

A BESZÉD TERMÉSZETESSÉGÉRŐL ALKALMAZOTT FONETIKAI SZEMPONTBÓL

Gósy Mária – Gyarmathy Dorottya – Horváth Viktória

Bevezetés

A beszédet az emberi kommunikáció legtermészetesebb formájaként tartjuk számon. Az embereknek általában fel sem tűnik, hogy milyen könnyen beszélnek – szemben például az olvasással vagy az írással. Ugyanakkor a szokatlan beszédhelyzet, illetve a váratlan beszédtema sokszor problémát okozhat; a beszélő nehezebben fejezi ki magát, elégedetlen azzal, amit mondott, avagy ahogyan azt megfogalmazta; a saját spontán beszédét nem a „megszokottnak” ítéli. A felnőttek egy részének a beszéde ilyenkor – az adott kommunikációs helyzetből függően – szorongóvá, feszélyezettebbé válik. Mások a beszédhelyzettől függetlenül, minden felkészülés nélkül, akár váratlan helyzetekben és témák esetén is képesek spontán módon jól beszélni, azaz jól követhető kiejtésben, megfelelő tagolással, egyértelműen elmondani azt, amit akarnak. A spontán beszédet sokan összekeverik a természetes beszéddel. Ez többféle okra vezethető vissza, például a definíciók sokfélesége, a megközelítések különbözőségeire, az elemzési célok eltéréseire. A jelen tanulmányban egyrészt annak a kérdésnek az elméleti körüljárását tűztük ki célul, hogy a beszéd spontaneitása és természetessége milyen összefüggéseket mutat, másrészt ezekkel a tényezőkkel kapcsolatos saját gyakorlati tapasztalatainkról számolunk be.

A spontán beszédéről

A beszédtudomány szakirodalma azt mutatja, hogy a különféle célú kutatók hosszú időn keresztül főként felolvasott vagy előre betanult szövegek vizsgálatán (szósorok, színészi, előadóművészi interpretációk, klisémondatok stb. alkalmazásán) alapultak. A spontán beszéd elemzésének igénye viszonylag új keletű a külföldi és a hazai kutatásokban egyaránt. Mikor spontán a beszéd? Úgy fogalmazhatunk, hogy a beszélő spontán módon beszél akkor, ha nem tudja (nem akarja, vagy nem áll módjában) közlendőjét részletesen átgondolni, a meghangosítást tartalmi és formai szempontból pontosan megtervezni (szemantikai relációk, grammatikai szerkezetek, szövegszerűség, lexikális válogatás, szegmentális és szupraszegmentális artikulációs kontroll stb.). A Levelt-féle beszédproduktions modell (1989) értelmében ez azt jelenti, hogy a beszédszándékot nem követi semmiféle felkészülés; a makro- és a mikrotervezés folyamatai előzetes tervezés nélkül működnek az adott beszéd-

helyzetben. A spontán beszéd fogalma egyértelműen definiálható ugyan, a szakirodalom mégis elsősorban azt emeli ki vele kapcsolatban, hogy megakadályozások tartják, valamint hogy a beszélők gyakran nem jól formált mondatokat használnak benne (pl. Ward 1989, Lustgarten et al. 2002).

A beszéd spontaneitása nem függ attól, hogy több szereplő közötti társalgásban történik-e a verbális kommunikáció, vagy valaki egy történetet (narratívát) mond-e el, de attól sem, hogy milyen a beszélő aktuális lelkiállapota (különböző spontán beszéd zajlik például az orvossal történő beszélgetés során, munkahelyi megbeszélésen, értekezlet alkalmával, családi, baráti élménybeszámolókor, vita vagy veszekedés esetén). A spontaneitásnak elméletileg akár fokozatai is lehetnének (mint például „nagyon spontán”, „alig spontán”, „részben spontán”, „túlnyomórészt spontán”...). Ennek a feltételezett skálának a fokozatai azonban nehezen (talán egyáltalán nem) meghatározhatók, még nehezebben általánosíthatók, és az alkalmazott fonetika területein szükségtelenek.

A beszéd természetességéről

Wolfson szerint a *természetes beszéd* nem más, mint megfelelő beszéd, azaz szemben áll a beszélő szándékának nem teljesen megfelelő beszéddel (1976). Úgy is fogalmazhatunk, hogy a beszéd akkor tekinthető természetesnek, ha a beszélőnek nem okoz feltűnő, észlelhető nehézséget a folyamat működtetése (például a dadogó beszédet vagy a rekedt hangú beszédet ezért állítják szembe a természetes beszéddel). (A nem patológiás beszéd esetén azonban probléma, hogy minek alapján állítjuk, hogy egy adott beszélőnek nincs különösebb nehézsége a verbális kommunikációban, vagy hogy a megnyilatkozás nem igényelt tőle többletenergiát.)

Az emberi agy kiválóan alkalmazkodott a verbális kommunikáció mellett számos új kommunikációs forma megtanulásához, újabb kommunikációs stratégiák kialakításához (ilyen például az olvasás: Csépe 2006). Képesség váltunk ezek „természetes” alkalmazására. Mit értünk tehát a természetességen a beszéd vonatkozásában? Nusbaum és munkatársai úgy fogalmazzák, hogy a természetesség a beszéd multidimenzionális, szubjektív minősége (1995). Ennek a gondolatnak az alapján elméletileg azt mondhatjuk, hogy a beszéd akkor természetes, ha megfelel a beszélő személyiségének, egyéniségének, szokásos beszédének. Nem tudunk azonban egzakt fonetikai paramétereket kapcsolni a természetes beszéd egyes jellemzőihez. Az sem segítene sokat, ha definiálnánk ilyen „kapcsolatokat”, hiszen nem volna etalon, amelyhez viszonyíthatnánk őket egy ismeretlen beszélő esetében. Egy adott történetet például különbözőképpen (más-más stílusban, szóhasználat, hangképzéssel, prozódiaival stb.) mesélünk el a szüleinknek, idősebb vagy fiatalabb idegennek, rokonunknak, a „haverunknak”, a közeli barátunknak. Ennek ellenére senki sem mondaná azt, hogy a felsorolt helyzetek bármelyike ne felülne meg a természetesség kritériumainak.

Hogyan közelíthető meg tehát a természetes beszéd fogalma? A beszédprodukciónak startégiák alkalmazását igénylő folyamat (Levelt 1989). A beszédstratégia voltaképpen egy általános séma, amely tartalmazza a beszélő szándékát, a beszédtervet és azt a módot, ahogyan a közlendőjét megformálja az adott hallgató(k)nak. Feltételezhetjük, hogy a természetességnek több tényezője van. Ilyen lehet a beszélő személyisége: van, akit jobban zavar egy adott kommunikációs helyzet, van, akit kevésbé vagy egyáltalán nem. A „közös” beszédtema, a kapcsolódó területek ismerete, és az átadás módja is hatással lehetnek a beszéd természetességre. Valószínűleg nem természetes például az a beszélgetés, amelynek során a beszélő folyamatos meghatározásokat, részletes elemzéseket kell, hogy adjon a hallgatónak a megértés érdekében. A beszélő – többnyire tudat alatt – alkalmazkodik a beszédpartnerhez, ez befolyásolhatja beszédének megformálását (vö. Bata–Grácsi 2009), s ez összefügghet a természetességgel is.

A nem feszélyezett (nem szorongó) viselkedés nyilvánvalóan kapcsolatos a természetes beszéddel; a szorongás mértéke és a beszéd természetessége közötti összefüggés azonban számos további kérdést vet fel (hiszen a mindennapi verbális kommunikáció során is adódhatnak olyan körülmények, amelyek között a beszélő nincs könnyű lélektani helyzetben). A Grice (1975) szerinti kooperációs elv szabályozza a társalgást, és valószínűleg biztosítja vagy lehetővé teszi a beszéd természetességét. Ez az elv négy maximából áll: a mennyiségi, a minőségi, a relevancia-, valamint a módmaximából. Grice nem tartotta rendszerét lezártnak, ő maga is kiegészítette ezeket további elvekkel, ilyenek például az esztétikai, a társadalmi vagy a morális maximák. Leech (1983) megalkotta az udvariassági elvet, amely újabb maximákat tartalmaz (mint például tapintatmaxima, nagylelkűség-maxima, jóváhagyás-maxima, szerénységmaxima, megegyezésmaxima, rokonszenvmxima).

A beszéd természetessége az alkalmazott fonetikai területeken

A beszéd természetességének az előadó-művészi, illetve a retorikai gyakorlatban nagy jelentősége van, mivel a beszélő (itt: a színész, a szónok) verbális viselkedése alapozza meg a megformált személy hitelességét (ehhez egyéb viselkedésformák is társulnak, pl. mimika, gesztusok). A színész azonban előre megtanult szöveget hangosít meg, a szónok pedig rendszerint felkészül a mondandójára, ezért a beszéd természetességét másként ítélik meg náluk, mint az átlagos, ún. naiv beszélő esetében. A beszéd természetessége az alkalmazott fonetikában az utóbbi évtizedekben elsősorban a következő kutatások kapcsán merült fel: *a*) a beszéd szintézis (azaz a beszéd mesterséges előállítás), *b*) a mesterséges beszéd felismerés, *c*) a törvényszéki fonetika (a beszélő személy felismerésének kérdése) és *d*) a beszéd korpuszok fejlesztése.

a) A természetesség a beszéd szintézisben. Az első szintetizáló rendszerek (nyelvtől függetlenül) erősen gépies hangzásúak voltak; a kezdeti technológia (főként a korlátozott tárkapacitás) nem tette lehetővé, hogy a fejlesztők a

szegmentális szerkezet meghangosításán túl egyéb paraméterekkel is hatósan foglalkozhassanak (Olaszy 1989). A formánsalapú szintetizátorok beszéde jól érthető, hangzásuk azonban erősen gépies (Tóth–Németh 2008). A technológia fejlődésével a hangzás egyértelműen javult, a hangzás gépiessége csökkent (van Santen et al. 1997, Olaszy et al. 2000). A beszédszintetizátorokkal előállított hangot az érthetőség, illetve a természetesség kívánalmaitól mentén szokták minősíteni. A beszédszintetizálási módszerek változása, valamint a tárkapacitás növekedése mindenekelőtt a prozódiai jellemzők beépítését tette lehetővé, és ezáltal a mesterséges beszéd már jobban megközelítette az emberit (Fék et al. 2005, Németh et al. 2006). Noha ez minőségi változást idézett elő a beszédszintézis hangzásában, a „természetesség” problémáját nem oldotta meg. Az újabb kutatások az érzelmi komponens beépítésével, illetőleg annak prozódiai leképezésével igyekeznek a beszéd gépies hangzását tovább csökkenteni, mind jobban az emberi beszéd hangzását szimulálni (Németh et al. 2006, Zainkó–Fék 2006).

A természetesség megítélését az észlelésben – számos egyéb tényező mellett – a zöngéképzés vagy a szótag prozódiai szerkezete befolyásolta a kísérleti adatok szerint (Nusbaum et al. 1995). Úgy tűnik, viszonylag egyszerű azt megmondani, hogy a mesterségesen előállított beszéd mennyire vagy milyen tekintetben természetes, hiszen az adatközlők a humán beszédprodukciónál viszonyíthatnak. Az ember-gép kommunikáció fokozódó igényei következtében napjainkban még jobban előtérbe került ez a kérdés. A legjobban érthető és a legtermészetesebbnek ítélt szintetizált beszédet korpuszalapú elemkiválasztásos módszerrel állítják elő (Tóth–Németh 2008).

b) A természetesség a mesterséges beszédfelismerésben. A beszédfelismerők jó hatásfokkal működnek izolált szavak, felolvasott szövegek, illetve egy adott személy beszédére betanított rendszerek esetében (pl. Becchetti–Ricotti 1999, Mihajlik et al. 2006, Vicsi et al. 2007, Tüske et al. 2007). A spontán beszéd létrehozásának következtében létrejövő sajátosságok (mint pl. a megakadásjelenségek) azonban problémát jelentenek a fejlesztőknek (Schaaf–Kemp 1997, Seneff 2004). A folyamatos spontán közlések mesterséges felismerése ugyanis rugalmas akusztikai modellálást, mondathatár-detektálást, rugalmas kiejtésmodellálást, megfelelő akusztikai és nyelvi modell alkalmazását és mindezek automatikus összegzését teszi szükségessé (vö. Furui 2005 áttekintését).

c) A természetesség a beszélő személy felismerésében. Az emberi hang magában rejtja a beszélő személy egyedi jellemzőit is, ennek kutatásával az utóbbi évtizedekben kezdtek behatóan foglalkozni elsősorban a törvényszéki fonetika területén (Braun–Köster 1995, Dankovičová–Nolan 1999, Butcher 2002, Nikléczy 1996). A kiinduló kérdés az volt, hogy a beszéd szegmentális vagy szupraszegmentális szintjén keresendők-e az egyéni beszédre utaló összetevők, avagy mindkettő tartalmaz ilyeneket. Ha ismerjük a beszélőt, akkor rendszerint meg tudjuk mondani, hogy természetesen beszél-e vagy

nem; nem biztos azonban, hogy tényeket is fel tudunk sorolni véleményünk alátámasztására. Egy nem ismert (idegen) személy esetében ráadásul nincs is mihez viszonyítanunk. Valamennyi segítséget jelent ebben az ún. alkati harmónia, amely a beszélő anatómiai felépítésének megfelelő hangadást jelent (Gósy 1999).

A beszélőre jellemző sajátosságok részben anatómiailag meghatározottak, részben a beszélőnek az adott helyzetben létrehozott artikulációs mozgásaival kapcsolatosak. Ez azt jelenti, hogy a létrehozott beszéd artikulációs, illetve akusztikai fonetikai jellemzői egyfelől szervi eredetű, másfelől tanult különbözőségekre vezethetők vissza. A szakirodalom prototípusmodellje egy prototipikus hangzást tételez fel a hallgató memóriájában, amelyhez a hallgató minden új beszélő esetében viszonyít (Lavner et al. 1999). A természetesség kérdésre azonban ez a teória nem ad választ. Nem lehet ugyanis tapasztalunk minden ember beszédéről ún. természetes és nem természetes környezetben. A kutatási eredmények szerint az agy bizonyos kérgi területe a különböző emberi zöngéhangok megkülönböztetéséért felelős (Belin et al. 2000). A kompenzációs adaptációs elmélet (Kock 2007) elsősorban a kommunikációs helyzet (és nem a beszéd) természetességével foglalkozik. Azt állítja, hogy a kommunikáció természetességének fokozatai attól függenek, hogy egy adott médiumban beszélők milyen mértékben alkalmazzák a beszédükben azokat az elemeket, amelyeket a szemtől szemben álló beszélgetőpartnerek szoktak alkalmazni. Ezeknek az elemeknek a megnevezésével azonban adós marad.

d) A természetesség a korpuszokban. Beszédkorpuszok fejlesztésekor felmerül, hogy a rögzített anyagok mennyire tekinthetők természetesnek, vagyis olyanoknak, amilyenek a mindennapi életben előfordulnak. Minthogy a beszéd-felvételek valamilyen mértékben mindig mesterséges környezetben zajlanak, az egyik legfontosabb kérdés, hogy a beszélő mennyire képes ekkor természetesnek, vagyis a megszokott módon beszélni. Vajon milyen mértékben kontrollálja a közléseit, mennyire képes függetlenedni a mikrofontól, illetve a számára többé-kevésbé ismeretlen beszélgetőpartner(ek)től? Itt említjük meg Labovnak a szociolingvisztikai interjúk rögzítése kapcsán bevezetett „megfigyelői paradoxon” fogalmát (1979).

A beszédkorpuszok fejlesztésének általános célja annak feltérképezése, hogyan beszélnek az emberek a mindennapokban. A kérdés az, hogy a korpuszok rögzített anyagai valóban a mindennapok verbális kommunikációját tükrözik-e. Ennek megítélésében sokszor keveredni látszik a spontán beszéd és a természetes beszéd fogalma. Először a természetesség kérdését közelítjük meg. Ha az adatközlők tudják, hogy éppen a beszédüket figyelik, akkor nyilvánvalóan jobban ügyelnek annak tartalmi és formai sajátosságaira. Ez pedig gátolhatja a beszélőt abban, hogy úgy beszéljen, mint a mindennapokban. Ha a mindennapok verbális kommunikációja természetes, akkor vajon ebben az értelemben természetes-e a beszédkorpuszok rögzített anyaga? „Minél jobban

biztosítjuk [...] az adat-, illetőleg a szövegrögzítés technikai föltételeit [...] a kísérleti effektus miatt annál inkább eltávolodnak az adatközlők a megszokott nyelvhasználatuktól, a kutatók pedig egyre távolabb kerülnek eredeti céljuk elérésétől” – írja Kiss Jenő (1995: 37). A megfigyelői paradoxon áthidalásának egyik lehetséges módja az, ha az adatközlőtől többféle típusú beszédet, többféle módon gyűjtünk (Kontra 1990, Wardhaugh 1995). Labov megoldásnak tekinti továbbá azt, hogy csoportinterjúkat készítsünk, mivel ezt tekinti a kommunikáció legtermészetesebb módjának. Véleménye szerint a megfigyelési hatás itt érvényesül a legkevésbé. Az interjút készítő személynek alárendeltként kell „vezetnie” a beszélgetést, és éreztetnie kell, hogy az ő tekintélye kisebb, mint az adatközlőké (Labov 1984).

A beszédkorpuszok felvételi módszereinek sokfélesége tehát azt a célt is szolgálja, hogy minél „természetesebb” helyzetbe kerüljön az adatközlő. A játékok, társasjátékok alkalmazása (pl. Szabó 2004) azért jó, mert valamennyire eltereli a figyelmet a beszédfelvétel tényéről. A feladat jellegű beszéd-rögzítések (pl. a térképmódszer: Horváth 2004) előnye, hogy az adatközlők hosszabb-rövidebb idő elteltével a feladatra és nem a saját verbális viselkedésükre koncentrálnak. A dialógusok, a narratívák, a társalgások, a viták, az értekezletek vagy a hallott történet, illetve látott film reprodukálásának vagy megbeszélésének természetessége sok tényezőtől függ [pl. az adatközlő(k) és az adott téma kapcsolata, mennyire zavarja a beszélőt a mesterséges helyzet, mennyire képes feloldódni, milyenek a beszédpartnerek stb.]. Nemegyszer a felvételi technika rovására rögzítenek beszédanyagokat éppen a természetesség mind jobb megközelítése érdekében (pl. otthonokban, mindennapi helyszíneken, utcán, iskolában). Kérdés azonban, hogy az ismert vagy szokásos helyszín, avagy a nem személyes kapcsolat (telefon) valóban növeli-e a beszéd természetességét (pl. Hutchinson–Pereira 2001, Keszler 1983, repülőjegy-rendeléssel kapcsolatos beszélgetések).

A beszédkorpuszok fejlesztése kapcsán felmerülő másik kérdés a spontaneitás. (A jelen tanulmányban nem térünk ki az olvasásalapú beszédkorpuszokra.) Láttuk, hogy különféle módjai vannak a nem olvasott beszédanyag-felvételeknek. Az ilyen módon rögzített anyag spontán beszéd, hiszen a beszélőnek nincs módja az adott témára felkészülni. A gondolatait abban a helyzetben, azonnal kell verbalizálnia. A mesterséges szituáció, illetve a beszéd-rögzítés ténye hatással lehet (általában negatívan) a beszéd folyamatosságára (pl. hezitálások arányának növekedése) vagy a megszokott légzéstechnikára. (Ezek a mindennapi kommunikációban tapasztalható jelenségek, mint amilyen a hezitálás is, úgy tűnik, egyszersmind a természetesség benyomását növelik, vö. Csapó et al. 2009.) A beszéd-rögzítéses helyzet befolyásolhatja továbbá a lexikai válogatást (ilyen helyzetben nyilvánvalóan jobban odafigyel az adatközlő arra, hogy milyen szavakat mond ki), valamint valamilyen mértékben a kiejtést is (vö. Lindblom 1990: hiperejtés és hipoejtés). Szende szerint (1973) egy ilyen helyzetben a beszélők törekszenek a nyelvi norma

megközelítésére, sokkal inkább, mint a mindennapi megnyilatkozásokban. Mindez rendszerint jogos feltételezés, de nehéz a megállapítást általánosítani és objektíven igazolni.

Spontaneitás és természetesség a BEA tapasztalatai alapján

Láttuk, hogy a spontán beszédet tartalmazó korpuszok esetében merül fel újra és újra a beszéd természetességének kérdése, amely azonban (mint említettük) gyakran keveredik a spontán beszéd fogalmával. A természetességet nemegyszer a spontaneitással rokon értelmű fogalomként kezelik, így összekeveredik a beszédtervezési folyamat jellege és a verbális nyelvhasználat minősége.

A BEA (BESzélte nyelvi Adatbázis, vö. Gósy 2008) célja különféle spontán beszédek rögzítése stúdiókörülmények között, amely alkalmassá teszi a beszédmintákat akusztikai fonetikai elemzésekre is. Egy felvétel közel 60 percig tart, különböző részekből áll, az adatközlőnek van ideje „hozzászokni” a felvételi körülményekhez, az interjúkészítőhöz, így beszéde és viselkedése egy bizonyos idő elteltével (közel) természetessé válik. A felvételi protokoll kialakításánál törekedtünk arra, hogy az egyes részek egymásra épülésének sorrendje már önmagában is segítse az adatközlőt az esetleges kezdeti szorongásának leküzdésében.

A **spontaneitás** nem sérül egy mesterségesen létrehozott kommunikációs helyzetben, például egy interjú során, hiszen a beszélő az adott pillanatban válogat a gondolatok között, és rendeli hozzájuk a nyelvi formát. Az adatközlő ekkor sokkal inkább a közlés tartalmára fordít figyelmet, kevésbé koncentrálna a nyelvi átalakításra és legkevésbé az artikulációra. A rögzített beszéd tehát egyértelműen spontán akkor, ha a beszélőnek a feltett kérdés(ek)re felkészülés nélkül azonnal reagálnia kell, illetve a társalgásban minden előzetes megbeszélés és tervezés nélkül vesz részt (a társalgás során felmerülő témák váratlanok a számára). A beszéd spontaneitásának mértéke természetesen változhat az adott téma vagy az adott feladat függvényében, de ez adatközlő-specifikus.

A BEA spontánbeszéd-anyagai is értelemszerűen különböző fokozatúak lehetnek egy „spontaneitási skálán”. A beszédprodukción feltehetően a legnagyobb mértékben akkor spontán, amikor az interjút készítő egy váratlan problémáról kérdezi az adatközlő véleményét. Némelykor az adatközlő erre reagál is, és ez a beszédének egyre természetesebb formájához vezet. Például: *jaa igen ez igaz na látod milyen praktikus kérdés tényleg hát nem tudom...* Valamivel kevésbé lehet spontán a narratíva típusú beszéd, amikor az adatközlő a családjáról, a munkájáról, a hobbjáról beszél, hiszen ilyenkor óhatatlanul aktiválja a témával kapcsolatos tartalmi és szerkezeti kliséket. A beszélő ilyenkor is „válogat”, átgondolja, hogy mit mondjon, mit hagyjon ki, miről beszéljen és mennyit. (Nincs tudományos kutatási eredmény arra vonatkozó-

an, hogy a beszédtervezési folyamatok milyen különbségeket mutatnak a mindennapi kommunikációban és egy beszédrögzítéses helyzetben.)

A BEA-felvétel egyik szakaszában két elhangzó szöveget kell az adatközlőnek saját szavaival összefoglalva elmondania. Ezek a narratívák egyértelműen spontánok ugyan, de a mondandó tartalma adott (hiszen azt a beszélő megelőzően hallotta), így a beszédtervezési folyamatok valamilyen mértékben másként működnek, mint az előzőekben említett narratívák vagy a társalgások során. Nyilvánvalóan spontán közléseket hallhatunk a társalgásokban, hiszen az adott témára nincs mód előzetesen felkészülni, ugyanakkor a társalgás sajátosságai mégis több időt hagynak az adatközlőnek a beszédtervezésre (pl. Markó 2007), mint a narratívák esetében. További kutatás adhat zámat arról, hogy mindennek vannak-e nyelvi és/vagy kiejtési következményei, és ha igen, milyenek.

A spontánbeszéd-adatbázis interjúinak készítésekor a beszédszándék nem a beszélőtől indul ki, hanem mesterségesen hozzák létre oly módon, hogy felkérjük a személyt a részvételre. Megjegyezzük, hogy az adatközlők nemegyszer átveszik az irányítást, ami növeli a beszéd spontaneitását, s egyúttal feltételezhetően a természetességét is. Például az egyik adatközlő, amikor a felvételvezető elővezette a beszéd témát, a következőképpen reagált: *na jó akkor kezdjék maguk*. Egy másik példa: *akkor viszont győri megkérdezem magától hogy a győri történetek közül tudja-e azt hogy a Széchenyi téren a bencés templommal szemben van egy Mária-szobor és ez a Mária-szobor arról nevezetes, hogy nagypénteken délben amikor harangoznak a Mária meghajtja a fejét miért*.

A szorongás hatással van a spontán beszéd folyamatosságára és a beszéd természetességére, de nem a spontán voltára. A mikrofon zavaró tényező (Szende 1973), hiszen azt jelzi, hogy az elhangzott közlés visszavonhatatlan, korrigálhatatlan, megőrződik. Ez (hosszú távon) nem jellemző a természetes beszédhelyzetekre, ezért erre az adatközlők nincsenek felkészülve, nincsenek stratégiáik az ebből fakadó zavar ellensúlyozására. A legtöbb emberben szorongást kelt, hogy felveszik és később különböző szempontokból elemzik majd a beszédét. Úgy érzik, hogy bizonyos elvárásoknak kell megfelelniük, bár a felvételeket megelőzően a kísérletvezető informálja az adatközlőt (biztosítja, hogy anonim marad a felvétel stb.), és ez általában csökkenti is a beszélő szorongását. A beszédrögzítés tényéből adódó zavar azonban nem egyformán érint mindenkit; akadnak, akikre ez nagyobb mértékben hat, míg mások (láthatólag) egyáltalán nem foglalkoznak vele, látszólag nem frusztrálja őket az adott helyzet. A mikrofon okozta esetleges beszédmódosulások az ő esetükben nehezen vagy egyáltalán nem megítélhetők.

A BEA-felvétel előtt a beszélők egy része bevallja, hogy nagyon izgul (sokszor a testtartáson, a mozdulatokon is látható a feszélyezettség). Ezek a külsődleges jelek viszonylag gyorsan megszűnnek, az adatközlők hamar feloldódnak. Ezt utólag (jóleső érzéssel) el is mondják. Nemegyszer kapunk

olyan visszajelzést, hogy mennyire jó volt beszélgetni, a beszélőnek – saját maga számára is – meglepően gyorsan sikerült feloldódnia, belehelyezkednie a beszélgetésbe, a témába; a kérdések, a társalgás, a beszédpartnerek és az egész légkör olyan volt, hogy nagyon gyorsan elterelte a figyelmét a hangrögzítés és az idegen helyzet tényéről. Egy beszédfelvétel hasonlóan bizalmi szituáció, mint például egy orvossal vagy egy pszichológussal történő beszélgetés. A közös munka csak akkor hatékony, ha a kommunikációs partnerek rövid idő alatt egymásra tudnak hangolódni, és kialakul közöttük ez a bizalmi viszony. Ez esetünkben azt is jelenti, hogy az adatközlő nem mérlegeli minden hangos megnyilatkozását, hanem a mindennapokhoz hasonlóan beszél. Ez pedig a beszéd természetességét növeli.

Mindebben a felvételvezetőnek óriási szerepe van. Egyfajta intuitív, empátikus készséggel kell rendelkeznie ahhoz, hogy igen rövid idő alatt elnyerje a beszélő bizalmát, és mintegy „régis ismerősként” beszélgessen vele. Már a beszélgetés kezdetén fel kell tudnia mérni a beszélgetőpartner „szorongási szintjét”, és ennek megfelelően igyekeznie kell oldani azt. Ennek érdekében lehetősége van például felcserélni az egyes részek sorrendjét. Többször előfordult, hogy az ismeretlenként odaérkező adatközlő a felvétel után megkérdezte, hogy vajon voltak-e róla esetleg előzetes információk, mert a felvételvezető olyan témákat hozott fel, amik őt nagyon foglalkoztatják mostanában. Egy másik esetben a beszélgetés során az adatközlő azt mondta, hogy bár tudja, hogy ezt a felvételt kutatók meg fogják hallgatni és elemzik, de akkor is elmondja, hogy... És még egy példa. Megtörtént, hogy a felvetett kérdésre az adatközlőnek az eredeti véleménye a beszélgetés végére az ellenkezőjére változott. Ez pedig csak úgy lehetséges, hogy közben nem a felvétel tényére figyelt, hanem teljesen lekötötte az adott téma, a véleménykülönbség, illetve a saját hozzáállásának a módosítása. Ezek a „vallomások”, vélemények, megjegyzések és tények megerősítik azt, hogy a felvételek legalábbis közel természetes beszédet tartalmaznak. Minden felvétel előtt van továbbá egy úgynevezett „próba felvétel”, ami egyfajta „bemelegítésként” szolgál az adatközlő számára, ezalatt van ideje hozzászokni a mikrofon jelenlétéhez.

A témaválasztás a mindenkor adatközlő valószínűsíthető személyes érdeklődéséhez igazodik. A felvételvezető legtöbbször közfelháborodást kiváltó, avagy a közvéleményt foglalkoztató aktuális (nem politikai jellegű) eseményekről kérdezi az interjúalany véleményét; figyelembe véve a labovi elvet (1984), végig „alárendelt” szerepben maradva. A téma felvetéskor is tekintettel vagyunk az adatközlőre; ha a választott téma nem tetszik neki, láthatólag nem akar beszélni róla, akkor az interjúkészítő azonnal vált. A témák a beszélőből gyakran érzelmeket is kiváltanak, ami nagyban hozzájárul ahhoz, hogy megfeledkezzenek a felvétel rögzítésének tényéről. Nemegyszer előfordult például, hogy a véleménykifejtés hevében trágár kifejezés „csúszott ki” az adatközlő száján, és csak utána döbbsent rá, hogy ez egy rögzített felvételkor történt: *és akkor mondtam hogy hát mondom akkor most elmennek a bűdös*

*pi**ba így széttéptem a bocsánat a felvétel jaj jaj jaj kiszaladt na mindegy.* Előfordul, hogy a beszélő a véleménykifejtés folyamán elakad, megszakad a gondolatmenete, illetőleg a felvételvezető úgy ítéli meg, hogy túl szűkszavúan nyilatkozott az adott témában. Ilyenkor – az adatközlő személyiségétől függően – szándékosan ellenkező nézőpontú kérdésekkel és véleményformálással igyekszik továbbléteni a társalgást. Mindez megint csak eltereli a figyelmet a mikrofon jelenlétéről.

Amikor azt fejtegetjük, hogy egy spontánbeszéd-felvétel során mennyire beszél természetesen az adatközlő, voltaképpen arra keressük a választ, hogy mennyire hasonlít a beszéde ahhoz, mint amikor informális helyzetben, a mindennapokban, például családi körben beszél. Ha lenne ilyen informális helyzetben gyűjtött beszédanyagunk, akkor lehetne válaszolni erre a kérdésre. Ilyen beszédanyag hiányában csupán a felvételvezetői tapasztalatra támaszkodhatunk.

Összegző megállapítások

A spontán beszéd mibenléte egyértelműen meghatározható, szemben a beszéd természetességének fogalmával, ez utóbbi – mint láttuk – nehezen definiálható, sok bizonytalanságot rejt magában. A kettő nemegyszer keveredik egymással, ami a beszédkutatásban számos félreértésre adhat okot. A felkészülés nélküli beszéd mindig spontán, jól jellemezhető sajátosságokkal. A beszéd természetességének megítéléséhez számos tényező figyelembevétele szükséges, és jelenleg nem rendelkezünk ehhez objektív kritériumrendszerrel. Az emberek eltérően beszélhetnek mind a spontaneitás, mind a természetesség tekintetében, noha ezen eltérések jellemzése valószínűsíthetően bonyolult feladat. Az ilyen irányú kutatások megalapozásához igyekeztünk a jelen tanulmánnyal hozzájárulni.

Irodalom

- Bata Sarolta – Gráczy Tekla Etelka 2009. Hatással van-e a beszédpartner életkora a beszélő beszédének szupraszegmentális jellegzetességeire. In Keszler Borbála – Tátrai Szilárd (szerk.): *Diskurzus a grammatikában, grammatika a diskurzusban*. Tinta Kiadó, Budapest, 74–83.
- Becchetti, Claudio – Ricotti, Lucio Prina 1999. *Speech recognition and C++ implementation*. John Wiley and Sons, Rome.
- Belin, Pascal – Zatorre, Robert J. – Lafaille, Philip – Ahad, Pierre – Pike, Bruce 2000. Voice-selective areas in human auditory cortex. *Nature* 403. 309–312.
- Braun, Angelika – Köster, Jens-Peter (eds.) 1995. *Studies in forensic phonetics*. Wissenschaftlicher Verlag, Trier.
- Butcher, Andrew R. 2002. Forensic Phonetics: Issues in speaker identification evidence. Paper at the Inaugural International Conference of the Institute of Forensic Studies: “Forensic Evidence: Proof and Presentation”. Prato, Italy 3–5 July (CD-ROM).

- Csépe Valéria 2006. *Az olvasó agy*. Akadémiai Kiadó, Budapest.
- Dankovičová, Jana – Nolan, Francis 1999. Some acoustic effects of speaking style on utterances for automatic speaker verification. *Journal of the International Phonetic Association* 29. 115–129.
- Csapó Tamás – Zainok Csaba – Németh Géza 2009. Szintetizált beszéd prozódiai változatosságának növelése spontán beszéd alapján. Konferencia-előadás. Beszédkutatás 2009, MTA Nyelvtudományi Intézet, Budapest.
- Fék Márk – Olasz Gábor – Szabó János – Németh Géza – Gordos Géza 2005. Érzelem kifejezése gépi beszéddel. *Beszédkutatás 2005*. 134–144.
- Furui, Sadaoki 2005. Recent progress in corpus-based spontaneous speech recognition. *IEICE Transactions on Information and Systems*. E88-D(3): 366–375.
- Gósy Mária 1999. Az egyéni hangszínezet és a beszélő felismerésének kísérleti- fonetikai megközelítése. *Magyar Nyelvőr* 123. 424–438.
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás 2008*. 194–207.
- Grice, H. Paul 1975. Logic and Conversation. In Cole, Peter – Morgan, Jerry L. (eds.): *Syntax and Semantics: Speech Acts*. Academic Press, San Diego, 41–58.
- Horváth Viktória 2004. Megakadályozások a párbeszédekben. *Beszédkutatás 2004*. 187–199.
- Hutchison, Ben – Pereira, Cécile 2001. *Um, one large pizza*. A preliminary study of disfluency modelling for improving ASR. In Lickley, Robin – Shriberg, Elisabeth (eds.): *Disfluency in spontaneous speech. Proceedings*. Edinburgh, 77–81.
- Keszler Borbála 1983. Kötetlen beszélgetések mondat- és szövegtani vizsgálata. In Rácz Endre – Szathmári István (szerk.): *Tanulmányok a mai magyar nyelv szövegtana köréből*. Tankönyvkiadó, Budapest, 164–187.
- Kiss Jenő 1995. *Társadalom és nyelvhasználat*. Nemzeti Tankönyvkiadó, Budapest.
- Kock, Ned 2007. Media naturalness and compensatory encoding: the burden of electronic media obstacles is on senders. *Decision Support Systems* 44. 175–187.
- Kontra Miklós 1990. Az élőnyelvi vizsgálatok tárgyáról és módszereiről. In Szabó Géza (szerk.): *II. DialSzimp*. Veszprém, 131–138.
- Labov, William 1979. A nyelv vizsgálata társadalmi összefüggésben. In Pléh Csaba – Terestyéni Tamás (szerk.): *Beszédaktus – kommunikáció – interakció*. Tömegkommunikációs Kutatóközpont, Budapest, 365–398.
- Labov, William 1984. Field methods of the project on linguistic change and variations. In Baugh, John – Sherzer, Joel (eds.): *Language in use. Readings in sociolinguistics*. Prentice-Hall, Englewood Cliffs, N. KJ., 30–53.
- Lavner, Yizhar – Rosenhouse, Judith – Gath, Isah 1999. The prototype model in speaker verification. In Olasz, Gábor – Németh, Géza – Erdőhegyi, Krisztina (eds.): *Proceedings of Eurospeech '99*. Budapesti Műszaki Egyetem, Budapest, 771–774.
- Leech, Geoffrey 1983. *Principles of pragmatics*. Longman, London–New York.
- Levelt, Willem J. M. 1989. *Speaking. From intention to articulation*. A Bradford Book. Cambridge, Massachusetts.
- Lindblom, Björn 1990. Explaining phonetic variation: a sketch of the H&H theory. In Hardcastle, William J. – Marchal, Alain (eds.): *Speech production and speech modeling*. Kluwer, Dordrecht, 403–439.

- Lustgarten, Paul C. – Juang, Biing-Hwang 2002. Naturalness in speech communications. <http://users.ece.gatech.edu/~juang/Publications/Naturalness> (Az utolsó leltöltés ideje: 2009. augusztus 26.)
- Markó Alexandra 2007. A társalgás hangtana. *Magyar Tudomány* 167. 54–57.
- Mihajlik Péter – Fegyő Tibor – Tatai Péter 2006. Új eljárás gépi beszédfelismerés környezetfüggő beszédhangmodelljeinek kialakítására. *Beszédkutató 2006*. 218–231.
- Németh Géza – Olasz Gábor – Fék Márk 2006. Új rendszerű, korpuszalapú gépi szövegfelolvasó fejlesztése és kísérleti eredményei. *Beszédkutató 2006*. 183–196.
- Nikléczy Péter 1996. Beszélő személy azonosítása szűkfrekvenciás szavak alapján. *Beszédkutató '96*. 20–31.
- Nusbaum, Howard C. – Francis, Alexander L. – Henly, Anne S. 1995. Measuring the naturalness of synthetic speech. *International Journal of Speech Technology* 1. 7–19.
- Olasz Gábor 1989. *Elektornikus beszédelőállítás*. Műszaki Kiadó, Budapest.
- Olasz Gábor – Németh Géza – Olasz Péter – Kiss Géza 2000. Profivox – a legkorszerűbb hazai beszéd szintetizátor és szövegfelolvasó. *Beszédkutató 2000*. 167–179.
- van Santen, Jan H. P. – Sproat, Richard W. – Olive, Joseph P. – Hirschberg, Julia (eds.) 1997. *Progress in speech synthesis*. Springer, New York–Berlin.
- Schaaf, Thomas – Kemp, Thomas 1997. Confidence measures for spontaneous speech recognition. *International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*. Vol 2. 875–878.
- Seneff, Stephanie 2004. The use of subword linguistic modeling for multiple tasks in speech recognition. *Speech Communication* 42/3–4. 373–390.
- Szabó Eszter 2004. Megakadásjelenségek nyelvi játék közben. *Beszédkutató 2004*. 200–208.
- Szende Tamás 1973. *Spontán beszédanyag gyakorisági mutatói*. Nyelvtudományi Értekezések 81. Akadémiai Kiadó, Budapest.
- Tóth Bálint – Németh Géza 2008. Rejtett Markov-modell alkalmazása magyar nyelvű gépi szövegfelolvasáshoz. *Beszédkutató 2008*. 182–193.
- Tüske Zoltán – Simon Márta – Mihajlik Péter – Fegyő Tibor 2007. Érzelmek automatikus felismerése a beszéd akusztikus jellemzői alapján. *Beszédkutató 2007*. 151–162.
- Vicsi Klára – Szaszák György – Németh Zsolt 2007. Folyamatos magyar beszéd mondatfajtáinak automatikus felismerése. *Beszédkutató 2007*. 162–172.
- Ward, Wayne 1989. Understanding spontaneous speech. In: *Proceedings of the Workshop on Speech and Natural Language*. Association for Computational Linguistics, Morristown, New Jersey, 137–141.
- Wardhaugh, Ronald 1995. *Szociolingvisztika*. Osiris–Századvég. Budapest.
- Wolfson, Nessa 1976. *Speech events and natural speech: Some implications for sociolinguistic methodology*. Cambridge University Press. Cambridge.
- Zainkó Csaba – Fék Márk 2006. Beszédatadabázis prozódiajának szerepe a gép beszéd hangzásában. *Beszédkutató 2006*. 208–217.