ahhoz, hogy a folyamat hibátlanul működjön. A gyermek -- életkorától függően -- az elhangzó beszédjelet különböző mértékben dolgozza fel. A fejlődés éppen abban jelentkezik, hogy csökken az alsóbb, növekszik a felsőbb szintek részvétele a megértésben, vö. 33 ábra. Kísérleti adataink ezt a következőképpen támasztják alá.

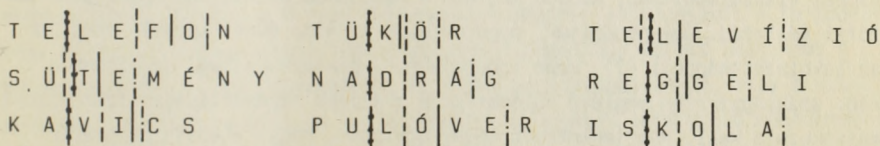


33. ábra

A felismerési kulcsok felhasználásának változása
az életkor függvényében

A szótalálási feladatban a háromévesek nagyobbik részének a teljes szó elhangzására szüksége volt a megértéshez. Kisebb részük az utolsó előtti hang elhangzása után a szavak jelentős részét már képes volt felismerni. A nagycsoportos óvodásoknál általában előbb bekövetkezett az azonosítás, sőt néhány szókezdetre több szót is meg tudtak nevezni. Például: te --> teve,

televízió, tele van vagy ke --> kemence, kell, kenyér, kendő. A második osztályosoknál -- várhatóan -- még előbb ismernek rá egy-egy szóra, és majdnem minden két hangból álló hangkapcsolatot ki tudnak egészíteni értelmes szóvá, például: sü --> sün, süllyed, sült, sütő. A fejlődés mégis kisebb mértékűnek tűnt annál, mint amit elvártunk volna. Az ötéveseket a hétévesektől nem csupán az írni-olvasni tanulás ténye választja el egymástól, hanem feltételezhetően nagyobb mértékű ismeretanyag-növekedés is. Továbbá a jelenlegi olvasztanítási módszerek egyenesen a gyermekek szóátalálási "tudására" építenek. Még nehezebb feladat lesz tehát az írás és az olvasás elsajátítása azoknak a gyermekeknek, akiknél ez a képesség az óvodáskori szintet nem vagy alig haladja meg. Nagy egyéni különbségeket találtunk a második osztályosok között: több gyermek az ötévesek átlagteljesítményét érte csak el. Az ötödik osztályosoknak nem jelentett problémát ez a feladat; az első (szókezdő) hangkapcsolatra számos értelmes szót tudtak mondani. Az egyes gyerekek közti különbség ebben az életkorban főként az idővel fejezhető ki: volt, aki gyorsan, volt, aki hosszabb gondolkodás után egészítette ki különböző szavakká az elhangzott jelcsoportot. A 34. ábra a fejlődést kilenc vizsgált szóval szemlélteti. A középiskolásoktól kezdve az egyes kísérleti személyek között csupán a "megtalált" szavak mennyiségében volt különbség.



34. ábra

Szóátalálási teljesítmény 3, 5, 7 és 10 éveseknél

(3 évesek: — — — — , 5 évesek: ————— , 7 évesek: — — — — , 10 évesek: —————)

b) Egy szó felismerése, majd pedig jelentésének megértése több tényező függvénye; gyakran a beszédhelyzet határozza meg, hogy közülük melyiknek lesz elsőbbsége (ld. a magyarra: Gósy 1984d). A kísérleti szavakat úgy választottuk ki, hogy azok már az óvodáskorúaknak is feltételezhetően az aktív szókincsébe tartozzanak. (A szójelentés ismerete, illetőleg a használat igen különböző lehet.) Eredményeinket a 38. táblázat tartalmazza.

Vizsgált csoportok	Helyes válaszok százalékban
3 évesek	47,5
5 évesek	51
7 évesek	63,3
10 évesek	67,1
14 évesek	73
főiskolások	80
felnőttek	84

Legnagyobb a fejlődés a 3--7 éves időszakban. Az iskolába kerülés, az első osztály elvégzése minőségi ugrást jelent: ebben a csoportban már igen ritkán fordult elő, hogy a gyermekek hibásan ismételték volna az elhangzott szót (pl. értelem nélküli hangsorként). Ugyancsak igen kicsi volt a "nincs válasz" aránya, amikor a gyermek semmit nem ismételt. Az előző, szótalálási feladat teljesítményéhez képest itt jelentős fejlődést tapasztaltunk. Ehhez képest visszafogottabb, kisebb mértékű a másodikos és ötödikes gyermekek közötti változás. Úgy gondoltuk, hogy nemcsak a három évnyi korkülönbség, de az iskolázottságbeli (?), azaz ismeretekben, gondolkodásban, szókincsben stb. szükségszerűen meglévő különbségnek a szavak azonosításában feltétlenül mutatkoznia kell. Nem tekinthető megnyugtatónak, hogy -- adataink szerint -- a nagyobb mértékű fejlődés a 14-20 év közötti időszakban jelentkezik, amikor az ekorúak egy része számára a tanulás aktív korszaka már lezárult (ipari tanulók) vagy éppen lezárulóban van (érettségizők).

Adataink rámutattak a szavak különböző hírértékének szerepére az egyes korcsoportokban. Az óvoda azonosítása 100 %-osnak tekinthető a két óvodai csoportban (a 3 évesek nem ismételték meg a szót minden esetben, de sosem tévesztették). Hasonlóképpen jó a felismerés a második osztályosoknál is; az ötödikesek azonban már majdnem 30 %-ban dolgozat-nak értik e szót, még a főiskolás csoportban is 25 %-os az óvoda dolgozat azonosítása. Hasonló okból jobb a hal szó felismerése az óvodásoknál, mint a többi életkori csoportban, ahol inkább falat értettek. (Az óvodákban hal jel van, fal természetesen nincs.)

A 39. táblázatban összegeztük az egy és több szótagú szavak helyes felismerésére kapott adatokat.

39. táblázat

Vizsgált csoportok	Helyes felismerés százalékban	
	egy szótagúakra	több szótagúakra
3 évesek	36,5	50
5 évesek	46,8	58,3
7 évesek	38,8	79,6
10 évesek	55,3	75
14 évesek	55	85
főiskolások	56,2	95,8
felnőttek	71,3	92,8

Az adatok alátámasztották azt a jelenséget, hogy zavart helyzetben könnyebb a több szótagú szavak azonosítása, mivel több elemből könnyebben "rekonstruálható" a hallott hangsor (Büky--Egyed--Pléh 1984, 288). Másfelől érdekes képet kaptunk a helyes felismerés arányának változásáról a fejlődésben. Az egy és több szótagú szavak átlagos azonosítása között legkisebb a különbség a két óvodás csoportban. Ez újabb bizonyítéka annak, hogy a kisgyermek kevesbé használják fel a magasabb szinteken lejátszódó működések eredményeit. Egyenletesebben változó és sokszor nagyobb mértékű a több szótagú szavak helyes felismerése a többi csoportban; vagyis növekszik a felsőbb szintek részvétele a megértési folyamatban. (Két negatív adatunkra, a "visszaesést" mutató azonosításokra nem találtunk elfogadható magyarázatot.) Valamennyi korcsoport teljesítményét összegezve, az egy szótagúakra kapott korrekt felismerés átlaga: 51,4 %, a több szótagúaké: 76,6 %.

c) A mondatazonosítási kísérletben szereplő 10 mondatból hétnek a jelentése általában érthető, háromé nehezen vagy egyáltalán nem volt érthető az óvodáskorúak számára. Mondatok azonosításakor felnőtteknél még jobban csökken az alsóbb szintek részvétele a megértési folyamatban; mivel a szintaktikai szabályokat hordozó beszédjelek, valamint maguk a szabályok is nagy segítséget jelentenek a globális egész, a mondatjelentés felfogásához (vö. Foss 1969).

A zajban elhangzó mondatokra kapott adatokat a pontosan azonosított szavak mennyisége szerint is elemeztük. Ezeket és a gyorsított mondatok felismerésére kapott eredményeket a 40. táblázat mutatja.

40. táblázat

Vizsgált csoportok	Helyes válaszok százalékban		
	természetes elhangzásban	zajban hangzó mondatokra	gyorsított mondatokra
3 évesek	45	10,0 -- 28,4*	10
5 évesek	85	36,2 -- 57,6	58
7 évesek	100	41,4 -- 62,5	59,3
10 évesek	100	46,4 -- 68,4	89,3
14 évesek	100	52,0 -- 78,1	93
főiskolások	100	62,9 -- 82,2	97,1
felnőttek	100	66,0 -- 80,3	98

* Bal oldalon a tíz mondat helyes felismerésének átlaga, jobb oldalon az összes mondatban felismert valamennyi szó átlaga szerepel.

Az izolált szavak azonosításához viszonyítva töretlenebb a mondatazonosítás fejlődése a számadatok tükrében. A gyermekek teljesítménye mindkét típusú feladatban gyengébb, mint a főiskolásoké vagy a felnőtteké. Feltűnő, hogy a mondatazonosításban a "minőségi ugrás" nem az óvoda/iskola választóvonalnál, hanem egyrészt a 3 és 5, másrészt a 14 és 20 évesek között látható. A mondatokban szereplő szavak megértésében egyetlen nagyobb ugrás van: a két óvodai csoport között. Kimondható ennek alapján, hogy a beszédmegértés fejlődésében a legnagyobb fordulópont az óvodai években zajlik, három- és ötéves kor között. Hasonló jelentős változás -- az izolált szavaknál ugyanezt tapasztaltuk -- van 14 és 20 év között is. További vizsgálatra szorul, hogy a főiskolások és a felnőttek teljesítménye közötti különbség szignifikáns-e.

Hároméves korra az anyanyelv elsajátítása alapfokon megszilárdult (Gósy 1984a). Kísérletünk ugyanakkor azt bizonyította, hogy az e korúak beszédészlelése még mindig nagymértékben függ az adott beszédhelyzettől. A nem megszokott téma és a feladat gyakorlatilag a nullára csökkenti a teljesítményüket. Meg kell jegyeznünk, hogy ugyanakkor néhány esetben egészen kiemelkedő,

meglepően jó eredményeket kaptunk: L.E. 3 éves kislány a zajban elhangzó mondatokat 30 %-ban jól ismerte fel (az életkori átlag 10 %), a pontosan azonosított szavak aránya 45,2 %; N.Zs. 5 éves kisfiú mondataazonosítása 50 %-os, a mondatokban szereplő szavak helyes felismerése 80,6 % -- a gyermek ezzel a 14 évesek átlagteljesítményét közelíti meg.

Kísérletek nélkül is valószínűsíthető, hogy a beszédészlelés (és a beszédmegértés) az életkorral párhuzamosan fejlődik; pontosabbá, részletezőbbé válik. Ebben az értelmi fejlődés mellett az ismeretek bővülésének, az aktív tanulásnak, a beszédtapasztalatnak (a különféle kommunikációs helyzetekben történő sikeres részvételnek) és még más tényezőknek is fontos szerepe van. Kísérletünk meggyőzött arról, hogy ez a fejlődés nem mindig fokozatos, számos esetben láthatók minőségi/mennyiségi ugrások. Szemléletesen bemutatva ezt a fejlődést a kapott adatok alapján, grafikonon ábrázoltuk a beszédészlelés alakulását 3 éves kortól felnőttkorig (35. ábra).

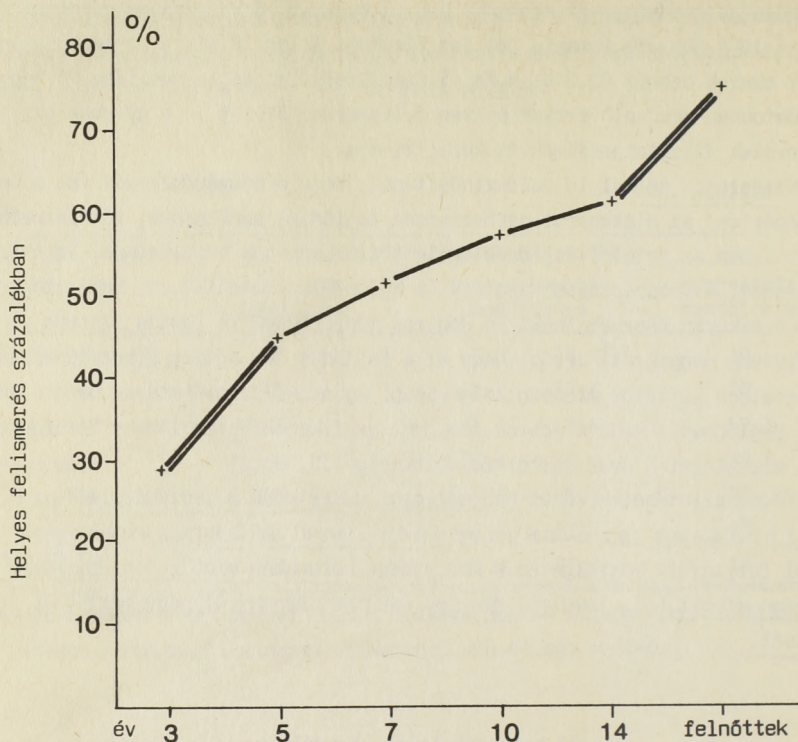
Következtetéseink közül hangsúlyozni szeretnénk a beszédészlelés fejlesztésének szükségességét. Annál is inkább, mivel a gyermekkel folytatott beszélgetések mennyisége korunkban mind a családi, mind a családon kívüli (pl. iskolai) életben (sajnos) rendkívül nagy mértékben lecsökkent.

E0016

A beszédészlelés vizsgálata

Bevezetés

Az előző fejezetben láttuk, hogy a gyermekek beszédelsajátítása folyamán szoros kapcsolatban fejlődik az ún. artikulációs és a perцепciós bázis. A megértés eleinte nagymértékben megelőzi a produkciót, majd a köztük levő nagy különbség egyre csökken (a fejlődés állandó kölcsönhatásban történik). Éppen ezért különösen nagy jelentősége van annak, hogy a hallás és a beszédészlelés, valamint a beszédprodukció épségét állandóan ellenőrizzük. Nem elegendő azonban a puszta (szülői) megfigyelés. Egzaktabb és rendszerezesebb vizsgálatok, szűrések lennének szükségesek. Különösen a perцепciós bázis működését illetően, mivel ennek a hibája gyakran rejtve marad; s mire nyilvánvalóvá válik, a javíthatóság, korrigálhatóság lehetősége erősen lecsökkent. Minél korábban és minél pontosabban meghatározhatók az ép működéstől



35. ábra

A beszédészlelési teljesítmény fejlődése
óvodáskortól felnőttkorig

való eltérések, annál nagyobb a gyógyíthatóság esélye.

A hallás épségének és a beszédmegértés megfelelő szintjének vizsgálata még klinikai körülmények között is számos nehézséggel jár, ha kisgyermekről van szó. Megoldatlan az óvodáskorúak rendszeres szűrése, a diffúz dyslaliás gyermekek és a mentálisan károsodott gyermekek vizsgálata. Az Egészségügyi Minisztérium szakmai irányelvben írja elő a 0--18 évesek audiológiai szűrővizsgálatát (7006/1983. Eü. K. 12). Az ehhez kapcsolódó szakmai vélemények és tapasztalatok azt mutatják, hogy a legnagyobb módszertani nehézség éppen a három, illetőleg a három--hat éves korosztályt érinti (Szabó E.--Szabó Gy.--Simon 1984).

Az elmúlt években végzett beszédészleléssel kapcsolatos kutatások lehetőséget teremtettek olyan új eljárások kidolgozására, amelyek az említett gondokat próbálják megoldani.

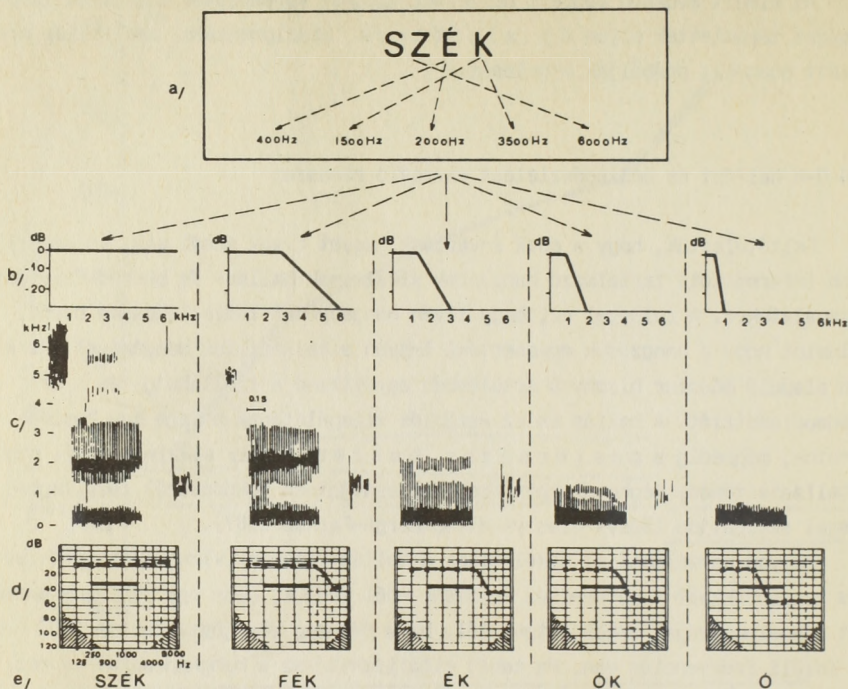
A G-0-H hallást és beszédészlelést vizsgáló rendszer

Feltételeztük, hogy a csak invariáns jegyet (vagy annál alig valamivel több információt) tartalmazó hangsorok alkalmasak hallás- és beszédészlelési vizsgálatokra. A feltétel az, hogy ilyen hangsorokat létre lehessen hozni, valamint hogy a hangzásuk egyenértékű legyen a természetes beszédével. Az ezen alapuló módszer bizonyos értelemben egyesítené a tisztahang- és a beszédaudiometriát. A hallás és az észlelés vizsgálatának anyaga így beszéd lehetne, mégpedig a mesterséges beszéd, amely alkalmas arra, hogy a hallástartomány meghatározott frekvenciasávjainak működéséről felvilágosítással szolgáljon (Gósy--Olaszy--Hirschberg--Farkas 1987).

Hogyan képzelhető el elméletileg a hallástartomány vizsgálata mesterséges hangsorokkal? A percepciós eredményekből tudjuk, hogy egy szá hangkapcsolat mássalhangzóját akkor értjük jól, ha a réshang zörejjösszetevője 4000 Hz-en felüli frekvencián van. Ha tehát előállítható ez a hangkapcsolat -- mesterségesen -- úgy, hogy a [s] zöreje például 5000 Hz-en helyezkedik el, s a frekvenciatartomány az adott időtartamban másutt nem tartalmaz a mássalhangzóra jellemző komponenseket, feltételezhető, hogy a felismerése abszolút helyes lesz. Tegyük fel azonban, hogy olyan kísérleti személy hallgatja, akinek 4000 Hz-től kezdve erősen csökkent a hallása (mindkét oldalon). A hallássérült személynél a jelfeldolgozás első szintjén zavar támad: az 5000 Hz-es zörejt ugyanis -- a halláscsökkenés mértékétől függően -- más-más hangként fogja észlelni, esetleg egyáltalán nem fogja meghallani. A szá felismerése sá-vá, fá-vá, há-vá, majd tá-vá, végül á-ra fog torzulni. Mindezzel szemben a természetes szá hangsort, akár férfi, akár nő ejtésében, ez a hallássérült személy a többletinformációk alapján még tökéletesen azonosítja (36. ábra).

A mesterséges szavak előállítás

Mesterséges hangsorok többféle módszerrel állítható elő; három nagy csoportjuk: a formánsszintézis, az ún. LPC és a hullámszintézis. Olyan hangsorok előállításához, ahol előre meghatározott - kikísérletezett -- értékre



36. ábra

A mesterséges (szintetizált) hangsor szűrőfunkciójának működése halláskárosodások esetében

- a) a szék szó felismeréséhez/megértéséhez szükséges frekvenciaértékek (az akusztikus invariáns jegy alkalmazása a szintézisben)
- b) szűrőállások: frekvenciaszűrés halláskárosodás imitálására
- c) hangszíneképek: a szék szó frekvenciaszűrés után megmaradt jelrészének akusztikai szerkezete
- d) audiogramok: a frekvenciaszűréssel imitált/igazolt feltételezett halláskárosodások mértékét szemléltető küszöbgörbék (a frekvenciatorzításnak megfelelően magashang-károsodás imitálása)
- e) a szék szó helyett -- a frekvenciaszűrésnek megfelelően -- azonosított hangsorok

kell az adott paramétereket beállítanunk, legmegfelelőbb a formánsszintézis eljárása.

A mesterséges szavakat több lépcsőben fokozatosan alakítottuk ki (vö. Gósy 1986). A fokozatok a következők voltak:

- a tesztszavak természetes, férfi ejtésben történő rögzítése magnetofon-szalagra,
- a természetes szavak akusztikai szerkezetének elemzése (frekvencia, intenzitás, idő tekintetében),
- a mesterséges hangsorok generálása az UNIVOICE szintetizáló rendszerrel,
- a mesterséges változatok akusztikai szerkezetének elemzése (frekvencia, intenzitás, idő),
- a természetes és a mesterséges hangsorok akusztikumának és hangzásának összevetése,
- a generált hangsorok akusztikai szerkezetének módosítása a természetes ejtés elemzett összetevőinek és hangzásának alapján,
- a mesterséges szavak hangminőségének javítása, kísérleti személyek véleménye alapján,
- az akusztikus invariáns jegy percepciójáról kapott eredmények beépítése a mesterséges hangsorok szerkezetébe, a redundancia csökkentése,
- a beszédhangok összetevőinek előre meghatározott frekvenciasávokra koncentrálása, például a mássalhangzó zörejjócának változtatása a frekvenciatengelyen,
- a módosított frekvenciaszerkezetű mesterséges hangsorok felismerésének vizsgálata -- tesztelés.

A több kísérletben tesztelt szóanyag végleges formájában alkalmassá vált a további vizsgálatokhoz (vö. a bú, a sír és a szűz szó előállítási adatait: 41., 42. és 43. táblázat).

A mesterséges szavak szűrőfunkciójának működése

Ép hallás esetén 60 dB-en 100 %-os klinikai körülmények között mérve az egyszótagú természetes ejtésű magyar szavak felismerése (Götze 1974). (Ez a tisztahang-audiogramon azt jelenti, hogy a küszöbgörbe 10 és 20 dB között helyezkedik el, klinikai körülmények között mérve; nem tekintve most bizonyos centrális hallászavarokból adódó eltéréseket.) Ha a küszöbgörbe lejjebb "csúszik", azaz halláskárosodás áll fenn, akkor a 60 dB-es értéken nem lesz

KÓDSZÁM: BÚ

Hangszept azonosító	IDŐ (ms)	IDA	AD (dB)	F0 (Hz)	F1 (Hz)	B1 (Hz)	F2 (Hz)	B2 (Hz)	F3 (Hz)
81	40	30	2	97- -92					
82	60	10	3- -2	100- -97					
83	40	1	8- -10	122	400- -300	60	654	62	2614
BU	40	10	12- -14	122- -116	300	60	654- -617	93	2614
U1	40	12	14	112- -109	300- -252	60	617	93	2614
U2	40	40	14	109	252	60	617	93	2614
U3	40	25	12-0	109- -112	252	60	617	93	2614

100 %-os a megértés. Tudjuk továbbá, hogy az egyes szavakat milyen szerkezetű beszédhangok építik fel, ezeknek az észlelése állandó hangerő mellett felvilágosítással szolgál a hallástartomány ép voltáról vagy zavart működéséről.

A mesterséges szavakból álló tesztanyagot a beszédaudiometriában szokásos módszerrel használtuk a kísérletekben. Ez azt jelentette, hogy a szavakat fejhallgatón keresztül, különböző erősségeken hallotta a gyermek, egyszerre csak az egyik fülében. Feladata azok elhangzása után azonnali ismétlésük volt. Az egyes szavakat 8 mp-es szünetekkel rögzítettük magnetofonszalagra; a tapasztalat az, hogy ez a szünettartam két szó között elegendő arra, hogy a gyermek felismerje és hangosan megismételje a hallott hangsort. A 40 mesterségesen előállított, szintetizált egyszótágú szót a 44. táblázat tartalmazza.

A gyermek hangos ismétléseit, visszamondásait a kísérletvezető azonnal leírta (rögzítve a gyermek életkorát, az aktuális intenzitásértéket, valamint a szóanyag típusát, ti. természetes vagy mesterséges változat volt-e). Éphalló felnőttekkel és gyerekekkel végzett kísérleteinkben a lejátszási hangerő konstans volt: 60 dB.

Kísérleti anyagunkat laboratóriumi, klinikai és óvodai

KÓDSZÁM: SÍR

HANGSZELET AZONOSÍTÓ	IDŐ (ms)	IDA	AD (dB)	F0 (Hz)	F1 (Hz)	B1 (Hz)	F2 (Hz)	B2 (Hz)	F3 (Hz)	B3 (Hz)	AC (dB)	AK (dB)	K1 (Hz)	K2 (Hz)
SI	60	30									6- -22	16	1796	2540
S2	10	70									28	16	1796	2540
S3	40	10									24	16	1796	2540
SI	40	1	0- -4	133- -122	212	60	1796- -1903	155	2614					
I1	40	10	6- -10	122- -116	212	60	1903- -2016	155	2614	500				
I2	40	12	10	116- -112	212	60	2016- -2263	186	2614	500				
13	40	12	10	112	212	60	2329	186	2614	500				
14	40	12	10	112	212- -245	60	2329- -2136	186	2614	500- -250				
IR	50	15	10	116	252- -300	60	2075- -1849	186- -93	2614					
R1	40	8	6- -10	112- -116	300- -400	60- -96	1796	93	2614					
R2	40	3	0- -4	106	400	60	1745		2614					
R3	40	4	6- -0	103	400		1745	2614						

43. táblázat

KÓDSZÁM: SZŰZ

HANGSZELET AZONOSÍTÓ	IDŐ (ms)	IDA	AD (dB)	F0 (Hz)	F1 (Hz)	B1 (Hz)	F2 (Hz)	B2 (Hz)	F3 (Hz)	AC (dB)	AK (dB)	K1 (Hz)	K2 (Hz)
SZ1	60	25								0- -20	22- -6	6041	11079
SZ2	20	30								22- -24		6041	11079
SZ3	40	30								25- -22		6041	11079
SZŰ	40	1	0- -12	112	206	60	647		2691				
Ű1	40	10	12	112- -109	206	60	1647		2614- -2540				
Ű2	40	10	12- -14	109- -106	206	60	1647		2468- -2397				
Ű3	40	35	14	106- -103	206	60	1647- -1796		2397				
ŰZ	40	10	14- -10	103- -100	206	60	1796- -1647		2397- -2486	0- -12		6041	9051
Z1	40	10	10	100	206- -238	60	1647	31- -93	2468- -2614	16		6041	9051
Z2	50	10	8- -6	100	259	96	1647	93	2691	14- -10		6041	9051
Z3	60	10	6-0	100	218	96-	1600	93	2614	16- -14		6041	9051

Sorszám	Szintetizált szavak			
1.	meggy	mos	bőr	bab
2.	síp	sír	csíp	sár
3.	búr	bók	szög	busz
4.	ász	szűz	cél	ősz
5.	gép	ágy	gáz	mák
6.	zsír	sző	szíj	sín
7.	bot	bor	bú	bál
8.	szél	csók	só	szék
9.	gyík	kút	dob	méz
10	szú	ész	zsák	cím

körülmények között próbáltuk ki (Gósy 1986). A laboratóriumi kísérletek egyik célja az volt, hogy megállapítsuk, vajon a mesterséges szavak hangminősége megfelelő-e, azaz éphalló, magyar anyanyelvű gyermek és felnőtt könnyen felismeri-e és megérti-e azokat. A másik cél annak a megállapítása volt, hogy a hangzásában egyenértékűnek tekinthető természetes és mesterséges szavak akusztikai szerkezetének különbsége a halláskárosodások jelzésére, tehát kiszűrésére alkalmas-e. A pozitív eredménnyel zárult laboratóriumi kísérletek alapján tesztanyagunkat klinikai körülmények között ép- és nagyothalló gyermekekkel hallgattattuk le. Feltételezésünk az volt, hogy a) az éphallók a természetes szavakkal megegyezően fogják a mesterséges szavakat visszamondani; b) a nagyothallók eltérően azonosítják a természetes és a szintetizált hangsorokat. Amennyiben a két feltételezést a klinikai kísérletek igazolják, akkor felállítható az a munkahipotézis, hogy a mesterséges hangsorokkal kimutathatók a halláskárosodások, mégpedig egyetlen hangerőn, azon, amelyen az éphallók beszédmegértése természetes szavakra 100 %-os.

A természetes és a mesterséges hangsorok akusztikai szerkezetének különbségét hallássérült gyermekek percepciójának megítélésére Angliában próbálták felhasználni (Hazan--Fourcin 1983); céljuk a gyermekek percepció fejlesztésének kimutatása volt. Kísérletükben stimulusként szintetizált magánhangzókat használtak. Feltételezik, hogy ha megállapítható, hogy az ép gyermek mikor sajátítja el a beszédhangok fő felismerési kulcsait, akkor megíthető a halláskárosodott gyermek elmaradása is (41--42). Egy finn kutatócso-

port ugyancsak szintetizált szavakkal kísérletezett: 50--160 szót adtak 1 perc alatt normál és hallássérült személyeknek (Rahko--Karjalainen--Laine--Lavonen 1979) visszamondásra. Eredményeikből úgy tűnik, hogy e módszer alkalmas a centrális hallópályák működésének megítélésére (a közleményből nem derül ki, vajon miért alkalmaznak a természetes szavakkal ismert módszerben mesterségeseket).

A 20 éphalló óvodás gyermek vizsgálatával kapott eredmények az alábbi tendenciákat mutatják:

- a gyermekek mind a természetes, mind a mesterséges szavakat visszamondták, azonosításukban különbség nem volt;
- a gyermekek számára érthetetlen szavak, amelyeket ők logatomokként mondtak vissza, jól jellemezték az észlelési szintjüket.

A 40 hallássérült gyermek halláskárosodásának típusa és mértéke eltérő volt. Az egy-egy fülbe juttatott 10 szó visszamondása után ellenőriztük, hogy a gyermek válaszai mennyiben egyeznek meg halláskárosodásának típusával, mértékével, illetőleg a szinuszhanggal mért küszöbgörbével. Majd összevetettük ugyanazon beteg gyermek válaszait a természetes és a mesterséges szavakra. Azt láttuk, hogy a szintetizált szavak felismerésének mértéke -- csak úgy, mint a laboratóriumi szűrési kísérletben, az imitált halláskárosodásoknál -- korrelál a küszöbgörbe értékeivel. Ez azt jelenti, hogy ha a küszöbgörbe az alsóbb frekvenciákon mutatott az éphez képest negatív eltérést, akkor a gyermek a fülébe juttatott 10 szó közül azokat mondta vissza hibásan, amelyeknek felismerési kulcsa az alsó frekvenciasávok valamelyikébe esett (például [u, o] vagy [b, m]). Ha a küszöbgörbe a felső frekvenciákon nem érte el az éphallás szintjét, akkor a gyermek azokat a szavakat tévesztette, amelyekben a beszédhangok felismerési kulcsa a magas frekvenciákon volt (például [i:, e:] vagy [s, f]). Vö. 46. táblázat. A percepciós kísérletekből láttuk, hogy a frekvenciaszűrés mértékétől függően törvényszerűen azonosítják a kísérleti személyek az egyes beszédhangokat más hangminőségeknek. A természetes beszéd esetén azonban ezek csak igen kis mértékben valósulnak meg, a fokozatok a nagyfokú redundancia következtében erősen lecsökkenten vannak jelen, illetve egészen eltűnnek. Vö. Gósy--Olasz--Hirschberg--Farkas 1985.

A megelőző laboratóriumi kísérleteknél már látható volt, hogy nagyobb frekvenciaszűrés esetén is jól vagy közel jól azonosíthatók a természetes ejtésű szavak, míg mesterségesen előállított megfelelőik már torzak vagy teljesen felismerhetetlenek. A hallássérült gyermekeknek a kétféle tesztanyagra adott válaszaiból ugyanez a kép tárul elénk: ugyanazon intenzitás-

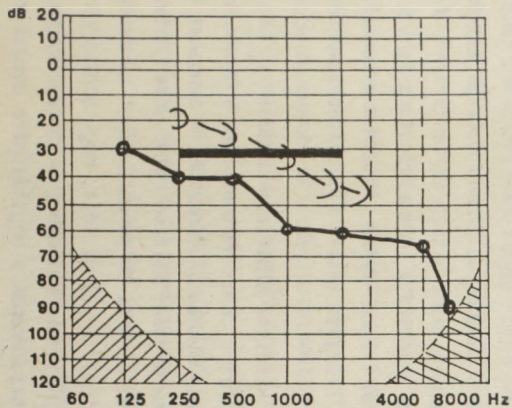
Nagyothallás foka/típusa	Vezetéses		Percepciós		Kevert	
	bea- dott szó	v á l a s z	bea- dott szó	v á l a s z	bea- dott szó	v á l a s z
enyhe (10-30 dB)	gép-dép, tép, kép, ék bók-pók, ók, ó búr-úr, ór, ól bab-gab, bag, bak, pad gyík-kép, sík mos-nos, os, mus, fus mű-bú, bi		síp-csíp, íp ész-és, ép, ét szúz-szül, ül, szék cím-ím, szink szú-sú, fú szíj-szó, só ész-ágy, áll, ás, át, áp		bók-pók, góp búr-úr, ø gyík-gyúk, ki mos-mus, vos mű-lí, vú	+ ész-és, ép, ét szúz-sün, fú cím-ó, ø szú-gó, fú szíj-síp, só ész-ás, áll, át, áp
közepes (30-50 dB)	gép-kép, szék bók-pók, hók, ó búr-dúr, túr, ól, ú bab-gab, pat, dap gyík-gép, kip mos-bos, vos, los, vos mű-vú, rú, lí, ø		síp-fűk, fik, ó, ø ész-és, ép, ét, ó stúz-súr, sír, ü, ø cím-szív, ír, il, ø szú-só, fú, dú, ø szíj-síp, hí, ó, ø ász-ágy, áj, ás, át, áp szék-ép, lék, ó, ø		bók-pók, bu búr-úr, ól gyík-gép, sik, csip mos-mus, vos mű-vú, rú meggy-legy, veggy gép-pép, ø	+ ész-és, ép, ét szúz-fül, tűz, ø cím-ó, ø szú-sú, fú szíj-ó, ø ász-ás, át, áp szék-lék, ló, ó
nagyfokú (50 dB felett)	gép-kék, ó, ø bók-buk, pók, ó búr-túr, úr, ø bab-ap, a gyík-ík, í, ó mos-vus, os, bos, bot mű-í, ø meggy-legy, vegy, begy 1000 Hz-nél		síp-gyík, ó, ø ész-ép, ét, é, ó, ø szúz-o/ cím-ø szú-tú, pú, ø szíj-út, ø ász-át, á szék-lék, ló, ó, ø 2000 Hz-nél		bók-hó, ó búr-ór, úr mos-vos, lus gép-ø meggy-lent-ve 4000 Hz-nél	+ ész-ét, ép, ó szúz-fül, ó, ø cím-fül, ó, ø szú-tú, ø szíj-fú, ø ász-ágy, át, á szék-hét, lék, ló, ó 6-8000 Hz-nél
"kút" esetén	csíp-só szíj-szó cím-ó		síp-fut csíp-só szék-szó		szú-sú, tú ész-és, ét, ép	cím-szív, ó szúz-fül, ø ász-ágy, áll, ás, át, á

Megjegyzés: A vezetéses enyhe esetében: 80 dB-en hallották; a közepesnél, a percepciós enyhe, közepesnél és a kevert enyhénél: 70 dB-en, a percepciós és kevert közepesnél 80 dB-en; a nagyfokúnál 90 és a fölötti dB-értékeken.

értéken adva a szavakat, lényegesen rosszabb a többletinformációt nem vagy csak kevésbé tartalmazó szintetizált hangsorok felismerése. A 37. ábrán egy a jobb oldali fülére nagyothalló gyermek válaszait láthatjuk, a jobb fül audiogramjával együtt. Az ábra jól szemlélteti, hogy míg például a síp szót természetes elhangzásban a gyermek az adott intenzitáson (100 dB) jól értette, addig a mesterséges változatot már só-ként azonosította. Az [i:] magánhangzót, amelynek a második formánsa 2240 Hz-en van, hallásromlása miatt nem tudta helyesen felismerni. A magánhangzó többletinformációt gyakorlatilag nem tartalmazott, ezért a döntést a gyermeknek az észlelt "megmaradó" frekvencia, vagyis az első formáns alapján kellett meghoznia.

Úgy is fogalmazhatunk, hogy az új eljárással a beszédaudiometriát közelítettük a tisztahang-audiometriához, amennyiben az előállított szavakkal egy vagy több frekvencián meg tudjuk ítélni a hallás épességét.

Ezek az eredmények vetették fel azt a gondolatot, hogy a szavak felismeréséből kikövetkeztethető a küszöbgörbe. A hallásmérés történetében eddig ennek a fordítottja volt ismert: a tisztahanggal végzett vizsgálati eredményből próbáltak a beszédmegértésre következtetni (Fletcher módszere: 1929; vö. Götze 1974, 154). Most a szófelismerés minőségéből próbáltunk a hallásküszöbre következtetni. Ezt az tette lehetővé, hogy a szavakat alkotó beszédhangok a felismerési frekvenciakulcsot mintegy 250 Hz-től 8000 Hz-ig tartalmazták, valamint hogy az értelmes szavak mellett eleve logatom jellegűek is előfordultak (a gyermeket így nem érte "készületlenül" az "értelmetlen" hangsor). A hallássérült gyermekek egy részénél a 10--10 szó visszamondatása után megpróbáltuk a küszöbgörbéjüket fölrajzolni, majd összehasonlítottuk a szinuszhanggal mért eredménnyel. Az összevetés azt mutatta, hogy 5 dB-os eltérésekkel közelítettük meg az eredeti -- audiométerrel mért -- görbéket. Lássunk erre egy példát! N.B. 7 éves kisfiú jobb fülében 60 dB-en hallotta a szintetizált szavakat, amelyeket az alábbiak szerint mondott vissza: meggy → meggy, síp → sut, búr → úr, ász → át, gép → got, zsír → súr, bot → bot, szél → fér, gyík → gú, szú → fú. Általános következtetések: az alsóbb frekvencián lévő szavakat pontosabban mondja vissza, mint a felsőbb frekvencián lévőket: a felsőbb frekvencián lévők között is van azonban különbség a visszamondásban. Feltételezhető, hogy nagyobb mértékben percepciós típusú, kisebb mértékben vezetékes típusú halláskárosodásról van szó. Az [o, u, m, b] helyes visszamondá-



BE



VÁLASZ

természetesre

mesterségesre

mos

mos

mos

meggy

meggy

led

síp

síp

só

cím

szén

ó

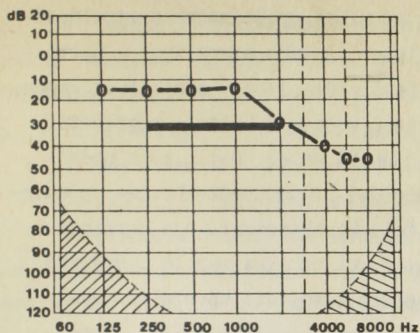
37. ábra

Ugyanazon szavak természetes ejtésű és mesterséges változatának eltérő felismerései halláskárosodás esetén (100 dB-en)

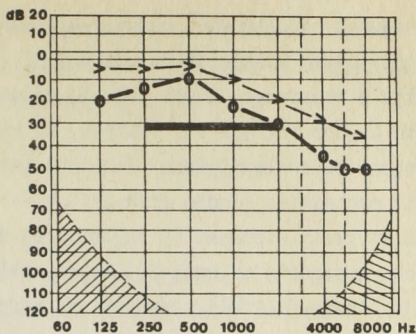
sa arra utal, hogy 250, 500 és 1000 Hz táján ép a hallás. Az [i:] és [e:], valamint a réshangok hibás felismerése valószínűsíti, hogy a gyermeknek 2000 Hz-től kezdve hallásromlása van. Az a tény, hogy az [i:] magánhangzót nem, a [ʃ]-t viszont helyesen azonosította a gyermek, feltételezi, hogy 1000 és 2000 Hz között kezd a hallása eltérni az éptől. Ez azért mondható így ki, mivel az [i:] felismeréséhez a 2330 Hz-es második formánst meg kell hallani -- ez nem történt meg, a [ʃ] percepcióját viszont a 2540 Hz-es komponens mellett egy 1500 Hz-es is biztosítja, aminek észlelése elegendő -- az adott hangerő mellett -- a mássalhangzó azonosításához. A [s] mássalhangzók különböző beszédhangként történt felismerése azt mutatja, hogy az adott frekvenciákon eltérés van a gyermek hallásában. A szű-ban szereplő 4000 Hz-es zörejt a gyermek [f]-ként, a szél-ben szereplő 6000 Hz-eset ugyancsak [f]-ként, míg az ász 7-8000 Hz-en lévő zörejjócat már zárhangként, [t]-ként ismerte fel. Ez a tény pedig -- a percepciók kísérletek tanúsága szerint -- azt jelenti, hogy az eredeti zörejből csak igen kis részlet volt feldolgozható, vagyis a gyermek csak keveset "hallott" belőle. Mindezek alapján a küszöbgörbe lefutása hozzávetőlegesen a következő: mintegy 1000 Hz-ig épnek tűnik, 1000 és 2000 Hz között kb. 15 dB-t, majd 2000--8000-ig további 10--15 dB-t csökken. A gyermek valószínű küszöbgörbéjét és tisztahang-audiogramját a vizsgált fülön a 38. ábra mutatja.

A görbékéből látható, hogy a valószínűsített és a szinuszhanggal mért eredmény között lényeges eltérés nincsen. Az eljárás tehát alkalmazható óvodáskorú gyermekek hallásszűrésére. Az a tény, hogy a mesterséges tesztanyag jobban mutatja a gyermek beszédészlelési szintjének különbségeit, mint a természetes szavak, felhasználhatóvá teszi egyúttal a 3--6 éves korosztály beszédészlelésének jellemzésére is. Ez annál inkább fontos, hiszen a beszédmegértés kellő fejlettségének megítélése elengedhetetlen (lenne) az iskolaérettség megállapításához.

A laboratóriumi és a klinikai kísérletek eredményei alapján anyagunkat -- szűrési céllal -- óvodai környezetben próbáltuk ki. Az első mérésorozatot 100 3--6 éves óvodással végeztük. A szavakat Uher magnetofonról közvetítettük a gyermekek jobb, illetve bal fülébe (felváltva), átlagosan 70 dB-es erősséggel (az eredeti 60 dB-es értéket az óvodai alapzaj miatt mintegy 15 dB-lel megemelten). A cél az volt, hogy megtudjuk: vajon egyetlen intenzitáson mérve, a szintetizált szavakból álló teszt alkalmas-e a gyermekek hallásának és beszédészlelésének a megítélésére óvodai körülmények között.



valószínű küszöbgörbe a
visszamondások alapján



audiómetérrel (MA 20-as
típus) mért küszöbgörbe

38. ábra

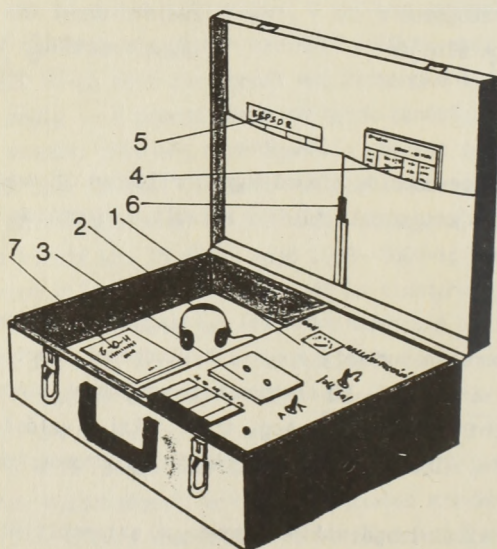
Hétéves, percepciós típusú nagyothallásban szenvedő
kisfiú jobb oldali füléről készült audiogramok

Mindegyik gyermeket -- szűrőaudiómetérrel -- tisztahang-audiometriával kontrolláltuk. 10 esetben nem tudtuk a kontrollt elvégezni, mert a gyermek nem működött közre, és mintegy 15--20 esetben bizonytalannak minősítettük az eredményt, mert eldönthetetlen volt, hogy mikor jelzi a valóságot a gyermek (ők elsősorban a kis, kisebb részben a középső csoportosok közül kerültek ki).

A 100 óvodás hallásvizsgálatának eredményei azt mutatják, hogy a tisztahang-audiometriával épnek ítélt gyermekek a szintetizált szavak megértésében lényegesen nem tévedtek. Az összes gyermeket figyelembe véve 15-nél tapasztaltunk kisebb-nagyobb eltérést a normálhoz képest (őket klinikai audiológiai vizsgálatra küldtük), a negatív irányú eltéréseket a szinuszhang-kontroll is alátámasztotta. Tapasztaltaink továbbá, hogy nem volt olyan gyermek, aki a szavakat jól visszamondta volna, ugyanakkor küszöbgörbéje nem esett az ép hallásnak megfelelő tartományba. A gyermekek szívesen közreműködtek, "játsoztak". Anyagunk óvodai környezetben szűrésre alkalmasnak bizonyult.

Az óvodai kísérlet egy sor technikai problémát vetett fel, mint például

a hangerő rögzítésének módja. Valamennyi technikai problémának a megoldására egy olyan készletet dolgoztunk ki, amely -- táskába szerelten -- biztosítja a tesztanyag óvodai alkalmazását szűrési céllal: a) kapcsolóval végezhető a jobb fül/bal fül váltása, b) kétállású kapcsoló kétféle rögzített hangerő használatát teszi lehetővé (átlagosan 55 és 65 dB-t), c) a beépített magnetofon minden frekvencián -- hasonlóan az audiométerhez -- egyenletes átvitelt biztosít (a lejátszófej átviteli karakterisztikája korrigálva van), a magnetofon csak lejátszásra alkalmas, d) az alkalmazott kisméretű fejhallgató átviteli karakterisztikája a szűrővizsgálati követelményeket teljesíti, mérete miatt a kisgyermek nem idegenkednek tőle (39. ábra).



39. ábra

A hallást és beszédészlelést vizsgáló kész-
lék és tartozékai

- 1 = a hangsorok elhangzását érzékelő/kijelző műszer
- 2 = kisméretű fejhallgató
- 3 = hangkazetta a speciálisan szintetizált hanganyaggal
- 4 = tesztlapok a mérés értékeléséhez (vö. 40. és 41. ábra)
- 5 = színes képsor beszédhibás gyermek vizsgálatához
- 6 = zöld és piros toll a jobb oldali és a bal fül jelzésére
- 7 = működési leírás, kezelési útmutató

A kísérleti eredményeket megpróbáltuk egy olyan összegzéshez felhasználni, amelynek segítségével az óvodások válaszait bárki (nem szakember is) értékelni tudja. Négy fokozatot határoztunk meg:

1. ép hallás -- jó beszédészlelés
2. enyhe hallás- vagy beszédészlelési probléma lehetősége
3. valószínű halláskárosodás
4. biztosra vehető halláskárosodás

Az ezeknek megfelelő helyre minden egyes szóhoz összeállítottuk a lehetséges válaszvariációkat (160 óvodás leggyakoribb ismétléseinek figyelembevételével). Így módon olyan tesztlapokat hoztunk létre, amelyeknek a segítségével fonetikai vagy percepcióval kapcsolatos ismeretanyag nélkül is megítélhető a gyermek hallása és beszédészlelése. A 40. és 41. ábra a 20--20 szóra feltételezett válaszlehetőségeket tartalmazó tesztlapokat szemlélteti (a tesztlapok használata igen egyszerű: a gyermek választát a mérést végző személy aláhúzással jelöli a tesztlap megfelelő hangsoránál, zölddel a jobb oldali, pirossal a bal oldali fület szemléltetve). Tekintettel arra, hogy a gyermekek -- a szűrés céljának megfelelően -- egyetlen hangerőn hallják a szavakat, az egyes oszlopok közötti valószínűsíthető halláscsökkenés mértéke mintegy 10 dB. A beszédészlelés megítélése szempontjából különös jelentősége van a második oszlopnak. Ezért itt úgy állítottuk össze a válaszlehetőségeket, hogy az eredeti szótól egyre távolabb eső hangsorok az egyre nagyobb eltérést mutatják (ez az értékelésnél fontos szempont). Például a dob-ra adható válaszok: dod, dud, dog, dug, bod, bob. A helyes választól legkevésbé a dod és legjobban a bob tér el.

A kialakított készlettel -- táskába szerelten -- a mesterséges hanganyaggal 242 3--7 éves gyermeket mértünk óvodai (tíz 3 évnél fiatalabbat bölcsődei) körülmények között. A 45. táblázatban foglaltuk össze a vizsgált gyermekek életkori és nemek szerinti megoszlását. 120 gyermek esetében történt azonnali tisztahang-kontroll vizsgálat (MA 20-as szűrőaudiométerrel).

A 242 óvodásból 18 gyermeket szűrtünk ki (12 fiút és 6 kislányt), és küldtünk klinikai hallásvizsgálatra; az ép hallású gyermekek száma: 224 fő, közülük 113 fiú és 111 kislány. (A kiszűrt gyermekek közül 4 kiscsoportos, 8 középsős és 6 nagycsoportos óvodás.)

Az elemzésekből a 10 bölcsődés gyermek anyagát kihagytuk, mivel az ő mérésük egyetlen célja az volt, hogy megtudjuk: a bölcsődés korú gyerekek vizsgálhatók-e az anyaggal. A legkisebb gyermek, aki a mesterséges szavakat visszamondta, így hallását értékelni tudtuk éppen 2 éves volt, beszédproduk-

A VÁLASZLEHETŐSÉGEK FOKOZATAI

Jobb fül - zöld szín
Bal fül - piros szín

Sor- szám	I. Jó hallás	II. Hallás- vagy beszédmegértési zavar állhat fenn, esetleg figyelmetlenség. A vizsgálat megismétlődő.	III. Hallászavar valószínű, orvosi vizsgálat javasolt.	IV. Hallászavar biztosra vehető, miejöbbi klinikai vizsgálat szükséges.
1.	meggy (megy)	hegy legy vegy negy	egy	bó od e ó -
2.	síp (sík)	sít sūt sūp szíp szép	zúg suk sut su só víd fut hó	kút út tú ú -
3.	bú	dú bók but bot bó pók pú púk dú gú	tú tó pú pó út	ó ú -
4.	ász	ház pászáz	ás ágy álláj	áf áh átáp á a ó ú -
5.	bot	but böt bó bu	pot put po pu ot ut op up lot dod tut	ó ú -
6.	szél	cél szép szil szír cil szék szül szöl	fél fél tér tél tól tór hé hó só	él ól ló ló tú ó -
7.	méz	méz méz misz niz miz műz	bé bó vél bű bō bóf ból bēl bōl	tóf tot tőt tú tó ó -
8.	zsír (sír)	szír szil zér zír síl sél zil zsi zse sí sé	fór fot fūr fúl fut fup fop	hú hó úr út ór ót up óp -
9.	bőr	pőr bōl pōl bér pér bél bír dōr ōr	bor por bur pur	bo bu po pu tó tu ó ú -
10.	szú	szó szūr szül szór szőr szű szūr	sú sū só fő fú	hó hú ló lú tó tú pú út ót ó ú -

11.	mos	mus hus bos vus vos mosz los lus nos mös mes musz	busz bosz búf vot vut mo tuf mot bot	po pu-op up -
12.	sír	síp sí szil sél sūr sé sér szír szér	fúr fül fut hír híl fot fup fop hú hó ér	úr út ór ót tú tó ú ó -
13.	bab	dad hob dab bag gab bo ba gad dag dod dog	pap pat pot top kap tap kup pup tup kop	tú tó út ót ó ú -
14.	szűz (szül)	zűsz sziz szü szűsz síz szil szé	só sú fül fūt fűz főz fő fű fú fő	hó hú tó tú ó ú -
15.	ágy	ág	áj aj ad ód	tó ót út út a ó -
16.	sző	ször szön szöm szöl szörp szūr szű szül sző	fő fű fő fú só sú ló	hó hú tő tó tú ó -
17.	bor	dur hul por pur pol pul dor	pú tú	ó ú -
18.	csók	sók csót csó só sūk csúk	fók fúk fő fú fű fut	tó tú ót út ó ú -
19.	kút	út kúk tút tú tuk puk kú	ó ú	ó -
20.	ész	tész pész méz méz ész	és ós ōs éf óf őf é	út tú ó ú -
NÉV:		Szül. idő:	Dátum:	

40. ábra

Tesztlap óvodai hallásméréshez

A VÁLASZLEHETŐSÉGEK FOKOZATAI

Jobb fül – zöld szín
Bal fül – piros szín

Sor-szám	I. Jó hallás	II. Hallás- vagy beszédmegértési zavar állhat fenn, esetleg figyelmetlenség. A vizsgálat megismétlendő.	III. Hallászavar valószínű, orvosi vizsgálat javasolt.	IV. Hallászavar biztosra vehető, mielőbbi klinikai vizsgálat szükséges.
21.	mák	bák nák vák	ma ba pa va mok bók pú pók muk mu bu mo bó	tó tú ó ú –
22.	csíp (síp)	csít csi sí csil csik sík sít csűp	sút fűp fűk fut fú fő hűp üt űk ki	pót út űp ót óp tú tó tút –
23.	bók	búk bóp bót pók púk but búp bó	dud dad gó du dó	tó tú ót út ú ó –
24.	szív	szín szil szék cij szű cű	fél fér fut fő fol fül fok hír hó foj fúj	ló lú tó tú ó ú –
25.	gáz	gász gál gá báz bász dáz dász záz ház	ász dág go gu bo az ág bú	tú tó ó –
26.	cél (szél)	cil szil szir cir ci zí	fél fér fő fő föl for fut fok tél tép dél fű	ló tó tú ó ú –
27.	busz	pusz pu buz bosz boz dusz	puff puff pus pos pu pul puk	tú tó ó ú –
28.	só	szó sű	fó fú hó hú	ló tó tú ó ú –
29.	zsák	zsáp zsát zsá ság sák sár sá zák zá záz	vár vá fá fa vak fát láp láb lá rák	ló lú hó hú ó –
30.	szög	szöb szöd szüg széék szűk szök szöp cöd zög	sög sök ső só fő fő fót főt fű fók só	tű tú tó ú ú hó ló –

31.	dob	dud dod dog dug bod bob	tó tú ót út	ó ő u –
32.	sín	sím szín szím szem szem sí sén sém sün	lín lím lő lű fű fő fi fő ló	hó hú ló un um tú út ú –
33.	bús	púsz pú búz bosz boz dús dusz gús	puff puff pú púl pút	tú tó ó ú –
34.	szék (szépek)	szél cél szik szil szít szí szűk szűt szök cül szé	fék fél lék fók fő fű hók hűk hó só hú fut	ló lú ó ú ol uk up –
35.	bál (Pál)	bár pár már dál mál tál bá	bor bol por pol bur pul bul pur bó pó bú pú	tó tú ó ű –
36.	cím	cíl cím cén cin szín szím szén szűm cű cí szil cib	fém fén fib fő fő fű lő hó hú hull ful fül	tó tú ó ú –
37.	gép	kép tép dép gíp díp típ	dóp góp dó bó gó	tó tú ó ú –
38.	sár	szár zsár száll sát sá	fáj fál fá fa foj	laj lap lab ló lú hó hú ó ú –
39.	gyík	gyíp gyép gyék gíp gép jík gyű gyüt	dép tép tyíp tyék zsík	gyó dó dú út úk –
40.	ősz (öz)	űsz üz isz	ős és őš ő ű is óf	őj uj oj ol ó é –

NÉV:

Szül. idő:

Dátum:

85

41. ábra

Tesztlap óvodai hallásméréshez

45. táblázat

Óvodai csoport (életkor)	Gyermekek száma			Valamennyi óvodás
	fiú	lány	összesen	
nagycsoportos (5;0--5;10)	62	63	125	
középsős (4--4;11)	37	40	77	242 fő
kiscsoportos (3;0--3;11)	26	14	40	

ciós szintjére jellemző, hogy az echolálás időszakában járt, a holofrázisos még nem jelentek meg beszédében.

Beszédészlelési adatok

A 224 éphalló gyermek visszamondásai alapján a következő megoszlást láttuk az életkor és a nemek szerint (46. táblázat).

46. táblázat

Nemek	A gyermekek helyes válaszai %-ban életkor szerint		
	kiscsoportosok	középsősök	nagycsoportosok
fiúk	60,2	65	66
lányok	64	66	67,6
összesen	62,1	65,5	66,8

A beszédészlelési szint megítélésekor nemcsak azt vettük figyelembe, hogy a 10--10 szóból hányat mond tökéletesen vissza a gyermek, hanem azt is, hogy milyen mértékben térnek el a helyestől a tévesztett variánsok. Ennek az utóbbinak a vizsgálata további különbségeket mutat az egyes gyermekek között. Az adatok alátámasztják azt a hipotézist, hogy a gyermek csak a teljes akusztikai-fonetikai-fono-

lógiai feldolgozás után kapcsol jelentést a hangsorhoz, mivel az értelmes szóanyag azonosítása nem volt nagyobb arányban helyes, mint az értelmetlené. Például: a jól ismert meggy, mos, ágy, szél szavak helyes megértése 80--90 %-os, hasonlóan jó a nagy valószínűséggel ismeretlen ász, sző, bál, cím, mák azonosítása (70--90 %). Feltételezhető, hogy ismert, mégis alacsony a gyík (32,4 %), a szög (21 %) felismerése, ugyanakkor relatíve magas a biztosan értelmetlenként észlelt szűz (48,8 %), a szú (71,7 %), a bál (30,6 %) vagy a bú (63,6 %) helyes percepciója. Természetesen könnyebb a felismerés, ha a jelentés is "segít": ezt jól példázza a gáz szó, amelyet (városi gyerekek!) már kiscsoportban is 87,5 %-ban helyesen azonosítanak.

A szavak és a logatom jellegű hangsorok felismerésére kapott válaszok a következő tendenciákat igazolják:

- az életkorral párhuzamosan fejlődik a beszédészlelés
- 3 és 4 éves kor között nagyobb a minőségi változás, mint 5 és 6 éves kor között
- nincs lényeges különbség a fiúk és a lányok teljesítménye között
- ugyanazon 10 szóval -- összehasonlítható mértékben -- nem vizsgáltuk a jobb és a bal fülbe adott szavak megértését, így a jobbfül-dominancia kérdéséhez érdemben nem tudunk hozzászólni.

Magánhangzók észlelése

A magánhangzók közül hibátlanul felismerik az elől képzetteket, valamint a leghangzósabb [a:] -t. Bizonytalanságok a hátul képzett [u, o, ɔ] -kkal vannak, ezek jobbára egymással cserélődnek fel a percepcióban.

Mássalhangzók észlelése

Zárhangok. -- A zöngés és zöngétlen felpattanó zárhangok, amelyeknek a felismerése a felnőtteknél is (más nyelvekben is) a legbizonytalanabb, a gyermekek esetében is a legkevésbé korrekt. Ha értelmes a gyermek számára a szó, akkor --várhatóan-- pontosabb az azonosítás; lényegesen pontatlanabb a logatom jellegűeknél: búr → dúr, gúr, bók → bót, bóp. A jelentés szerepét, illetőleg részvételét az azonosításban néha igen nehéz a válaszból megítélni. Gyakori a kút szó út felismerése, a gyík gép értése, a síp sít felismerése stb. Adataim alapján legbiztosabbnak a dentális felpattanó zárhangok percepcióját találtam. Jellemző tendencia, hogy a [k] -t palatális magánhangzót tartalmazó szavakban inkább [p] -nek, veláris magánhangzót tar-

talmazókban inkább [t]-nek azonosítják (gyík -- gyíp; csók -- csót). Ez nyilvánvalóan összefügg az artikulációval is: a hátul képzett magánhangzó után egyszerűbb a kisebb nyelvváltoztatást igénylő dentális zárhangot ejteni.

Az egyszeri előfordulású [n] kivételével csak a [m] található meg a szavakban. A bilabiális nazális azonosítása jó, 10 % alatti arányban tapasztalható, hogy [b, n, l] mássalhangzónak észlelik.

Réshangok és zár-rés hangok. -- A [s] percepciója csaknem hibátlan valamennyi korcsoportban. Nyilvánvalóan az értelmezéssel kapcsolatos az ősz őz értése, annál is inkább, mivel főként nagycsoportban fordul elő. A [t̂s] felismerését valószínűleg szintén szemantikai tényezők befolyásolják, semmint percepciósak: a cím-ben pontos az azonosítás, a cél szónál nagyobb arányú a feltehetően magasabb hírértékű szél értés. Az alveoláris rész- és zár-rés hangok felismerése csaknem 100 %-osan helyes. A gyermekek harmadánál megfigyelhető, hogy réshangként azonosítják a zár-rés mássalhangzót: [t̂s] -- [s]. Jó a zöngések felismerése a zsír sír értését megint jelentés-tani okokra vezethetjük vissza. Bizonytalan a szóvégi [j] azonosítása (mint általában a szóvégi mássalhangzóké).

A [l] és a [r] azonosítása. -- E két mássalhangzó percepcióját -- a visszamondások alapján -- igazában csak a nagycsoportosoknál ítélni lehetjük meg, mivel a kisebb korcsoportokban ejtésük még gyakran hibás. Pontatlan azonosításuk kismértékű, jellemzően egymás közt érthetők félre.

Az éphalló és hallássérült gyermekek tesztlapjai alapján összeállíthatjuk azokat a szempontokat, amelyek irányadók a szűrések elvégzéséhez. Ezek a következők:

- a teljesítmény az életkortól függően változik,
- meg kell győződni a gyermek beszédének tisztaságáról, hiszen az ejtésihiba még nem jelent minden esetben percepciósz zavart is,
- 6 éveseknek tíz szóból 6--7-et hibátlanul vissza kell mondaniuk, 3--4 hangsor kerülhet a második oszlopba; 5 éveseknek legalább 5 szót kell jól ismételnük (öt jelölhető a második oszlopban); 3--4 éveseknek legalább 4 szót kell pontosan reprodukálniuk, a többi kerülhet a második oszlopba,
- ha a visszamondott szavak, illetve hangsorok az első három oszlopba kerülnek, de arányuk a második és a harmadik "javára" tolódik el, valamint ha az ismétléseket a második és harmadik oszlopba jelöltük, nagy a hallásprobléma valószínűsége,

- haladéktalanul orvosi vizsgálat szükséges, ha a szavak és hangsorok jelölései a harmadik, a harmadik--negyedik vagy a negyedik oszlopba kerültek,
- beszédhibás gyermek vizsgálata a speciális képsorral végezhető el (Gósy 1985).

Az óvodáskorúak beszédészlelési szintjének vizsgálata ily módon szűrés-ként is megoldható, ami mind a (jövendő) iskolai munkában, mind a logopédiai ellátásban nagy jelentőségű.

A beszédészlelés egyes szintjeinek vizsgálata

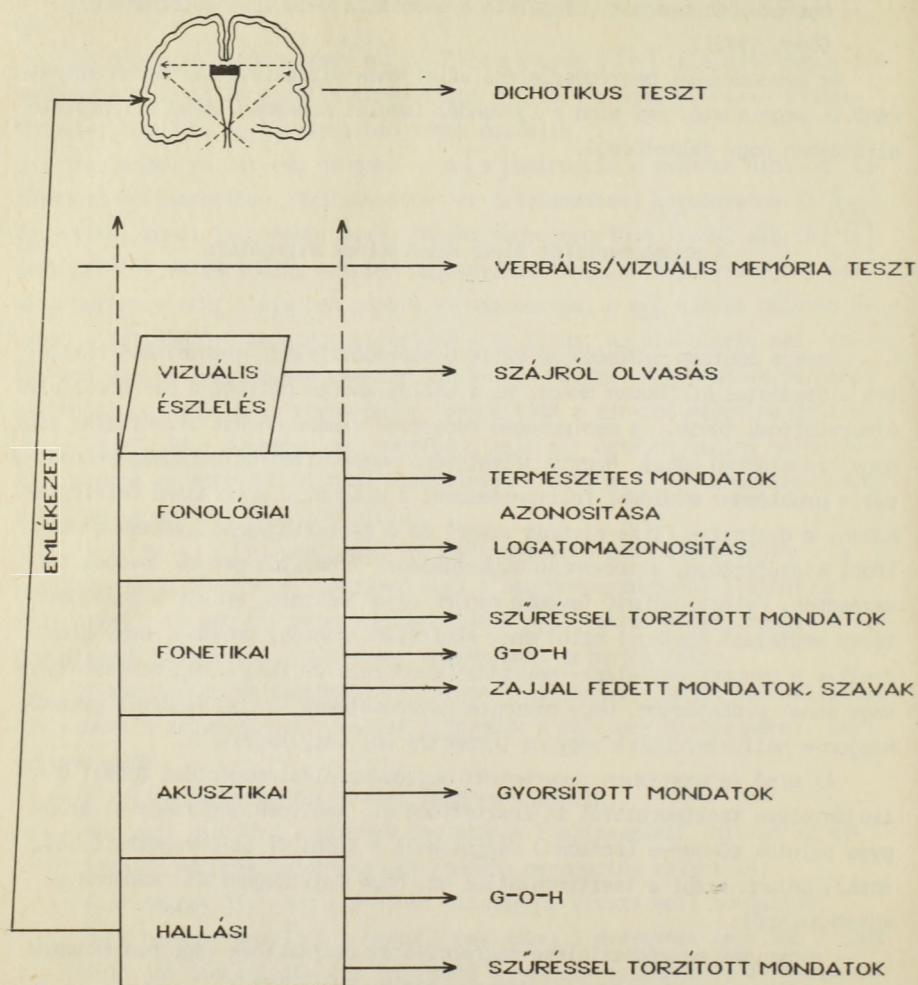
Számos esetben szükségessé válik a beszédészlelési mechanizmus részletes vizsgálata; különösen akkor, ha a hallás épsége mellett a beszédészlelés elmaradottnak tűnik, a mesterséges hangsorok visszamondása bizonytalan vagy nagy százalékban hibás. Nagy a jelentősége ennek bizonyos betegségek (centrális problémák) mielőbbi felismerésében, a diagnózis egyértelmű felállításaiban; a dyslexiás (írás-olvasás zavar) és a dyslexiagyánús gyermekek mielőbbi kiszűrésében, a prevenció megkezdésének irányításában és további tervezésében. Felhasználható továbbá minden olyan esetben, amikor a beszédészlelés megfelelő életkori szintjének megítélése fontos, például, hogy elég érett-e a gyermek észlelési--megértési rendszere az iskola megkezdéséhez, a vagy annak eldöntésére, hogy bizonyos határesetekben a hallássérült gyermek kapjon-e hallókészüléket vagy ne (utóbbira vö. Gósy 1989).

Az első fejezetekben ismertetett beszédészlelési/megértési modell alapján olyan teszt sorozatot fejlesztettünk ki, amelynek segítségével az egyes szintek működése (részben) függetlenül a többitől is vizsgálható (42. ábra). Célunk ezzel a teszt sorozattal az, hogy felvilágosítást kapjunk a következőkről:

- a gyermek beszédészlelése megfelel-e az életkorának vagy nem (korspecifikus normák meghatározásának igénye: Speak 1984),
- a beszédészlelési mechanizmusban hol jelentkezik elmaradás vagy zavar,
- az elmaradás vagy zavar milyen mértékű és kiterjedésű,
- milyen kompenzáló stratégiákat alkalmaz a gyermek,
- melyek a korrigálás lehetőségei (egyfelől metodológiai, másfelől szervezési szempontból: otthoni vagy logopédiai segítség).

MODELL

TESZTEK



42. ábra

A hierarchikus észlelési modell szintjei és a
tesztrendszer összefüggése

A beszédészlelés és a beszédmegértés elmaradásának vagy zavart voltának egész életre kiható következményei lehetnek. A legjelentősebbeket az alábbiakban foglaljuk össze:

- beszédképzési nehézségek,
- gondolkodási, logikai korlátok,
- dyslexia,
- tanulási, feladatmegoldási nehézségek,
- kommunikációs kudarcok,
- magatartási problémák,
- személyiségzavarok.

A beszédészlelési szint megállapítását célzó tesztsorozatot a beszédészlelési kísérletek eredményeinek ismeretében terveztük meg. A közvetlen célú adaptálás egyidejű kísérletek lefolytatásával, az újabb eredmények integrálásával történt. A tesztsorozattal a hallástól a teljes beszédmegértésig elemezhető a gyermek teljesítménye 3 éves kortól (a továbbiakban csak a beszédészlelésre vonatkozó tesztrendszer ismertetjük). Az anyagot úgy alakítottuk ki, hogy 7 éves korra az egyes résztesztekben elérendő átlagteljesítmény 100 % legyen ép működés esetén. Korcsoportonként 50 gyermekből álló (szociológiailag heterogén) csoportot vizsgáltunk meg, összesen 250 3-7 év közötti gyermeket. A kísérleteket csendes térben egyénileg végeztük a gyermekekkel.

A véglegesre kialakított tesztsorozat 10 altesztből áll (beleértve a G-0-H-val történő méréseket is). A vizsgálat időtartama 20 perc, ez az adott gyermek életkorától és kooperációs készségétől függően kismértékben változhat.

A beszédészlelési tesztsorozat

A tesztsorozat ismertetésével egyidejűleg közöljük az ép 3-7 éves gyermekek átlageredményeit.

1. A hallás ép működésének megítélését a G-0-H készülékkel végezzük. (Ennek hiányában 6 m-es sügött beszéddel tájékozódunk a gyermek hallásának épségéről.)

2. Az észlelés globális megítélése ugyancsak a G-0-H készülékkel történik az ott megadott szempontok és életkori eltérések figyelembevételével, külön a jobb és külön a bal fülbe adott mesterséges szavak felismerése alapján.

3. Az észlelés részletes vizsgálata zajjal elfedett 10 bővített magyar mondat megismételtetésével, visszamondatásával kezdődik. A zaj/jel viszony 4 dB. A mondatokat férfi bemondóval rögzítettük magnetofonszalagra (vö. 47. táblázat).

47. táblázat

Sorszám	Zajjal elfedett mondatok
1.	A sütemény nagyon finom volt.
2.	A rádióban zene szól.
3.	Terítsétek meg az asztalt!
4.	A repülőgép most szállt le.
5.	Menjünk holnap kirándulni?
6.	Az őzikét kergeti az oroszlán.
7.	Rakjátok össze a játékokat!
8.	A strand ma be van zárva.
9.	Ki akar lemenni vásárolni?
10.	Tavasszal sokat esik az eső.

Az ép óvodásokkal végzett kísérletek eredményei szerint e 10 mondat korrekt felismerése átlagosan a következő: 3 éves korban 30 %, 4 éves korban 55 %, 5 éves korban 70 %, 6 éves korban 90 %. A százalékos adatok azt mutatják, hogy a tíz mondatból átlagosan hány mondatot képesek a gyermekek az adott életkorokban hibátlanul felismerni. Bármilyen kis hiba esetén (ahol egyértelműen észlelési és nem produkciós hibáról van szó) a mondatazonosítást tévesnek kell minősítenünk. (A gyakorlati alkalmazásban természetesen nagy a jelentősége a kis hibával vagy kevés tévesztéssel felismert és ismételt mondatoknak is.)

4. A szófelismerést hasonló módszerrel, fehér zajjal elfedett izolált szavak megismételtetésével vizsgáljuk. A zaj/jel viszony ugyancsak 4 dB. A felhasznált szavak: oroszlán, meggy, csillag, kendő, szita, csörgő-kígyó, eper, száj, étterem, ablak. A szavak "fonetikailag kiegyenlítették", azaz a lehetőségeken belül követik a magyar beszéd beszédhang- és hangkapcsolat-gyakoriságát (a szótagszám is a magyar beszédre jellemző arányokat mutat). A gyermekek észlelésére -- mint korábban láttuk -- jellemző, hogy

a részletek megfigyelése és elraktározása döntő fontosságú. Ebből az következik, hogy az önállóan előforduló szavak azonosítása valamivel könnyebb feladat, mint például a mondatok felismerése. A kapott eredmények 3 éves korban: 35 %, 4 éves korban: 60 %, 5 éves korban: 75 %, 6 éves korban: 100 %.

5. A felismerési kulcsok észlelésének biztonságát vizsgáljuk frekvenciaszűrővel torzított 10 bővített mondat felismertetésével. Az ugyancsak férfi bemondóval hangszalagra rögzített mondatokat úgy szűrtük meg, hogy 2200 és 2700 Hz között maradjon csak jel, a levágási meredekség: 36 dB/oktáv. A mondatokat a 48. táblázat tartalmazza. Ép hallás és ép beszédészlelés esetén a gyermekek teljesítménye 3 éves korban 80 %, majd az életkor növekedésével fokozatosan eléri a 100 %-os pontos felismerést, vö. 43. ábra grafikonját. Ha a vizsgált gyermek eredménye itt lényegesen gyengébb a közölt-

48. táblázat

Sorszám	Frekvenciaszűrővel torzított mondatok
1.	A vonat nyolc óraker indul.
2.	Mikor lesz az esküvője?
3.	A tejfölt a macska itta meg.
4.	Kapcsold be a televíziót!
5.	A munkások estig dolgoznak.
6.	Holnap moziba megyünk.
7.	Megint leszakadt a gombod?
8.	Dobd a papírt a szemétkosárba!
9.	A kulcs a zsebemben van.
10.	Télen a medvék a barlangjukban alszanak.

nél, akkor komoly észlelési problémára kell gyanakodnunk. Kontrollként grammatikai és szemantikai tekintetben a torzítottéhoz hasonló 10 mondatot -- ugyanazon bemondóval -- torzítás nélkül rögzítettünk magnetofonszalagra. Ennek a 10 mondatnak a felismerése kivétel nélkül 100 %-os már 3 éves kortól. Abban az esetben, ha a gyermek teljesítménye -- azonos lejátszási hangereő mellett -- a természetes ejtésű (torzításmentes) mondatokra gyengébb, mint a frekvenciaszűrővel torzítottakra, akkor nagy valószínűséggel hallászavar áll fenn.

6. Az észlelési mechanizmus időzítési sajátosságait vizsgáljuk 10 kismértékben gyorsított mondat felismertetésével (a mondatokat a 49. táblázat tartalmazza). Tekintettel arra, hogy a beszéd időszerkezetének felbomlása sajátos működésre készíteti az észlelési rendszert, a gyorsított mondatok visszamondatásával hű képet kapunk az ép vagy zavart mechanizmusról.

49. táblázat

Sorszám	Gyorsított mondatok
1.	Az irígység rossz tulajdonság.
2.	Őt is beidéztek a tárgyalásra?
3.	A forgalmat rendőrök irányítják.
4.	Ne gyártsatok selejtet!
5.	A galamb a szabadság jelképe.
6.	A katonák felesküdték a zászlóra.
7.	Ki akar számot adni a munkájáról?
8.	Átkokat szórt mások fejére.
9.	Gyorsan megittatták az állatokat.
10.	Fejtsétek ki a véleményeteket!

Eltérően az eddigi mondatcsoportoktól, az itt szereplő mondatok tartalma több esetben -- elsősorban a jelentés, a szóhasználat és a szintaktikai felépítés következtében -- nem vagy csak részlegesen érthető a gyermekek számára. Varyspeech segítségével a mondatokat az eredeti tempójuk 1,4-szeresére gyorsítottuk; a gyorsítás mértéke gyakorlatilag nem haladja meg az átlagos spontán beszéd sebességét. Kísérleti adataink szerint a mondatok tökéletes észlelése 3 éves korban 30 %, 4 éves korban 55 %, 5 éves korban közel 80 %, 6 éves korban pedig csaknem 100 %.

7. A rövid idejű memóriát mind a verbális, mind a vizuális teljesítmény szempontjából vizsgáljuk (szó-, illetve képsorral).

a) A verbális emlékezet megítéléséhez a következő 10 szót használjuk: mackó, labda, hold, óra, hóember, ház, szív, egér, csónak, cipő, hal, tojás. Az életkor változásával alig jelentkezik különbség a teljesítményben; adataink alapján kimondható, hogy 4 éves kortól a szakirodalomban közölt átlagérték az irányadó (Richardson 1983). Ez minimálisan 5 szó felidézését jelen-

ti a -- jelen esetben -- 12 közölt szóból. Az átlagos (életkortól független) felidézhetőség: 7 ± 2 elem. Az óvodáskorúak vizsgálatával kapott eredményeink a következők: a 3 éveseknél 4,3 elem, az 5 éveseknél 5,7 elem, a 6 éveseknél 5,1 elem.

b) A vizuális memória vizsgálatát hasonló módszerrel végeztük; ekkor egy kartonra ragasztott 12 kép megnézetése után kell a gyermeknek visszaidéznie a látott képeket. A képeknek megfelelő szavak a következők: csillag, béka, vár, rendőr, torta, síp, dinnye, kút, malom, kakas, hegedű, szék. Az eredmények ugyancsak alig változnak 4 éves kortól (megközelítik a 7 ± 2 -es átlagot): 3 éves korban 4,3 elem, 5 éves korban 5,3 elem, 6 éves korban 6,2 elem.

A visszaidézett elemek (szavak, illetve képek megnevezései) számán kívül fontos annak megfigyelése, hogy az ismétlés milyen sorrendben történik: rendezett-e a felidézés vagy nem. Rendezettségéről akkor beszélünk, ha a közölt elemsor első és utolsó tagjai idéződnek fel elsőként, s csak ezeket követően a középen elhelyezkedők. Jellemző életkori változás, hogy a kisebb gyermekek elsőként a frissen (tehát utolsóként) elhangzott/látott elemeket idézik fel. Az életkor növekedésével az elsőként közölt/látott elemeket idézik fel elsőként, ezt követik az utolsók, majd végül a középsők. Ezt az eltolódást mutatja az életkor függvényében az 50. táblázat az első felidézett szóra kapott százalékos értékekkel.

50. táblázat

Közölt szavak	Életkori csoportok adatai (%)				
	3 évcsek	5 évesek	6 évesek	7 évesek	12 évesek
1. mackó	15	18	49	70	77
2. labda	7	6	5	5	0
3. hold	3	0	0	0	0
4. óra	7	0	0	0	0
5. hőember	0	0	0	5	0
6. ház	7	9	4	5	0
7. szív	0	0	0	0	0
8. egér	3	0	6	0	0
9. csónak	7	0	0	0	0
10. cipő	0	12	8	0	0
11. hal	7	18	6	5	0
12. tojás	41	34	17	10	23

8. Az a j a k m ű k ö d é s v i z u á l i s é s z l e l é s é n e k v i z s g á l a t á t (száj-ról olvasás) célozza a következő alteszt. Állatok neveit közöljük a gyermekkel pusztán a szájomozgás alapján (előzetesen felhívja a figyelmet arra, hogy állatneveket kell majd "kitalálnia"). A felhasznált állatnevek a következők: macska, oroszlán, elefánt, mókus, egér, kenguru, viziló, tehén, darázs, liba. A teljesítmény erősen függ az életkortól; adott esetben -- különösen 3 éveseknél -- az esetleg ismeretlen állatnevek is okozhatnak nehézséget. 5 éves kortól legalább 4-5 szót fel kell a gyermeknek ismernie, a táblázati arány az életkorral nő. Az igen gyenge teljesítmény hátterében leggyakrabban ingerszegény környezet, hibás artikulációjú szülők, illetőleg szegényes szókincs áll. A feltűnően jó teljesítmény hallásproblémára utal!

9. A lateralizációt dichotikus teszttel vizsgáljuk (vö. Pléh 1981), amelyben 1-1, illetőleg 2-2 különböző szót hall a gyermek egyidejűleg a két fülében. Feladata annyinak az ismétlése, amennyire képes. A felhasznált szavakat az 51. táblázat tartalmazza.

51. táblázat

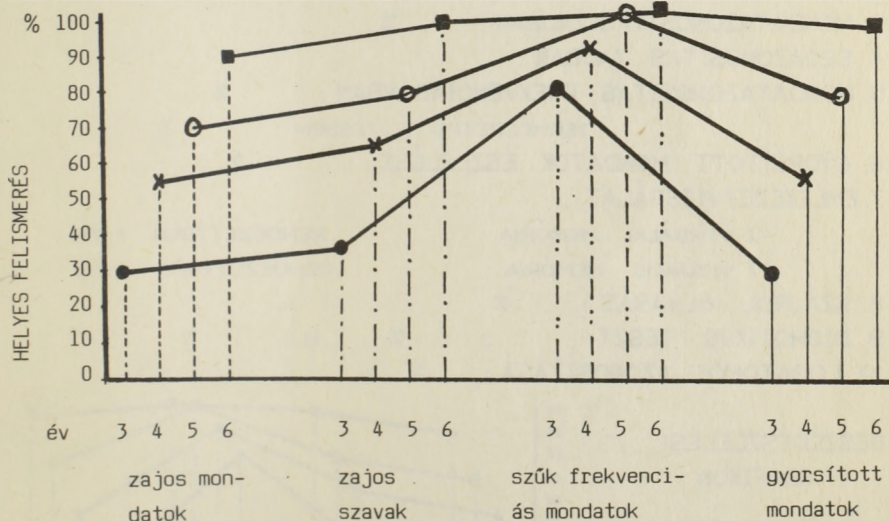
A dichotikus teszt szóanyaga	
bal fülbe közöltek	jobb fülbe közöltek
vár	ház
bolt	gomb
kacsa	kutya
majom	malac
kenyér	körte

madár, kacsa	majom, kutya
malac, kecske	mackó, teve
mókus, cica	majom, csirke
bogár, róka	bárány, liba
ajtó, kenyér	ablak, körte

A nagyobb mértékben egyik fülben hallott szavak visszamondása jelzi a félteke-dominanciát, amely különösen a bal-, illetőleg a kétkezes gyermekek beszédészlelési szintjének megítéléséhez fontos információ.

10. A fonetikai és a fonológiai szintű észlelés vizsgálatát segíti a logatomok ismételtetése. A visszamondás alapján megítélhető, hogy az anyanyelvi fonotaktikai szabályok milyen mértékben érvényesülnek már a gyermek beszédészlelési rendszerében. A felhasznált logatomok a következők: galalajka, zseréb, trankün, veréndekek, bakógy. Kiscsoportos óvodás korban még gyakori a tévesztés, különösen a hangsorépítési szabályoknak ellentmondó logatomok ismétlésekor (pl. trankün). 6 éves korban a visszamondás közel hibátlan.

A tesztsorozat eredményeit egy erre a célra tervezett lapon jelöljük (44. ábra). A grafikon az ép 3--6 éves gyermekek átlagos teljesítményét mutatja, amely mindenkor útmutatóul szolgálhat a logopédus (óvónő, tanítónő) számára az eredmények értékeléséhez (43. ábra).



43. ábra

Gyermekek beszédészlelési teljesítményének grafikonja
(a tesztrendszerrel vizsgálva)

(3 évesek: ●—●, 4 évesek: x—x,
5 évesek: o—o, 6 évesek: ■—■)

BESZÉDÉSZLELÉSI EREDMÉNY

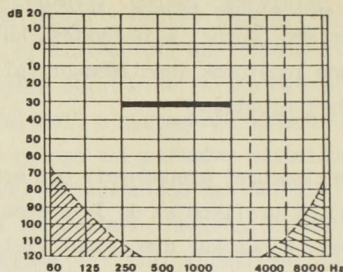
NÉV :

DÁTUM:

SZÜLETÉSI IDŐ :

ELŐZMÉNY :

1. HALLÁS (G-O-H) :



2. ÉSZLELÉS (G-O-H) :

JOBBI : %

BAL : %

3. MONDATAZONOSÍTÁS ZAJBAN : %

4. SZÓAZONOSÍTÁS ZAJBAN :

5. MONDATAZONOSÍTÁS FREKVENCIASÁVBAN : %

TERMÉSZETES EJTÉSSEN : %

6. GYORSÍTOTT MONDATOK ÉSZLELÉSE : %

7. EMLÉKEZETVIZSGÁLAT

a) VERBÁLIS MEMÓRIA :

RENDEZETTSÉG : I/N

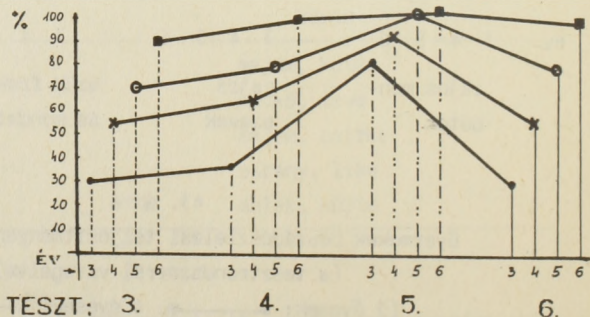
b) VIZUÁLIS MEMÓRIA :

RENDEZETTSÉG : I/N

8. SZÁJRÓL OLVASÁS : %

9. DICHOTIKUS TESZT : J : % B : %

10. LOGATOMOK AZONOSÍTÁSA : %

BESZÉDÉSZLELÉSI
GRAFIKON :BESZÉDPRODUKCIÓ :
JAVASLAT :

44. ábra

BEFEJEZÉS

A szokásos befejező gondolatokat és a végső összegzést a t á r g y -
m u t a t ó hivatott helyettesíteni.

*

A beszédészlelés a teljes beszédmegértési folyamat része; ennek a mű-
ködését érintő valamennyi kérdés megválaszolása, a hipotézisek mind egyér-
telműbb igazolása -- további kutatási feladat.

- Abbot, G.: A new look at phonological 'redundancy'. *ELT Journal* 40/4, 1986, 299--305.
- Ainsworth, W.A.: Perception of stop consonants in synthetic CV syllables. *L & S LI*, 1968, 139--156.
- Ainsworth, W.A.: Duration as a cue in the recognition of synthetic vowels. *JASA LI*, 1972, 648--651.
- Ainsworth, W.A.: Co-operative vowels and competitive consonants? *ISPS 9*. 1979, Vol. I. 219--220.
- Barton, D.: Phonemic discrimination and the knowledge of words in children under 3 years. *Rap. Rep. Child Lang. Dev.* 11. 1976, 61--68.
- Bauer, H.: Phoniatische Therapie der Gaumenspalten Sprache unser besondere Berücksichtigung neuerer Erkenntnisse. *Folia Phoniatr.* 24. 1972, 388--400.
- Beasley, D.S.--Beasley, D.C.: Auditory reassembly abilities of black and white first- and third-grade children. *J. of Speech and Hearing Res.* 16. 1973, 213--221.
- Beasley, D.S.--Bratt, G.W.--Rintelmann, W.F.: Intelligibility of time-compressed sentential stimuli. *J. of Speech and Hearing Res.* 23. 1980, 722-732.
- Beasley, D.S.--Maki, H.E.: Time- and frequency-altered speech. In: *Contemporary Issues in Experimental Phonetics*. Ed. N.J. Lass. New York, San Francisco, London 1976, 419--458.
- Békésy, G.: *Experiments in Hearing*. New York, Toronto, London 1960.
- Bever, T.G.--Lackner, I.--Kirk, R.: The underlying structure sentence is the primary unit of immediate speech processing. *Percep. Psychophys.* 5. 1969, 225--234.
- Blom, J.G.: The application of information theory to vowel-recognition experiments. *Int. Cong. Phon. Sci.* VI. Praha 1967, 189--196.
- Blumstein, S.E.--Stevens, K.N.: Perceptual invariance and onset time spectra for stop consonants in different vowel environments. *JASA* 67. 1980, 648--661.
- Boers, P.M.: Formant enhancement of speech for listeners with sensorineural hearing loss. *IPO Annual Progress Report* 15. 1983, 21--29.
- Bolla Kálmán: A nyelvészeti fonetika szakágazatai. *NyK* 74. 1976, 292--299.
- Bolla Kálmán: A magyar magánhangzók akusztikai analízise és szintézise. *Magyar Fonetikai Füzetek* 1. 1978, 53--78.

- Bolla Kálmán (szerk.): Fejezetek a magyar leíró hangtanból. Budapest 1982.
- Bolla Kálmán: Magyar hangalbum. In: Fejezetek a magyar leíró hangtanból. Szerk. Bolla Kálmán. Budapest 1982, 165--175. oldalak között.
- Bondarko, L.V. et al.: A model of speech perception in humans. Working Papers in Linguistics 6. Computer and Information Science Research Center, Ohio State Univ. Columbus, Ohio. Technical Report 1970, 70--92.
- Borovičková, B.--Malác, V.: Towards the basic units of speech from the perception point of view. ZPhon XXI, 1968, 83--88.
- Büky Béla--Egyed András--Pléh Csaba: Nyelvi képességek -- Fogalomkincs-- Megértés. Budapest 1984.
- Carterette, E.C.--Møller, A.: The perception of real and synthetic vowels after very sharp filtering. Proc. Speech Communication Seminar - 1962. II. Stockholm: Speech Transmission Laboratory, Royal Institute of Technology 1963.
- Cherry, C.: On Human Communication: A Review, A Survey, and a Criticism. Massachusetts 1957.
- Cherry, C.: Some experiments on the recognition of speech with one and two ears. JASA XXV, 1953, 975--979. Idézi: Csisztovics-Kozsevnnyikov 1965.
- Chilla, R.--Gabriel, P.--Kozielski, O.: Die Göttinger Kindersprachverständnistests. Die Sprachheilarbeit 1. 1977, 1--11.
- Chistovich, L.A.: Auditory processing of speech. L & S 23. 1980, 67--74.
- Clark, H.H.--Clark, E.V.: Psychology and language. New York, Chicago, San Francisco, Atlanta 1977.
- Cohen, A.--t'Hart, J.: Speech synthesis of steady-state segments. Proc. Comm. Sem. II. Stockholm 1963.
- Cooper, F.S.--Delattre, P.C.--Liberman, A.M.--Borst, J.M.--Gerstman, L.J.: Some experiments on the perception of synthetic speech sounds. JASA 24. 1952, 597--606.
- Cutting, J.E.--Pisoni, D.B.: An information-processing approach to speech perception. In: Speech and Language in the Laboratory, School and Clinic. Ed. J.F. Kavanagh, W. Strange. Cambridge, Massachusetts and London 1978, 38--73.
- Cutler, A.--Carter, D.: The prosodic structure of initial syllables in English. In: European Conference on Speech Technology. Vol.I. Szerk. Laver, J.-Jack, M.A. Edinburgh 1987, 207--210.
- Csányi Yvonne: A siketek általános iskolája; tantárgypedagógia. Budapest 1981.
- Csisztovics, L.A.--Kozsevnnyikov, V.A.: Реč, артикулјација и воспријатије. Moskva--Leningrad 1965.

- Darwin, C.J.: The perception of speech. In: Handbook of Perception. Szerk. Carterette, E.C.--Priedman, M.P. New-York, San Francisco, London 1976, 175--227.
- Delattre, P.: Acoustic or articulatory invariance? *Glossa* 1. 1967, 3--25.
- Delattre, P.--Liberman, A.--Cooper, F.: Acoustic loci and transitional cues for consonants. *JASA* 27. 1955, 769--773.
- Denes, P.B.--Pinson, E.N.: The Speech Chain: The Physics and Biology of Spoken Language. New York 1973.
- Dittman, A.T.--Llewellyn, L.G.: The phonemic clause as a unit of speech decoding. *Journal of Personality and Social Psychology* 9. 1967, 341--346.
- Dorman, M.F.--Raphael, L.J.: Distribution of acoustic cues for stop consonant place of articulation in CVC syllables. *JASA* 67. 1980, 1333--1336.
- Dzsaparidze, Z.N.: Ob etalonah perceptivnoj bazi jazika. *Voprosi analiza reci* IV. Tbiliszi 1973, 130--142.
- Dzsaparidze, Z.N.: Perceptivnaja baza jazika i vopros o kategorialnosti vosprijatija zvukov reci. *Uconije Zapiski* 60. 1971, 250--257.
- Eilers, R.E.: Context-sensitive perception of naturally produced stop and fricative consonants by infants. *JASA* 61. 1977, 1321--1336.
- Eilers, R.E.: Infant speech perception: history and mystery. *Child Phonology* 1980, 23--29.
- Eilers, R.E.--Wilson, W.R.--Moore, J.M.: Developmental changes in speech discrimination in infants. *JSHR* 20. 1977, 766--780.
- Eimas, P.D.--Siqueland, E.--Juszyk, P.--Vigorito, J.: Speech perception in infants. *Science* 171. 1971, 303--306.
- Fant, G.: Models of speech perception. *ZPhon* 21. 1968, 5--9.
- Fant, G.: Analysis and synthesis of speech process. IN: *Manual of Phonetics*. Szerk. Malmberg, B. Amsterdam 1968, 173--278.
- Fant, G.: Stops in CV syllables. *Speech Transmission Laboratory, QPRS* 4. Stockholm 1969, 2--25.
- Fant, G.: *Speech Sounds and Features*. Cambridge, Massachusetts and London 1973.
- Farkas Zsolt--Gósy Mária--Hirschberg Jenő: Kisgyermekkori beszéd-audiometria. *Fül--orr--gégegyógyászat* 29. 1983, 93--102.
- Fára, M.--Sedláčková, E.--Klasková, O.--Shrivnavková, J.--Kmelová, A.--Šupáček, J.: Primary pharyngofixation in cleft-plate repaire. Survey of 46-year experience with and evaluation of 2073 cases. *Plast*.

Rekonstr. Surg. 45. 1970, 449--458.

- Feldstein, S.F.--R.N. Bond: Perception of speech rate as a function of vocal intensity and frequency. *Language and Speech* 24/4, 1981, 387--395.
- Flanagan, J.L.: *Speech Analysis, Synthesis and Perception*. Berlin--Heidelberg, New York 1965.
- Fleischer, K.: *Hals-Ohren-Heilkunde für das Krankenpflegepersonal*. Stuttgart 1976.
- Fletscher, H.: *Speech and Hearing in Communication*. New York 1919.
- Fónagy Iván: *A hangsúlyról*. Bp. 1958.
- Fónagy Iván: Electro-physiological correlates of stress and stress perception. *J. Speech and Hearing Research* LX, 1966, 231--244.
- Fónagy, I.--Magdics, K.: Speed of utterance in phrases of different lengths. *Language and Speech* 1960, III, 179--193.
- Fónagy Iván--Szendé Tamás: Zárhangok, réshangok, affrikáták hangszínképe. *NyK* LXXI, 1969, 281--343.
- Foulke, E.--T.G. Sticht: A review of research on time compressed speech. *Psychological Bulletin* 72, 1969, 50--62.
- Fourcin, A.J.: Acoustic patterns and speech acquisition. In: *The Development of Communication*. Szerk: Waterson, N.--Snow, C. Chichester, New York, Brisbane, Toronto 1978, 47--72.
- Foss, J.D.: Decision processes during sentence understanding. *J. Verb. Learning and Verb. Behavior* 8. 1969, 452--462.
- Fry, D.B.--Abramson, A.D.--Eimas, P.D.--Liberman, A.M.: The identification and discrimination of synthetic vowels. *L & S* 5. 1962, 171--189.
- Fry, D.B.--Denes, P.: The solution of some fundamental problems in mechanical speech recognition. *L & S* 1. Part. 1. 1958, 35--58.
- Fujimura, O.: Analysis of nasal consonants. *JASA* 34. 1962, 1865--1875.
- Fujimura, O.: On the second spectral peak of front vowels: a perceptual study of the role of the second and third formants. *L & S* 10. 1967, 181--193.
- Fujimura, O.: Syllable as a unit of speech recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. Vol. ASSP-23, No. 1. February 1975, 82--87.
- Fujimura, O.--Macchi, M.J.--Streeter, L.A.: Perception of stop consonants with conflicting transitional cues: a cross-linguistic study. *L & S* 21. 1978, 337--346.

- Fujisaki, H.: Some remarks on recent issues in speech-perception research. L & S XXIII, 1980, 75--80.
- Fujisaki, H.--Kawashima, T.: On the modes and mechanisms of speech perception. Sogoshikenjo-Nenpo Vol. 28. 1969, 69--74.
- Fujisaki, H.--Kawashima, T.: A quantitative model for the mechanisms of discriminating speech sounds. 7th Int. Cong. on Acoustics. Budapest 1971, 21 H 2, 433-437.
- Fülöp Géza: Ember és információ. Budapest 1983.
- Gerstman, L.J.: Noise duration as a cue for distinguishing among fricative, affricate, and stop consonants. JASA 28. 1956, 982--986.
- Gimson, A.C.: An introduction to the pronunciation of English. London 1962.
- Glucksberg, S.--Danks, J.H.: Experimental Psycholinguistics, An introduction. New York, Toronto, London, Sydney 1975.
- Goldman-Eisler, F.: On the variability of the speed of talking and its relation to the length of utterances in conversations. British Journal of Psychology 45/2. 1954, 92--98.
- Gordos Géza--Takács György: Digitális beszédfeldolgozás. Budapest 1983.
- Gósy Mária: Az intonáció percepciója összehasonlító vizsgálatban. Magyar Fonetikai Füzetek 5. 1980, 100--109.
- Gósy Mária: A szegmentális hangszerkezet percepciójáról. Magyar Fonetikai Füzetek 8. 1981, 87--104.
- Gósy Mária: A [b,d,g] hangok percepció vizsgálat. Magyar Fonetikai Füzetek 10. 1982, 84--110.
- Gósy Mária: A beszédhangok kialakulása a gyermeknyelvben. Magyar Nyelv LXXIX, 1983, 59--71.
- Gósy Mária: Hangtani és szótani vizsgálatok hároméves gyermekek nyelvében. Nytud. Ért. 119. Budapest. 1984a.
- Gósy Mária: Szófelismerés az akusztikai szerkezet és a magyar nyelv hangsor-építési szabályainak függvényében. Magyar Nyelv LXXX, 1984b/4, 463--476.
- Gósy Mária: A beszédmegértés kezdetei. Nyelvtudományi Közlemények LXXXVI, 1984c/1, 23--25.
- Gósy Mária: A hallás és beszédmegértés kapcsolatáról. Magyar Nyelvőr CVIII, 1984d/3, 319--324.
- Gósy Mária: A kategoriális percepció kérdés tipológiai vonatkozásban. Magyar Fonetikai Füzetek 13. 1984e, 121--137.
- Gósy Mária: Beszédhibás gyermek beszédmegértésének vizsgálata. Gyógypedagógiai Szemle 2. 1985, 93--100.

- Gósy Mária: Magyar beszédhangok felismerése, a kísérleti eredmények gyakorlati alkalmazása. Magyar Fonetikai Füzetek 15. 1986, 3--100.
- Gósy, M.: High frequency speech perception: phonetic aspects and application. Proc. of Int. Cong. of Phon. Sci., Tallinn 1987, 88--92.
- Gósy Mária: Tempóészlelés és beszédmegértés. Műhelymunkák V, 1988, 87--123.
- Gósy, M.: Speech perception performance: evidence for or against a hearing aid. Proc. of Speech Research '89 Int. Conf. Budapest 1989.
- Gósy Mária: A vizuális észlelés hatása a beszédpercepcióra. Magyar Pszichológiai Szemle 5. 1989.
- Gósy, M.--Graaf, T. de: Differences in the perception of Dutch and Hungarian vowels. 1983. (megjelenés alatt)
- Gósy Mária--Olaszy Gábor: A gépi beszéd megértése. (Az UNIVOICE magyar nyelvű, azonos idejű, számítógépes rendszer percepciós vizsgálata.) Nyelvtudományi Közlemények LXXXV, 1983/1, 83--105.
- Gósy Mária--Olaszy Gábor--Hirschberg Jenő--Farkas Zsolt: Szintetizált szavak használata a beszédaudiometriában. I.--II. Fül-orr-gégegyógyászat 31. 1985,92--96, 229--233.
- Gósy, M.--Olaszy, G.--Hirschberg, J.--Farkas, Zs.: Phonetically based new method for audiometry: the G-O-H measuring system using synthetic speech. Proc. of Int. Cong. of Phon. Sci., Tallinn 1987, 185--189.
- Gósy Mária--Laczkó Mária: A beszédmegértési teljesítmény fejlődése óvodás-kortól felnőttkorig. Magyar Nyelvőr 111/4, 1987, 444--452.
- Götte Árpád: A magyar beszéd-audiometria I--II. (Kandidátusi értekezés.) Budapest 1974.
- Gutzman, H.: Untersuchungen über das Wesen der Nasalität. Archiv für Laryng. und Rhinol. 27. 1913, 59--125.
- Hadding, K.--Studdert-Kennedy, M.: An experimental study of some intonation contours. In: Intonation. Szerk. Bolinger, D. Baltimore 1972, 348--359.
- Hahlbrock, K.H.: Sprachaudiometrie. Stuttgart 1970.
- Halm Tibor: Hallástan. Budapest 1963.
- Harris, K.S.: Cues for the discrimination of American English fricatives in spoken syllables. L & S 1. 1958, 1--8.
- Hazan, V.--Fourcin, A.J.: Interactive synthetic speech tests in the assessment of the perceptive abilities of hearing impaired children. Speech, Hearing and Language 1983, 41--57.
- Hegedűs Lajos: A beszédhangok természete. Nyelvtudományi Közlemények XLIX, 1935, 337--349.
- Hegedűs Lajos: Hangnyomásmérések. Nyelvtudományi Közlemények LVI, 1954, 181--215.

- Hegedűs Lajos: Újabb vizsgálatok a magyar affrikáták köréből. Nytud.Ért. 17. 1958, 153--183.
- Hoequist, Ch.E.: Parameters of speech rate perception. AIPUK 23. Juli 1983, 99--138.
- Hoequist, Ch.E.--K.J. Kohler: Summary of speech rate perception research in Kiel. AIPUK, April 1986, 5--29.
- Horger Antal: Általános fonetika. Budapest 1928.
- House, D.: Identification of nasals: an acoustical and perceptual analysis of nasal consonants. Working Papers 22. 1982, 153--163.
- Hörmann, H.: Psycholinguistics. Berlin--Heidelberg--New York 1971.
- Janota, P.: Audiometrie řeči u dětí. Sborník vysoké školy pedagogické v Praze. Celostátní defektologická konference 1957. Státní pedagogické nakladatelství Praha 1958, 272--277.
- Jóó Tibor: Egy kiadatlan régi magyar lélektan. MPSzichSz 1929/3-4, 90--100.
- Karlgén, H.: Statistical methods in phonetics. In: Malmberg, B.: Manual of Phonetics. Amsterdam 1968, 129--155.
- Kassai Ilona: Időtartam és kvantitás a magyar nyelvben. Nytud.Ért. 102. Budapest 1979.
- Kassai Ilona: Essai pour une méthode applicable a la comparaison des systemes phonétiques et phonologiques du français et du hongrois: Studia Romanica 1974, 38--44.
- Kassai Ilona: A szünet kérdésköre a szöveglejegyzésben. In: Beszélt nyelvi tanulmányok. Linguistica, Series I. Szerk. Kontra Miklós. Budapest 1988.
- Kempelen, W. von: Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine. Wien 1791.
- Kent, R.D.: Auditory-motor formant tracking: a study of speech imitation. J. of Speech and Hearing Research 17. 1974, 202--223.
- Kenyeres Elemér: A gyermek első szavai és a szófajok föllépése. Budapest 1926.
- King, K.M.--Weston, P.L.: The effect of the percentage of time compression, sentence length, and age on children's recall performance of well-formed sentences. Paper pres. at Annual Convention of the American Speech and Hearing Association, Las Vegas, Nevada, Nov. 5--8, 1974.
- Kittel, G.: Rhinophonie und ihre ärztliche Behandlung. Sprache, Stimme, Gehör 3. 1979, 87--91.
- Kittel, G.: Dysphonie im Kindesalter. Laryng. Rhinol. Otol. 63. 1984, 208--211.

- Klatt, D.: A strategy for the perceptual interpretation of durational cues in English sentences. WP MIT, 1982, 1, 83--91.
- Kohler, K.J.: Parameters of speech rate perception in German words and sentences. AIPUK April 1986, 137--178.
- Kollár Dezső: Fül-orr-gégebetegségek gyermekkorban Budapest 1977.
- Konkle, D.F.--D.S. Beasley--F.H. Bess: A study of time-compressed speech with an elderly population. Presented Paper at the Annual Convention of the American Speech and Hearing Association, Las Vegas, Nevada, November 5--8, 1974.
- Kontra Miklós--Gósy Mária: Azonos dallam két forrásból. Műhelymunkák III, 1987, 19--30.
- Kryter, K.D.: The effects of noise on man. New York, London 1970.
- Kuhl, P.: Speech perception in early infancy: perceptual constancy for vowel categories. JASA 60. Suppl. 1. S 90. 1976.
- Kuwabara, H.: Perception of CV-syllables isolated from Japanese connected speech. L & S Vol. 25. 1982, 175--198.
- Ladefoged, P.--Broadbent, D.E.: Information conveyed by vowels. JASA 29. 1957, 98--104.
- Lafon, J.Cl.: Auditory basis of phonetics. In: Manual of Phonetics. Szerk.: Malmberg, B. Amsterdam 1968.
- Lajtha György: Rendszertechnika I. Budapest 1964.
- Lane, H.: The hue and cry concerning categorical perception (and conversely). ZPhon XXI, 1968, 109--116.
- Lassen, N.A.--Larsen, B.: Cortical activity in left and right hemisphere during language related brain functions. Proc. of 9. Int. Cong. of Phon. Sci. Vol. III. 1979, 137--145.
- Lawrence, D.L.--Byers, V.: Identification of voiceless fricatives by high frequency hearing impaired listeners. JSHD 12. 1969, 426--434.
- Laziczius Gyula: Fonétika. Budapest 1944.
- Lehiste, I.: Suprasegmentals. London 1970.
- Lehiste, I.: The units of speech perception. In: Speech and cortical functioning. Szerk. J.H. Gilbert. New York and London 1972, 187--237.
- Lehiste, I.: The many linguistic functions of duration. In: New Directions in Linguistics and Semiotics. Ed. J.E. Copeland. Houston 1984, 96--223.
- Lieberman, A.M.: Some results of research in speech perception. JASA 29. 1957a, 117--123.

- Lieberman, A.M.--Delattre, P.--Cooper, F.S.--Gerstman, L.: The role of consonant-vowel transitions in the stop and nasal consonants. *Psychol. Monographs* 68. No.8. 1954, 379--395.
- Lieberman, A.M.--Harris, K.S.--Hoffman, H.S.--Griffith, B.C.: The discrimination of speech sounds within and across phoneme boundaries. *J. Exp. Psych.* 54. 1957b, 358--368.
- Lieberman, A.M.--Harris, K.S.--Eimas, P.--Lisker, L.--Bastian, J.: An effect of learning on speech perception: the discrimination of durations of silence with and without phonemic significance. *L & S* 4. 1961, 175--196.
- Lieberman, A.M.--Cooper, F.S.--Harris, K.S.--MacNeilage, P.F.: A motor theory of speech perception. *Proc. Comm. Sem. II. Stockholm* 1963.
- Lieberman, A.M.--Cooper, F.S.--Shankweiler, D.P.--Studdert-Kennedy, M.: Perception of the speech code. *Psychological Review* 74. 1967, 431--461.
- Lieberman, Ph.: *Intonation, Perception and Language*. Cambridge, Massachusetts 1967.
- Licklider, J.C.R.--G.A. Miller: The perception of speech. In: *Handbook of Experimental Psychology*. Ed. S.S. Stevens. New York, London 1951, 1040--1075.
- Lindner, G.: *Experimentelle Phonetik*. Berlin 1969.
- Lindner, G.: *Hören und Verstehen*. Berlin 1977.
- Linell, P.: How misperceptions arise? In: *From Sounds to Words*. (Essays in Honor of Claes-Christian Elert 23 December). Umeå 1983, 179--193.
- Ling, D.: *Speech and the hearing-impaired child: theory and practice*. Washington 1977.
- Lotz, J.--Abramson, A.S.--Gerstman, L.J.--Ingemann, F.--Nemser, W.J.: The perception of English stops by speakers of English, Spanish, Hungarian and Thai: a tape cutting experiment. *L & S* 3. 1960, 71--78.
- Luria, A.R.: *Traumatic Aphasia*. The Hague 1970.
- Luriya, A.R.: *Jazik i soznanije*. Moskva 1979.
- Macchi, M.J.: Identification of vowels in isolation versus vowels spoken in consonantal context. *JASA* 68. 1980, 1636--1642.
- MacNeilage, P.: *The control of speech production*. *Child Phonology* 1980.
- Magdics Klára: *A magyar beszédhangok akusztikai szerkezete*. Budapest 1965.
- Magdics 1970 = Wodarz-Magdics 1970.
- O'Malley, M.H.--Peterson, J.E.: An experimental method for prosodic analysis. *Phonetica* XV, 1966, 1--13.

- Martini, A.--Morra, B.--Cornacchia, L.--Tobey, E.--Miller, C.: Influence of pre-voicing duration on dichotic performance. *Audiology* 22. 1983, 162--166.
- Maruszewski, M.: *Language communication and the brain*. The Hauge--Paris--Warszawa 1975.
- McNeill, D.: *The acquisition of language*. New York--Evanston--London 1970.
- Menyuk, P.: *The acquisition and development of language*. Massachusetts 1971.
- Miller, G.A.: *Language and Communication*. New York, Toronto, London 1951.
- Miller, G.A.: Decision units in the perception of speech. *IRE Transactions on Information Theory*. II-8, 81--83. 19.
- Miller, G.A.--Galanter, E.--Pribram, K.N.: *Plans and the structure of behavior*. New York 1960. Idézi: Csisztovics--Kozsevnyikov 1965.
- Miller, G.A.--Isard, S.: Some perceptual consequences of linguistic rules. *J. verb. Learning verb. Behavior* 2. No. 3. 1963, 217--228.
- Miller, G.A.--Johnson-Laird, Ph.N.: *Language and Perception*. Cambridge, London, Melbourne 1976.
- Miyawaki, K.--Strange, W.--Verbrugge, R.--Liberman, A.M.--Jenkins, J.J.--Fujimura, O.: An effect of linguistic experience: the discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics* 18. 1975, 331--340.
- Molnár József: *A magyar beszédhangok atlasza*. Budapest 1970.
- Motley, M.T.: *Nyelvbotlások*. *Tudomány* 3. 1985, 88--93.
- Neisser, U.: *Megismerés és valóság*. Budapest 1984.
- Nemser, W.: *An experimental study of phonological interference in the English of Hungarians*. Hága 1971.
- Ochiai, K.--Fujimura, O.: Vowel identification and phonetic context. *The University of Electro-Communications, Chofu, Tokyo*, 22--2, December, 1971, 103--111.
- Oden, C.G.--Massaro, D.W.: Integration of featural information in speech perception. *Psychological Review* Vol. 85. No 3. 1978, 172--191.
- Olaszy Gábor: *A magyar mássalhangzók és a mássalhangzó--magánhangzó-kapcsolódások akusztikai szerkezetének analízise és szintézise*. *Magyar Fonetikai Füzetek* 10. 1982, 46--84.
- Olaszy Gábor: *A magyar beszéd leggyakoribb hangsorépítő elemeinek szerkezetének és szintézise*. *Nyud.Ért.* 121. Budapest 1985.
- Onishi, M.: Perception of quality against quantity. *Int. Cong. of Phon. Sci.* Praha 1967, 713--715.

- den Os, E.: Perception of speech rate of Dutch and Italian utterances. *Phonetica* 42, 1985, 124--134.
- Osgood, Ch.--Sebeok, T.A.: *Psycholinguistics*. Bloomington--London 1965.
- Osmanné Sági Judit--Erdélyi Alisza: *Mi a neuropszichológia?* Budapest 1982.
- Osser, H.--F. Peng: A cross-linguistic study of speech rate. *Language and Speech* 7, 1964, 120--125.
- Pauka Károly: A beszéd megértése. In: *Fejezetek a leíró magyar hangtanból*. Szerk. Bolla Kálmán. Budapest 1982.
- Pauka Károly: *Halláslélektan -- a beszédmegértés alaptényezői*. Budapest 1980.
- Pettersen, T.: Electroconvulsive therapy and the lateralization of linguistic function. *Working Papers* 26. 1983.
- Pisoni, D.: Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics* Vol. 13. 2. 1973. 253--260.
- Pisoni, D.: Auditory short-term memory and vowel perception. *Memory and Cognition* Vol. 3. (1) 1975, 7--18.
- Pisoni, D.B.: Some current theoretical issues in speech perception. *Cognition* 10. 1981, 249--259.
- Pisoni, D.B.--Sawusch, J.R.: Some stages of processing in speech perception. In: *Structure and Process in Speech Perception*. Szerk. Cohen, A.--Nooteboom, S. Heidelberg 1975, 16--34.
- Pléh Csaba: Mondat és emlékezet. *MPSzichSz* 31. 1974.
- Pléh Csaba: A visszautaló (anaforikus) nyelvi eszközök értelmezése. *MPSzichSz* 3. 1981a, 215--230.
- Pléh Csaba: The role of word order in the sentence interpretation of Hungarian children. *Folia Linguistica* V, 1981b, 332--343.
- Pléh Csaba: Különböző szórendű mondatok értelmezése és a dichotikus aszimmetriák 3--6 éves gyermekeknél. *MPSzichSz* 3. 1981c, 365--395.
- Pléh Csaba: A megértés és a szövegalkotás pszichológiája. In: *Büky Béla--Egyed András--Pléh Csaba: Nyelvi képességek -- Fogalomkincs -- Megértés*. Budapest 1984, 271--375.
- Price, P.J.--H.J. Simon: Perception of temporal differences by two normal-hearing groups: the effect of age and level of presentation. *MIT WP* 1983, 129--149.
- Port, R.F.: Invariance in phonetics. In: *Invariance and Variability of Speech Processes*. Szerk.: Perkell, J.--Klatt D. Hillsdale 1984, 540--558.

- Rahko, T.--Karjalainen, M.A.--Laine, V.K.--Lavonen, S.: Speech audiometry by a speech synthesizer. I. A preliminary report. Arch. Otorhinolaryngol. 222. 1979, 85--89.
- Ranschburg Pál: Az értelem (Az emberi elme). Budapest 1923.
- Remington, R.: Processing of phonemes in speech: a speed-accuracy study. JASA LXII, 1977, 1279--1290.
- Richardson, S.O.: Differential diagnosis in delayed speech and language development. Folia Phoniatr. 35. 1983, 66--80.
- Rietveld, A.C.M.--Gussenhoven, C.: Perceived speech rate and intonation. J. of Phonetics 15, 1987, 273--285.
- Risberg, A.--Agelfors, E.: On the relation between frequency discrimination ability and the degree of hearing loss. STL-QPSR 1984/4, 59--71.
- Sarbó Artúr: A beszéd összes vonatkozásaiban, különös tekintettel a gyermekkorra. Budapest 1906.
- Saussure, F.de: Bevezetés az általános nyelvészetbe. Budapest 1967.
- Sawusch, J.R.--Pisoni, D.B.: Response organization in selective adaptation to speech sounds. Perception and Psychophysics Vol. 20 (6) 1976, 413--8.
- Scheffers, M.T.M.: The role of pitch in perceptual separation of simultaneous vowels. IPO Annual Progress Report 14. 1979, 51--55.
- Scholes, R.J.: Perceptual categorization of synthetic vowels as a tool in dialectology and typology. L & S XI, 1968, 194--107.
- Schüller Gabriella--Cs.Dévai Margit--Kádár Júlia: A beszédészlelés modelljei. MPSzichSz XXIV, 1967/2, 243--245.
- Scott, B.L.: Temporal factors in vowel perception. JASA 60. 1976, 1354--1365.
- Scott, B.L.: Speech as patterns in time. In: Perception and Production of Fluent Speech. Ed. R.A. Cole. Hillsdale, New Jersey 1980, 51--75.
- Selnes, O.A.--Whitaker, H.A.: Morphological development of the auditory system. In: The Neuropsychology of Language. Szerk. Rieber, R.W. New York and London 1976, 125--157.
- Shannon, C.W.: A mathematical theory of communication. Bell Syst. Techn. J. 27. 1948, 379--443. Idézi: Miller--Isard 1963.
- Shapley, M.: Prosodic variation and audience response. Papers in Pragmatics 1. No. 2. 1987, 66--80.
- Simons, W. F.: Transformation of the vowel triangles of different speakers. IPO Annual Progress Report 15. 1980, 29--33.
- Speaks, Ch. E.: Measurement of central auditory integrity. In: ASMA Reports 14. Szerk. Elkins, E. Rockville Pike 1984, 70--77.

- Spreng, M.--Young, N.B.: Számítógéppel előállított magánhangzók alkalmazása az objektív audiometria céljára. Zeitschrift für Hörgeräte Akustik. 1978. IX. (fordítás).
- Stern, W.L.: Psychische Präsenzzeit. Zeitschrift für Psychol. 13. 1887, 325--348.
- Stevens, K.N.: Toward a model for speech recognition. JASA 32. 1960, 47--55.
- Stevens, K.N.: The potential role of property detectors in the perception of consonants. In: Auditory Analysis and Perception of Speech. Szerk. Fant, G.--Tatham, M.A.A. New York 1975, 303--330.
- Stevens, K.N.--Halle, M.: Remarks on analysis by synthesis and distinctive features. In: Wathen-Dunn, W. (szerk.): Models for the Perception of Speech and Visual Form. Cambridge 1967.
- Stevens, K.N.--House, A.S.: Speech perception. In: Tobias, J. (szerk.) Foundations of Modern Auditory Theory. Vol. II. New York 1972, 1--62.
- Stevens, K.N.--Liberman, A.M.--Studdert-Kennedy, M.--Öhman, S.E.G.: Cross-language study of vowel perception. L & S XII, 1969, 1--23.
- Stevens, K.N.--Blumstein, S.: Invariant cues for place of articulation in stop consonants. JASA 64. 1978, 1358--1368.
- Stickt, T.G.--B.B. Gray: The intelligibility of time-compressed words as a function of age and hearing loss. J. of Speech and Hearing Res. 12, 1969, 443--448.
- Stevens, P.: Spectra of fricative noise in human speech. L & S III, 1960, 32--49.
- Stricker, S.: Studien über die Sprachvorstellungen. Wien 1980.
- Studdert-Kennedy, M.: Speech perception. L & S 23. 1980, 45--66.
- Sumbly, W.M.--Pollack, I.: Visual contribution to speech intelligibility in noise. JASA 26, 1954, 212--215.
- Summary of the discussion on speech perception. L & S 23. 1980, 81--91.
- Szabó Imre--Wallis, J.: A beszéd vagy hangképzés. (Sonorum Formatio) MGyógyged. II, 1910, 23, 54, 83, 113, 149.
- Szabó Elemér--Szabó Gyula--Simon György: Az alapellátás közreműködése a 0--18 éves csökkenthallásúak föl kutatásában. Medicus Universalis XVII, 1984, 135--139.
- Szende Tamás: A beszéd folyamat alaptényezői. Budapest 1976.
- Szépe György: Az alsóbb nyelvi szintek leírása. ÁltNyT VI, 1969, 359--466.
- Szirtes József--Diekmann, V.--Rothenberger, A.--Jürgensen, R.: Adalékok a beszédészlelés mechanizmusához. Pszichológia 1981/2, 171--182.

- Szirtes József: A beszédészlelés vizsgálata a pszicholingvisztikában és a pszichológiában. MPSzichSz 1983/3, 385--425.
- Tallal, P.: Implications of speech perceptual research for clinical populations. In: Speech and Language in the Laboratory, School and Clinic. Ed. J.F. Kavanagh--W. Strange. Cambridge, Massachusetts, and London 1978, 73--89.
- Tarnóczy, T.: Resonance data concerning nasals, laterals, and thrills. Word IV, 1948, 71--77.
- Tarnóczy, T.: Die akustische Struktur der stimmlosen Engelaute. AlinguH IV, 1954a, 313--349.
- Tarnóczy Tamás: Az érthetőségvizsgálatok magyar szövegmintái. Magyar Híradástechnika 3--4. 1954b, 43--51.
- Tarnóczy Tamás: A nazalitás fizikai jellemzői. In: A Magyar Fonetikai Foniátriai és Logopédiai Társaság Kongresszusa. Budapest 1968, 6.
- Tarnóczy Tamás: Hangnyomás, hangosság, zajosság, Budapest 1984.
- Tarnóczy, T.--Vicsi, K.: Some remarks on the perception of unvoiced stop-consonants. Acustica 43. 1979, 167--173.
- Trenschel, W.: Das Phänomen der Nasalität. Berlin 1977.
- van Uden, A.: Die Welt der Sprache für gehörlose Kinder. Neckar 1976.
- Vaane, E.: Subjective estimation of speech rate. Phonetica 39, 1982, 136--149.
- Varga Ervin: Új módszer a beszédpercepció kísérletes vizsgálatára pszichiátriában. MTA Biol. és Orvosi Tudományok Oszt. Közleményei 11. 1960, 73--93.
- Vértes O. András: A gyermek nyelve. Budapest 1954.
- Vértes O. András: A magyar leíró hangtan története az újgrammatikusokig. Budapest 1980.
- Vértes O. András: A magyar beszédhangok akusztikai elemzésének kérdései. In: Fejezetek a magyar leíró hangtanból. Szerk. Bolla Kálmán. Budapest 1982, 71--115.
- Vicsi Klára: Az időtartam szerepe néhány mássalhangzó-típus hallás alapján történő megkülönböztetésében. Magyar Fonetikai Füzetek 7. 1981, 59--67.
- Villiers, J.G. de--Villiers, P.A. de: Language Acquisition. Cambridge, Massachusetts and London 1979.
- Wallis, J. ld. Szabó Imre 1910.
- Willeford, J.A.: Sentence tests of central auditory dysfunction. In: Handbook of Clinical Audiology. Szerk. Katz, J. Buffalo, New York 1978, 252--262.

- Williams, L.--Bush, M.: The discrimination by young infants of voiced stop consonants with and without release bursts. *JASA* 63. 1978, 1223--1225.
- Wingfield, A.: The intonation-syntax interaction: prosodic features in perceptual processing of sentences. In: *Structure and Process in Speech Perception*. Szerk. Cohen, A.--Nooteboom, S.G. Berlin 1975, 146--160.
- Wingfield, A.--Buttet, J.--Sandoval, A.W.: Intonation and intelligibility of time-compressed speech. Supplementary report: English vs. French. *J. of Speech and Hearing Research* 22. 1979, 708--716.
- Wodarz-Magdics, K.: Experiments in Hungarian vowel recognition. *Linguistics* 56. 1970, 64--87.
- Woodrow, H.: Time perception. In: *Handbook of Experimental Psychology*. Ed. S.S. Stevens. New York, London 1951, 1224--1236.
- Woodworth, R.S.--Schlosberg, H.: *Experimental Psychology*. New York 1954.

TÁRGYMUTATÓ

- abszolút döntés 71
ABX, AX kísérletek 23
agyi kiváltott válasz 63
agykapacitás 28, 46,
affrikáták azonosítása 139--142
aktív--passzív modell ld. analízis szintézissel modell
akusztikai szint 57, 59, 195, 221--222
akusztikai kulcs 108, 122, 196
akusztikus invariancia 39--40, 46
alsóbb szintek 57--58, 195--196
analízis szintézissel modell 24--25
anyanyelvi percepciós bázis 15, 16--17
anyanyelvi percepciós bázis kialakulása 190, 193
artikulációs invariancia 39--40
artikulációs konfiguráció stabilitása 27, 38
asszociációs mezők 25, 44, 189, 199
audiogram 54, 204--205, 214--215
audiométer 54, 215, 217
állandó ld. invariancia
- belülről generált képzetek 24, 25
beszédaudiometria 53, 55, 181, 206, 212
beszédaudiogram 55
beszéddallam észlelése 160--163, 164--180, 186--187
beszédérthetőség 56
beszédeszlelés szűrése 223
beszédeszlelés óvodáskorban 194--201, 210
beszédeszlelés kezdetei 183--188
beszédeszlelés megítélése G-O-H készülékkel 217--223
beszédeszlelés zavarának következményei 225

beszédfrekvenciák 75--76

beszédhallás 182, 194

beszédhelyzet szerepe 44, 183, 188, 197, 200

beszédtempó és tempóészlelés összefüggése 157

beszélőgép 11

bináris döntés 36, 65, 72--73

centrális zavarok 56

centrális hallás 49

click-kísérletek 32--33

cochlea ld. csiga

cochleáris implantáció 57

Corti-szerv 50, 54, 57

csecsemők beszédészlelése 182, 184--188

csiga 13, 50, 181

csontvezetéses hallás 52

csontvezetéses hallás vizsgálata 54

dichotikus vizsgálat 29, 230

diszkriminációs képesség 186--187

echoikus memória 36

elfedés 20, 98

elsajátított különbözőség 183

elsajátított azonosság 183

elsődleges (felismerési) kulcs 101--102, 109, 115, 142

elsődleges hallási elemzés 59

emlékezet 35, 36, 65, 139, 224, 228--229

engram 35

etalon 33

expresszív hatás 170, 179

ép hallás 54, 205

ép hallás és beszédmegértés összefüggése 49, 194, 212

értés 158

észlelési és megértés viszonya 158

észlelési tesztlap 232

- felismerési kulcs 74--77, 137, 155, 196, 209, 227
 féltéke-dominancia ld. jobbfül-dominancia
 fonemikus hallás 16, 192
 fonemikus különbség észlelése 106
 fonetikai osztályozás 27, 60--62
 fonetikai szint 60--63, 157--158, 221--222, 231
 fonetikai tér 40--41
 fonémaátcsapás 23, 106, 141
 fonémadöntés 32, 64, 65--67, 71--72, 109
 fonológiai szint 63, 221--222, 231
- globális percepció (gyermekkorban) 188
 G-0-H hallást és beszédészlelést vizsgáló rendszer 216
 G-0-H tesztlap 218, 219
- halláselméletek 52
 hallásérzékenység 155, 181--182
 hallási memória 63
 halláskárosító tényezők 53
 halláskárosodás 9, 52, 54, 57, 76, 203, 205, 209, 211, 213
 hallásmechanizmus 51
 hallássérült 65, 76, 203, 209, 212
 hallásszűrés 56, 215
 hallásterjedelem 52
 hangátmenet 11, 61--62, 109, 117--119, 186
 hangmagasság észlelése 167--168
 hangmenetek észlelése 168--169
 hangsúlyélmény 175
 hangsúlyészlelés 174--180, 187
 hangutánzó szavak 16
 háttérismeret/-tudás 42, 45
 hierarchikus modell (globális beszédmegértés) 26, 35, 47--49,
 57--68, 224
 hosszú idejű memória 65

időtartam mint felismerési kulcs 100--107, 138--142
 időzítési kulcsok 150
 immanens artikuláció 22
 információsűrűség 46
 intenzitás szerepe 142--149
 invariancia-jelenségek 42--47
 invariancia 25, 26, 36--42
 invariáns jegy 47, 74, 76, 203, 205
 irrelevancia 45

jelentés hatása 44, 139, 141
 JND-éppen észlelhető különbség 155
 jobbfül-dominancia 14, 28--30, 194, 230

kategoriális percepció 23--24, 106
 késleltetett kötés 27, 36, 47
 kétféle reprezentáció 189--190
 Klatt modellje 150
 kommunikációs helyzet 7
 kommunikációs lánc 15
 kontextus szerepe 27, 43, 64, 67, 68, 70--71, 74, 108, 119--120,
 123, 127, 129, 136, 142, 143
 "kódolt" gondolatsor 9
 kulcsszó-stratégia 189
 küszöbfeletti vizsgálatok 53, 54
 küszöbgörbe 54, 205, 210, 212, 213, 214--215

lateralizáció ld. jobbfül-dominancia
 légvezetéses hallás vizsgálta 54
 lég-csont-köz 54
 légvezetéses hallás 52
 libermani tétel 22

[m-1] átcsapás 127

Markov-lánc 60

magánhangzó-azonosítás 68--71, 77--107

magasabb frekvenciák szerepe 77, 78, 94--98, 125

második formáns szerepe 77--107, 108--121, 123--128
másodlagos (felismerési) kulcs 78, 115, 133, 142
mássalhangzó-felismerés 17, 22--23, 63, 107--149
mondatazonosítás 194, 199--201, 226
mondatérthetőség 159
motoros komponens 12, 20, 24, 25, 61
motoros teória 20

nazálisok azonosítása 121--128
neurális spektrográfia 100

objektív audiometria 52--53

percepció bázis 16--19, 181--182, 193, 201

percepció egység 32--36

perifériás hallás 49

prototípus 67

prozódiai észlelés 164

redundancia 45, 46, 64--66, 100, 203--206, 210

részhangok azonosítása 128--142

rövid idejű memória 65

súgott beszéddel történő hallásvizsgálat 53

süketnéma 11, 31, 57

szegmentálás 27--28, 32, 36, 60--61, 192

szimultán működések 26, 35, 60

szintetizált magánhangzók azonosítása 100--107

szintetizált mássalhangzók 109--121, 123--128, 132--141

szintetizált szavak 203--205, 207--209

szintetizált szavak és hallásszűrés 209--215

szófelismerés 194, 195--199, 205, 209, 226--227

szótagszám és felismerés 199

szótalálás (kísérletben) 194

szubjektív audiometria 52--53

szűréssel torzított magánhangzók azonosítása 78--100

szűrőssel torzított mássalhangzók azonosítása 122--123, 128--132

tárolás 35, 63, 64, 65

tárolt hallási kép 35

tempóészlelés 149--164

tisztahang audiometria 53, 54, 212, 215

tisztahang-hallás 54, 56, 186--188

többletinformáció ld. redundancia

többtényezős azonosítás (több felismerési kulcs) 74, 100--107, 117,
117, 119

történeti áttekintés 10--14

transzformációk 40, 47

újszülöttkori észlelés 183--184

visszacsatolás 28, 47, 60

vizuális észlelés 30--31, 230

VOT az észlelésben 108, 185

zárhangok azonosítása 108--121, 139--142

zörejfrekvencia szerepe a réshangok azonosításában 132--138

SPEECH PERCEPTION

Introduction

The study of the process of speech perception is the subject of various disciplines. The linguistic-phonetic study of speech inevitably involves the investigation of perception and comprehension as parts of the entire communicative chain. Works on this topic, highly increasing in number in the international literature since the 1950', exhibit a three-fold, articulatory--acoustic--perceptual approach. Despite their valuable results, no unambiguous, language-specific description of the speech perception process has become available as yet.

Research on speech perception appears in the Hungarian literature rather sporadically. Nevertheless, the articles published report important partial results and provide models, problems, and a basis for further work.

The topic of this book is the investigation of (i) the speech perception process (ii) perceptual properties of Hungarian speech sounds, and (iii) the description of the applications of the results so obtained. The results of the study are based on experiments directed at the investigation of elementary perceptual units, the acoustic cues of Hungarian vowels and consonants, the development of speech perception, and hearing and perception examinations in children. The linguistic material was prepared by the following methods: frequency filtration, masking with white noise, and synthesized speech signals. Subjects of the experiments were native speakers of Hungarian, (i) adults from Budapest, differing in age and education, and (ii) children's groups of heterogeneous social composition. More than 1000 adults and children took part in the experiments. Methods of test procedures also varied with the specific goals: (i) in individual and group testing in a silent room, the prepared material was administered to the subjects either through a loudspeaker or through headphones, (ii) in clinical and kindergarten environment, the sound signals were given through headphones.

Results

The results are, on the one hand, linguistic-phonetic in nature:

speech processes were interpreted in the framework of articulation--acoustics--perception. The experimental data and the conclusions based on them contribute to a more faithful understanding of the quality of the speech sounds examined. On the other hand, our investigations made it possible to work out useful procedures which have been applied in practice. The specific results presented in the book can be summarized as follows.

The phenomenon of invariance

The author thinks that the phenomenon of invariance cannot be defined at a single level of speech. Corresponding to the hypothesized levels of speech understanding, different invariance phenomena are assumed to exist and are described by the author: (i) invariance at the level of (background) knowledge (ii) invariance in the actual speech situation, (iii) contextual invariance, (iv) invariance in speech sounds. The last-mentioned type is discussed in detail in order to demonstrate that the experimental results support the author's multi-level-invariance theory and to prove the primacy of acoustic structure in this respect. The invariance phenomenon can be defined by 'identification/perception cues'; in particular, by the acoustic-phonetic cues of a given speech sample, i.e. those acoustic parameters which are absolutely necessary for its identification. The experimental data show that there are 'multiple cues' and 'secondary cues' as well. The main consequences of the investigations of acoustic-phonetic cues can be summarized as follows: (i) the hearing system is able to categorize speech-like stimuli independently of language-specific characteristics; (ii) the speaker/hearer's native perception basis affects the working of his 'hearing system'; (iii) the acoustic-phonetic cues of speech sounds are defined by their main properties: some of them can be defined on the basis of their steady-state phase (like vowels) while others only in sound combinations (like stops); (iv) the participation of the three acoustic parameters in perception varies depending on the articulation of speech sounds; (v) there are a number of languages - including Hungarian - for which the acoustic-phonetic cues of speech sounds have been defined by experiments.

Perception models

A survey of the most characteristic speech perception models is followed by an attempt to form the author's own model. Experimental data concerning the "lower" hierarchical levels led to the conclusion that the elementary perceptual unit of the phonological level is the CV or VC sequence, and/or the sequence of one vowel and several consonants. The investigation of the identification of different types of speech sounds suggests that classification according to phoneme classes can be either absolutely precise, or less precise, even conditional, given that identifiability is determined by articulation and the parametric relations in the acoustic signal sequence thus obtained.

Identification of Hungarian vowels and consonants

Perceptual results for the Hungarian speech sounds show that the, or one of the, primary cues to identification can be expressed in terms of frequency values. The first two formants (F1 and F2) determine the Hungarian vowels correctly. The frequency range of the first formant changes with the increase of duration: as a rule, it narrows down/decreases, except for palatal labials, which are stable in this respect. Frequency components are characteristic of the quality of vowels at higher frequencies too; their role in perception is nevertheless much debated. According to the experimental data these components ensure the identification of vowels, if the intensity of the first two formants confirming the exact identification is decreased to a great extent. In such cases the phoneme decision is based on higher-number formants available.

Perceptual data for voiced stops ([b, d, g]) partly supported the results for the corresponding English consonants. Only the second formant of the burst of the bilabial plosive can be considered to be the primary cue to identification. The identification of dental plosives is influenced by the following vowel in the following way: the value of F2 is decisive before palatals, but before velars, intensity also has a great importance (multiple cues). It is characteristic of [g] in every sequence with vowels that F2 does not exclusively determine whether it can be precisely identified or not. Before front vowels, intensity value of F1 and internal durations are also important factors; before back vowels, intensity, F1

and internal duration are decisive, while the value of the second formant is secondary. The F2 value-range of the plosion of stops was expressed by implications which guarantee correct identification (in every CV combination).

The data concerning the identification of nasal consonants ($[m, n, \eta]$) did not support the claim in the literature that, similarly to stops, the invariant features of nasals of different places of articulation can be acoustically expressed by the value of F2. Hungarian nasal consonants cannot be unambiguously characterized by a unique primary cue to identification (e.g. as the frequency of the first formant of $[m]$ was increased with 150 Hz, its identification changed into $[l]$). According to the author's data, initial $[m]$ can be identified the most positively; $[m]$ and $[\eta]$ alternate with each other and with the dental nasal consonant in recognition, but $[n]$ is often mistaken also for $[l]$ and $[r]$ by the subjects.

The frequency-range of noise of the dental and alveolar fricatives in CV sequences with every vowel has been determined. The results have been expressed with implications. For example $[s]$ is identified before the vowel $[a:]$ if the noise is ≥ 4000 Hz, or in case of two foci, if one of them is 3500 Hz, and the other is > 4000 Hz, respectively; $[\zeta]$ is identified if the noise is between 1000 and 3500 Hz, or in case of two foci, if one of them is ≤ 2500 Hz, and the other is > 4500 Hz, respectively. The frequency range between 3000 and 4000 Hz seems to be uncertain for correct discrimination of dental and alveolar fricatives. In case of two noise foci in this range, their distance from each other also influences identification. On the basis of synthesized material exact values for the correspondance of frequency and duration concerning the identification of fricatives, affricates, stops in sound sequences and words could be provided.

Perception of speech tempo, melody, and stress

The perception of suprasegmentals has also been analyzed. Experiments were carried out concerning Hungarian speech tempo, fundamental frequency, and stress. A good many studies have demonstrated that the perception of speech tempo is a function of three components: articulation rate and number and duration of pauses. However, very careful experiments have sug-

gested the contribution of other factors like fundamental frequency, intensity, meaning, and frequency of words. The present experimental results can be summarized as follows: (i) subjective decisions on speech tempo have supported that the Hungarian speech tempo has been speeded up; (ii) utterances having peaks or an increasing fundamental frequency curve were judged faster than utterances with a decreasing F0-curve; (iii) judgements of the tempo of monotonous utterances depend on the actual articulation rate: until a certain value they seem to be judged slower than all the others while after this value they are perceived as faster (synthesized material was used).

The aim of the perceptual experiments concerning speech melody was to find out how the informants perceive length values in the melodic patterns, peaks and peak changes, register, range and pitch intervals, and also to see what relationship exists between the acoustic features and linguistic content. Psychoacoustic analyses showed that subjects were able to perceive the length of the given melody, the places of the peaks, and also the changes in the frequency contours. On the basis of their judgements, correspondences could be drawn between the melody patterns and Hungarian melody categories. Stress judgements were based on peaks and/or higher intensity in the imitated melody curves. An increase in fundamental frequency with or without an accompanying increase in intensity was perceived as stress. Main stresses were indicated on the basis of intensity alone in merely 20 % of the cases.

Development of speech perception

The author has made an attempt to analyse the process of speech perception in infants and children. After a survey about the investigations carried out on this topic so far, the author's own experiments are presented. She deals with the mechanism and evolution of hearing, the role hearing plays in understanding, and with the interrelation of hearing and speech perception. 'Perception base' as a fundamental degree of development occurs at the age of 3 years. However, the actual performance of children in speech perception is very uncertain, the understanding process is still reinforced by extra- and paralinguistic factors of verbal communication. With the participation of more than 300 children the development of speech perception could be defined and expressed in terms of identification scores between the ages of 3 and 10 (cf. Figure 35). The results obtained

have confirmed the author's hypothesis that the development of understanding involves a specific change in the process of speech perception/understanding. The decisions/operations on the lower levels of the speech understanding mechanism -- the operations on the acoustic and phonetic levels -- are more crucial for little children than the higher levels. The development in this respect corresponds to the change of the operations on the hypothesized levels of the understanding mechanism, as it could be seen in the strategies of word identification across age-groups (199).

Practical application of perception results

The experimental results concerning both the identification of Hungarian speech sounds and the development of the speech perception process in children provided a possibility for their practical application. One of them is the G-O-H hearing and speech perception level screening procedure while the other is a test-package for a thorough assessment of the speech perception level of normal-hearing children. Both methods have been developed at the Phonetics Laboratory of the Hungarian Academy of Sciences.

Using artificial speech, a new method was developed with which the functioning of the hearing mechanism and the level of speech perception can be estimated at the same time. It has its significance where the traditional methods are imperfect or their results are uncertain, and in mass-screening procedures. The hypothesis was first practically tested under laboratory conditions by means of frequency filtration of both natural and synthesized monosyllabic Hungarian words. Different degrees and types of hearing impairment were simulated which verified the working of the filtering function of artificial words (cf. Figure 36). The same test material was then tested in clinical and kindergarten environments in children with normal and impaired hearing. Both series of experiments (more than 300 children were involved) suggested that this new method -- the G-O-H -- provides a real solution for the mass-screening of hearing, and for the assessment of the speech perception level of kindergarten children (cf. in English: Gósy--Olaszy--Hirschberg--Farkas 1987).

It seems to be very difficult to tell whether a child's speech perception level meets the age-specific requirements or not. In normal communication situations (home or school life) there is no opportunity to detect the child's perception deficiencies; these problems remain frequently hidden

because of various strategies used by children in the speech understanding process.

In solving this problem the criteria of an age-specific test-package for assessing the speech perception level of normally developed children (ages between 4 and 10) have been developed. On the basis of the experimental results, a special diagram of speech perception development has been set up for Hungarian. The test-package consists of 10 subtests (cf. Figures 42 and 44) which examine the levels of the perception mechanism almost separately from hearing until the phonological level. The test-package can also be used in decisions for or against a hearing aid, in the examination of possible dyslexic children, or to add exact data for obtaining better knowledge on central hearing problems.

