

Beszédképzés és nyelvi reprezentáció

1. B e v e z e t é s. – KEMPELEN FARKAS a XVIII. század végén így határozta meg a beszédet: „A legtágabb értelmezés szerint a beszéd az a képesség, hogy érzéseinket és gondolatainkat másokkal jelek útján közöljük” (WOLFGANG VON KEMPELEN, *Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine*. Wien, 1791. Magyarra fordította: MOLLAY KÁROLY. Bp., 1982. 27). Ez a definíció évszázadokon át élt különféle megfogalmazásokban; az utóbbi évtizedek beszédkutatásának szemléletváltása azonban a beszéd fogalmát a beszédfeldolgozásra is kiterjesztette. Ezen a módon nemcsak

a beszélő, hanem a mindenkori hallgató tevékenységét, az aktív beszédpercepciók folyamatait is tekintetbe vette. A kommunikáció tágabb értelmezése is a jeladás és a jelfeldolgozás összefüggéseit jelenti, akár az állatok nyelvére, a gyermeknyelvre, avagy a patológiás beszédfolyamatokra, például az afáziára gondolunk. A beszéd képzése a beszédtervezési folyamat végeredménye, amelynek során a szándékolt közlést artikulációs mozgásokká transzformáljuk. A XIX. századi újgrammatikus iskola nagy alakja, HERMANN PAUL szerint a beszélők fizikai folyamatokon keresztül adják át gondolataikat. PAUL a mentális folyamatokat megkülönböztette a fizikai értelemben vett átvivő folyamatoktól. Úgy vélte, hogy számos mentális folyamat nem akaratlagos. Azt is feltételezte, hogy ezeknek a nem akaratlagos folyamatoknak egy jelentős része nyelvvel kapcsolatos jelenségeket tartalmaz. PAUL ezzel a hipotézisével a ma is elfogadott, modern elméletek előfutára lett.

A beszédprodukciónak tervezési része közvetlenül nem tanulmányozható, modellekben megfogalmazott hipotézisek mutatják be a mechanizmust (GÓSY MÁRIA, *Pszicholingvisztika*. Bp., 2005. 80–4). A legáltalánosabban elfogadott LEVELT modellje (WILLEM J. M. LEVELT, *Speaking. From intention to articulation*. Cambridge, Massachusetts, 1989.). Ez a modell vázlatosan a következő. A beszélést egyrészt az átadásra szánt gondolat megtervezése, másrészt az annak megfelelő aktuális nyelvi forma hozzárendelése előzi meg. A kettő rendszerint olyannyira egyszerre zajlik, hogy a beszélő nincs is tudatában a kétféle folyamatnak. Az első fázis a makrotervezés, ez egyfelől a szándékot, másfelől a grammatikai megformálás előkészítését tartalmazza. Ezt követi, rendszerint átfedésben, a mikrotervezés, amelynek során a beszélő az egyes beszédaktusokat nyelvi formába teszi, megkezdi a lexikális egységek válogatását a mentális lexikonból, és bizonyos időrendet tervez meg az elhangzásra vonatkozóan. A makrotervezés tehát inkább maga a szándék; a mikrotervezés pedig már pontosan definiálja a kiválasztott, használandó nyelvi szerkezeteket. A makro- és a mikrotervezés között kisebb-nagyobb átfedések és párhuzamos működések is lehetnek. A szemantikai szerkezet az a tartalom, amit a beszélő közölni szándékozik. Ez utóbbi jelenik meg a szintaktikai és a fonológiai struktúrák közvetítésével. Bizonyos alapstratégiák működése biztosítja a beszédprodukciónak sikerét; ezeken belül a beszélő olyan transzformációs szabályokat követ, amelyek szoros kapcsolatot tartanak a tárolt (nyelvi és beszéd-) elemekkel. A fonológiai szabályok érvényesülését követően létrejön a fonetikai struktúra, amely az artikulációs gesztusok megformálását készíti elő; ezen a szinten zajlik az artikulációs tervezés. Majd megtörténik a beszédhangsorozatok kiejtése, amelynek eredménye az az akusztikai jelsorozat, amelyet a hallgató dekódol.

Az artikulációs gesztusok következményei a levegő közegében mint akusztikai jelek továbbítódnak, ezek feldolgozása történik a hallgató dekódolásának különböző szintjein. Számos transzformáció történik eközben; az eredeti gondolat, a tartalom, a forma az artikulációs mozgásokban – zöngképzésben vagy éppen a hangszalagműködés hiányában, a különféle beszédképző szervek mozgásában, működési eredményeiben – megőrződve, majd az idő, a frekvencia és az intenzitás paramétereivel jellemezhetően akusztikai hullámformává alakulva jut el a hallási feldolgozás területére. Ez a beszédfeldolgozás kapuja, amelytől a percepciók folyamatok kezdődnek. A hallás épsége feltétele a rá épülő folyamatok elvárt működésének; a jó hallás önmagában azonban nem biztosítja a percepciók folyamatokat. Az ép hallás és a rá épülő, jól működő szintek nem elegendőek ahhoz, hogy a hallgató az agy megfelelő területein ugyanazt a tartalmat, formát dekódolja, mint ami a beszélő szándéka volt. Ez történik akkor, ha a nem (vagy nem jól) ismert idegen nyelv hallgatására gondolunk. Szükség van arra ugyanis, hogy közös legyen a beszélő és a hallgató mentális reprezentációja. A mentális reprezentáció többféleképpen értelmezhető; gondolatok, eszmék, vágyak, elképzelések, illetőleg észleletek, amelyek rendelkezésre állnak, előhívhatók. Ezen belül a specifikusan 'nyelvi' olyan mentális reprezentáció, amely az

A magánhangzók tulajdonságai a hangminőségtől, a hangkörnyezettől, a beszélőtől, a tempótól függően változnak, ugyanakkor a normalizációs folyamatok biztosítják a felismerést a fonetikai különbözőségek ellenére. Ezek a normalizációs folyamatok a beszédészlelés során érvényesülnek elsődlegesen, de módosulhatnak a feldolgozási modell felsőbb szintjein. Tapasztalatból tudjuk azonban, hogy a működések nem hibátlanok; a hibalehetőségek sokfélék, és a mindennapi kommunikációban elő is fordulnak. Megállapítható az is, hogy a beszéd motoros kontrollja nem eredményez feltétlenül invariáns akusztikai paramétereket, amelyek a hallgatót egyértelműen segítenék a feldolgozásban.

B) Talán nincs még egy olyan beszédhangunk, amellyel kapcsolatban annyi ellentmondás, vita és tagadás lett volna és lenne ma is tapasztalható, mint a [h]. Vitatják a képzési helyét, bizonytalan, hogy van-e és milyen (ha van) a zöngés megfelelője, sőt még az is kérdéses, hogy voltaképpen mássalhangzó-e vagy magánhangzó? Mindezek után az már nem is meglepő, hogy a kiejtése bizonyos kontextusban csaknem meghatározhatatlan és előjelezhetetlen. A „kezdetekben” a [h] mássalhangzónak még a léte is kérdéses volt. A XVIII. század elején például pusztán lehelésnek tartották, nem beszédhangnak. VIRÁG BENEDEK, a költő hívta fel a figyelmet olyan szóalakzatokra, amelyek csak ennek a hangnak a meglétében vagy hiányában különböznek, mint *adatom* és *adhatom*. A mai szakszóval élve, fonológiai alapon igazolta a [h] önálló létét. Később zöngétlen magánhangzónak (XIX. század) is tekintették, azt gondolván, hogy a zöngétlenség ténye e hang lényegéhez tartozik (LAZICZIUS GYULA. Fonétika. Bp., 1944. 76). Jóllehet tapasztalták az „ún. zöngés h-kat” (pl. gégetükörrel), sőt voltak közlések ilyen mássalhangzóról különböző nyelvek esetében (pl. örmény), és azt is elismerték, hogy ilyeneket lehet képezni, mindennek ellenére „afféle fonetikai különlegességnek” tekintették őket. HORGER ANTAL például elismerte, hogy a [h]-féle hangok zöngésedhetnek, azonban a magyar laringális résmássalhangzókat zöngétlennek tartotta. A magyar [h] zöngés változatát elsőként MEYER igazolta, majd GOMBOCZ ZOLTÁNNAL végzett kutatásában is megerősítette (E. A. MEYER – GOMBOCZ ZOLTÁN. Zur Phonetik der ungarischen Sprache. Uppsala, 1909.). Ők írják le azt – mintegy 20 felvétel alapján –, hogy a „h-hang” két magánhangzó között zöngéssé válik (ún. leheletes zöngéről írnak, amit a „hauchstimmhaft” szóval definiálnak). LAZICZIUS elfogadja a leheletes vagy mormolt zöngét, a zöngésedést azonban pontosítja. Nemcsak két magánhangzó, hanem általában két zöngés hang közötti előfordulásra is érvényesnek véli. Azóta ez a megállapítás gyakorlatilag megkérdőjelezetlenül jelen van a szakirodalomban.

A /h/ realizációját glottális réshangra és szótagkezdő pozícióra korlátozó fonológiai megkötés a XVII. századtól már nem érvényesül kötelezően; a megfelelő fonetikai helyzetben nem hajtódik végre valamennyi lexémában automatikusan. Ezt mutatja egyrészt az *ihlet* szó, amely minden valószínűség szerint kihalt, majd a XVII. században a régi források alapján ismét használatossá vált, de nem *ilhet* vagy *illet*, hanem az „eredeti” *ihlet* hangalakban (BÁRCZI GÉZA – BENKŐ LORÁND – BERRÁR JOLÁN, A magyar nyelv története. Bp., 1967. 128). Ezt látszanak igazolni másfelől a XVII. század után nyelvünkbe került jövevényszavak, pl. *pech*, *technika* hangalakja, valamint az, hogy megjelennek a [h]-s változatok a *cseh*-féle csoporton belül is.

A *h*-végű névszótövek hangtani sajátosságaira vonatkozó megállapítások a XIX. század közepén jelennek meg a fonetikai leírásokban, ekkor olvashatunk először arról, hogy például a *tereh*, *juh* és *méh* szavakban ún. „nyugvó *h*”-t ejtünk (VÉRTES O. ANDRÁS, A magyar leíró hangtan története az újgrammatikusokig. Bp., 1980. 99). A „néma *h* szabály” legkorábbi megfogalmazásaival a nyelvvelőli gyakorlatban találkozunk. 1856 óta az akadémiai helyesírási ajánlásokban, majd később „A magyar helyesírás szabályai”-ban rögzítik, hogy a *h*-ra végződő szótövekben az ortográfiai és ortoépiai szabály eltér (SZEMERE GYULA, Az akadémiai helyesírás története [1832–1954]. Bp., 1974. 124–5). A nyelvhasználatra

vonatkozó megfigyelések és vélekedések azonban ellentmondóak. Az ellentmondásokra vagy éppen az ingadozás következtetlenségére (?) jó példa a *méh* 'rovar' szó megítélése. A G. VARGA (Alakváltozatok a budapesti köznyelvben. Bp., 1968. 235) kérdőívében szereplő négy szóalak közül a „néma *h* szabály” szerinti ejtés jóval ritkább volt a *méh*, *méhnek* 'rovar' szóalakok, mint a *cseh*, *csehtől* esetében. Ezzel szemben ELEKFI szerint a *méh*^{1,2} homonimák *méh* (*mé[h]*, *mé[h:]jel*) 'női szerv', ill. *méh* (*m[e:]*, *m[e:]vel*) 'rovar' szerinti ejtése nem kötelező, de ezek szokásos kiejtési megkülönböztetését regisztrálja (ELEKFI LÁSZLÓ, Homonimák felismerhetősége toldalékos alakok alapján: Nyr. 2000: 146–63). A kísérleti adatok azt igazolták továbbá, hogy a magyar anyanyelvű adatközlők nyelvhasználati tapasztalata, eltérő észlelési stratégiájuk, valamint a „norma” szigorúbb vagy elnézőbb alkalmazásai is hatással vannak az elemzett mássalhangzó viselkedésére. Olyan következtetések is megfogalmazhatók, amelyek hozzájárulhatnak a *h*-végű szavak nyelvi státusának mind jobb megismeréséhez. Kiderült például, hogy a gyakoriság a vártnál jobban meghatározza a vizsgált szavak produkciójának és észlelésének sajátosságait (GÓSY MÁRIA – KOVÁCS MAGDOLNA, Nyelvi tudatosság, norma és rendszerkényszer a *h* esetében: Nyr. 2003: 439–56). Másként jelennek meg az ejtésben és a percepcióban is a *h*-végű tőszók, a toldalékolt szavak és a szóösszetételek. A „néma *h* szabály” visszaszorulása legerőteljesebben a szóvégi pozíciót érinti, kevésbé a szóösszetételeket és a legkevésbé a toldalékolt alakokat, ami a morfológiai és nem a fonetikai kontextus erősebb hatását tükrözi. Az észlelés ugyan a konkrét akusztikai-fonetikai adatok alapján történik, azonban a beszélők nyelvi tudatossága, a mentális reprezentáció hatással van döntéseikre.

A szóvégi /h/ ejtése artikulációsán sokféle lehet, s ennek megvannak az akusztikai következményei. Ezek az akusztikai következmények érintik a fizikai időtartamot, a zörej intenzitását és frekvenciaspektrumát is. A kísérleti eredmények szerint (GÓSY–KOVÁCS i. m.) az értelmes szavakban és a logatomokban szereplő [h] hangok megítélése jelentős különbséget mutat az észlelésben. Statisztikailag igazolható, hogy az értelmetlen hangsorok azonosításakor a résztvevők kevesebbszer azonosítottak [h] hangot. Amíg a *méhkirálynő* szóban például 90,9%-ban észleltek [h]-t, addig az *éhki* logatomban csak 36%-ban. Ez azt a beszéd-percepcióban ismert jelenséget igazolja, hogy az értelmetlen hangsorok azonosításában a felsőbb feldolgozási szintek nem vesznek részt, döntően tehát az akusztikai, fonetikai, fonológiai szintek működnek. Szavak megítélésekor ugyanakkor a kísérleti személyek mégis működtetik a felsőbb szinteket, tehát lexémákat, morfológiai szerkezeteket stb. azonosítanak, s ez befolyásolja az észlelési működéseket. A tapasztalatunk mindezek ellenére az, hogy az ejtési bizonytalanság és a gyakorta előjelezhetetlen variációk, illetőleg az akusztikai sokféleség ellenére a hallgatók képesek a szavak helyes azonosítására, még izolált ejtésben is.

Kísérletsorozatban vizsgáltuk meg, hogy vajon valóban zöngés-e a [h] két magánhangzó között (mintegy 3000 adat alapján). Az 5. ábra szemlélteti a [h] ejtésének akusztikai következményeit ugyanabban a hangsorban (*ki hova*), különböző személyek ejtésében (GÓSY, A /h/ zöngésedése két magánhangzó között: Beszédkutatás 2005: 1–21).

Az akusztikai-fonetikai elemzések alapján a következő megállapítások tehetők. A magyar intervokális helyzetű /h/ fonéma realizációjakor a kiejtett beszédhangok zöngések, illetve zöngétlenek; sőt, akár vokalizálódhatnak is, azaz törlődhetnek mint mássalhangzók. A zöngések zöngésségben eltérhetnek, azaz különféle fokozatokban jellemző rájuk az alaphang megjelenése, valamint a zöngé intenzitása. A zörejesen (is) realizálódók esetében is különféle mértékben és frekvenciákon jelentkeznek a zörejösszetevők. A zöngésedés nem független az artikulációs tempótól; ugyanakkor jellegzetes eltéréseket tapasztaltunk az egyéni ejtéstől függően. Arra is választ kerestünk, hogy vajon a magánhangzó-kontextus, pontosabban a magánhangzók minősége befolyással van-e a zöngésség alakulására (6. ábra).

Az adatok igazolták a feltételezéseinket. Spontán beszédben két szó találkozásakor a szükséges zöngésségi hasonulás a két szó közötti szünet időtartamértékének függvényében realizálódik. Ez az érték fiataloknál szignifikánsan rövidebb, mint az idősebb beszélőknél, ami nyilvánvalóan az idősebbek lassú beszédtervezési folyamataival van kapcsolatban. A fonológiai működést tekintve egy további kérdés is felmerül a szabályalkalmazás és a mentális lexikon viszonyában: hol, illetőleg hogyan történik a zöngésségi hasonulás mint fonológiai szabály végrehajtása? Két lehetőség adódik.

Az egyik esetben a mentális lexikonban a lexikai egységek ún. alapalakban vannak tárolva, s az egyes beszédhangokat érintő változásokat a fonológiai szint aktiválása következtében a kódolt szabályok eredményezik. A másik lehetséges esetben a mentális lexikon tartalmazza a lexikai egységeket alapalakban és a fonológiai szabályok következtében módosult alakban egyaránt. A fonológiai szabály ebben az esetben azt határozza meg, hogy melyik alak lehívása szükséges a mentális lexikonból az adott kontextusban. Ez tehát azt jelenti, hogy az adott lexikai egységek mind a zöngés, mind a zöngétlen változatban tárolódnak. Ennek megfelelően a *hogy* kötőszó mind *hogy*, mind *hoty* alakban kódolva van, vagy például a *veréb* szó *verép* alakban is megtalálható a mentális lexikonban. További kísérletek szükségesek annak eldöntésére, hogy vajon mi módon történik a zöngésségi hasonulás kivitelezése a spontán beszédben.

3. Követkeltések. – A beszédképzés és a mentális nyelvi reprezentáció összefüggéseinek bemutatására több kísérletsorozat eredményeit tárgyaltuk. Elemeztük a magánhangzók ejtési különbözőségeit, a [h] intervokális zöngésedésének és a [r] koartikulációjának jellegzetességeit, valamint egy nyelvi szabály, a zöngésségi hasonulás megvalósulásának kritériumait a spontán beszédben. Minden esetben megállapítható volt, hogy a nemegyszer nagymértékű akusztikai-fonetikai változatosság ellenére a nyelvi reprezentáció állandó, és ez biztosítja a hallgató számára a percepciót a verbális kommunikáció során. Feltételezzük, hogy a nyelvi reprezentáció az agyban egyfajta neurális hangszínképként fogható fel, amely kellőképpen rugalmas ahhoz, hogy igazodjon a folyamatosan érkező akusztikai jelsorozat nyelvspecifikus sajátosságaihoz, ezáltal lehetőséget teremtve az invariáns elemek mindenkor feldolgozhatóságára. Mindennek bizonyítására további alapkutató-sok szükségesek, amelyekben az artikuláció – akusztikum – észlelés együttes szemléletében történnek az elemzések.

A beszédképzés és beszédfeldolgozás összefüggéseire irányuló kutatások eredményei felhasználhatók a nyelvtudományon belül a nyelvtörténetben, a szociolingvisztikában, a dialektológiában, avagy a fonológiában. Meghatározó következménnyel járnak a beszédtechnológiában a fonetikai alapú beszélőazonosítás és az automatikus beszéd felismerés területén. Az eredmények közvetve pedig felhasználhatók a pedagógiában is, például az anyanyelvi fejlesztésekben, a korszerű olvasás tanításban, a beszédhibák pontosabb megismerésében, a terápiák kidolgozásában, sőt akár az idegen nyelvek oktatásában és tanulásában.

GÓSY MÁRIA

Speech production and mental representation

The process of speech comprehension consists of two large phases: the perception of vocal phenomena corresponding to the system of linguistic signs, and the interpretation of that code system. Both phases are made up by several levels that collectively ensure the decoding of speech phenomena in a regular cooperation with one another. Speech comprehension is an active process whereby the

hearer interprets the speech phenomena s/he has perceived at successively higher levels. For that process to work, certain mental representations must be shared by speaker and hearer. The notion 'mental representation' can be understood in various ways; it includes thoughts, ideas, wishes, percepts, conceptions, etc. Specifically 'linguistic' mental representations, on the other hand, are such that they contain linguistically relevant signs and functions of the individual's patterns of knowledge. Certain portions of linguistic mental representations may keep changing or being modified throughout the individual's lifetime. The present paper analyses, in several series of experiments, cases in which objective acoustic phonetic parameters seem to contradict the corresponding mental representations. The aim is to highlight the relationship between articulatory/acoustic differences and the invariant features that underlie them as well as that between mental representations and objective parameters. Four areas are investigated with respect to Hungarian: (a) the production, acoustic patterns, and perception of vowels; (b) the variants of the phoneme /h/; (c) the coarticulatory behaviour of [r]; and (d) a phonological rule of coarticulation applying in spontaneous speech. On the basis of the results obtained, the author attempts to answer the question of how the (apparent) paradox of the contradiction between the objective physical reality of speech and its mental representation might be resolved.

MÁRIA GÓSY