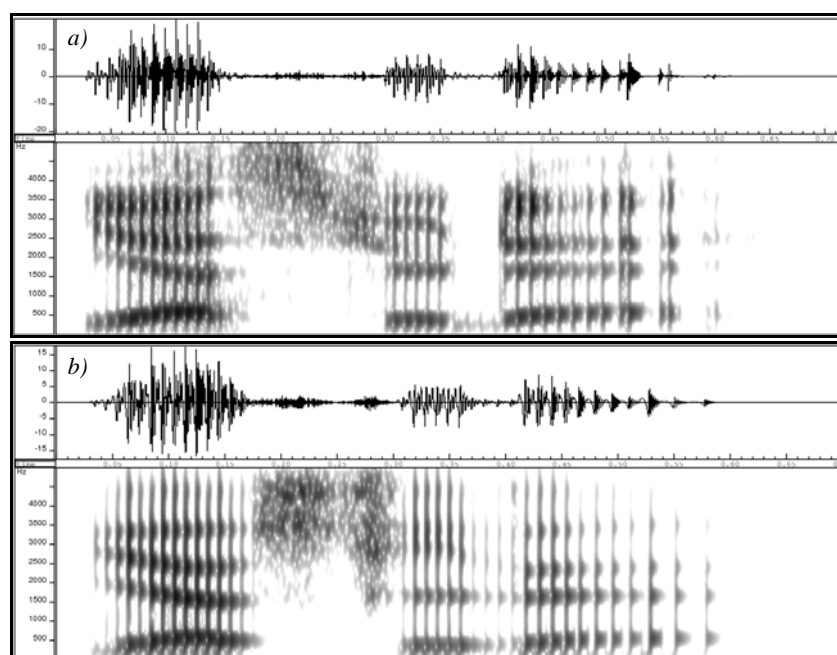


A GLOTTALIZÁCIÓ SZEREPE A BESZÉLŐ SZEMÉLY FELISMERÉSÉBEN

Bóhm Tamás

Bevezetés

A magánhangzók és zöngés mássalhangzók artikulációjának alapja a zöngé (Gósy 2004: 28). A zöngéképzés, a fonáció során a hangszalagok általában közel periodikusan (kváziperiodikusan) rezegnek. Ilyenkor a hangszalagok nagyjából állandó időközönként összecsapódnak – a rezgés reguláris. Rövidebb-hosszabb ideig azonban ez a rezgés lehet irreguláris. Ekkor a hangszalagok összecsapódásai között eltelt idő széles határok között ingadozik, és általában jóval hosszabb. Ezt a jelenséget nevezzük glottalizációnak (1. *a*) ábra).



1. ábra

A *yesterday* szó férfiejtésben, glottalizált végződéssel:

a) természetes beszéd, *b)* másolatszintézis (l. a hanganyag ismertetésénél)

A glottalizációt a szakirodalomban többféleképpen definiálták (Laver 1980: 93–95; Huber 1992; Slifka 2000: 100–101; Redi–Shattuck-Hufnagel 2001). Ebben a tanulmányban az „érzékeltetően irreguláris hangszalagrezgést” tekintjük glottalizációnak. Az irregularitás jelentkezik a rezgés pillanatnyi frekvenciájában, amplitúdójában vagy mindkettőben. Az ingadozás mértékének elég nagyok kell lennie ahhoz, hogy a hallgató számára érzékeltetően eltérjen a reguláris zöng hangzásától. Ez a kikötés azért fontos, mert reguláris zöngképzés esetén is csak kváziperiodikus rezgésről beszélhetünk, aminek frekvenciája és amplitúdója kis mértékben ingadozik (Gordos–Takács 1983: 24).

A glottalizációt – a fenti definíció szerint – akusztikai jelenségnek tekintjük, nem feltételezünk semmit a produkciójáról vagy percepciójáról (az említett érzékeltetésen felül). Ettől függetlenül természetesen vannak produkciós és percepció-s vonatkozásai is, ezeket a következőkben foglaljuk össze.

A glottalizáció produkcióját hagyományosan a hangszalagok szoros összeszorításával magyarázzák (Laver 1980: 122–126), ami a reguláris fonációnál erősebb záródást eredményez. Ezért a rezgés kevésbé stabil, frekvenciája és amplitúdója ingadozhat. Slifka (2000) azonban a tüdő nyomásszintjének közvetett mérésével kimutatta, hogy glottalizációt a hangszalagok összeszorításával és ellazításával is elérni lehet.

A glottalizációt érdekes, recsegő hangként érzékeljük. Az angol nyelvű szakirodalomban a hozzá tartozó érzeti jelenséget, a hangminőséget számos névvel illetik (*creaky, vocal fry, pressed* stb.). Catford megfogalmazásában a glottalizáció ahhoz hasonlít, amikor valaki egy rudat végighúz egy fémkerítésen (idézi Laver 1980: 122). Egy magyar tanulmány is kiemeli, hogy a szélsőségesen mélyen futó alaphfrekvencia-görbét nem a hang mélyüléseként, hanem nyikorgóvá, torzzá válásaként érzékeljük (Markó 2005: 61–62).

Több angol nyelvre és egy magyar nyelvre végzett vizsgálat is nagy különbségeket mutatott ki az egyes beszélők fonációs szokásaiban. Ezekből a tanulmányokból arra következtethetünk, hogy a glottalizáció gyakorisága jellemző a beszélő személyre. Redi és Shattuck-Hufnagel (2001) 14 amerikai angol anyanyelvű bemondó beszédét elemezte. Volt olyan adatközlő, aki a vizsgált esetek 88%-ában, míg egy másik személy csak 13%-ában glottalizált. A négy amerikai személy, akik Slifka (2000: 100–103) vizsgálatában szerepeltek, a bemondások 5–93%-ának a végén glottalizált. Valószínűleg magyar beszélők között is hasonló különbségek vannak: Markó (2005: 61) egy adatközlő spontán beszédében gyakran, a többi három felvételeiben csak néhányszor észlelt „nyikorgó zöngét”.

A glottalizáció az amerikai angolban gyakori jelenség intonációs frázisok elején (Redi–Shattuck-Hufnagel 2001), de több tanulmány szerint legnagyobb arányban a bemondások végén jelenik meg (Henton–Bladon 1987; Redi–Shattuck-Hufnagel 2001). Ez utóbbi kiemeli, hogy a beszélők között ezen a pozíción a legnagyobb a különbség. Slifka (2000) ezzel összhangban az egyes személyek közötti különbségeket szintén a bemondások végén mutatta ki.

Összefoglalva, a glottalizáció előfordulási aránya jellemző a beszélő személyére (elsősorban a bemondások végén). Ennek hátterében valószínűleg fiziológiai, szociolingvisztikai és egyéb hatások állnak.

Hipotézisek és a kísérlet alapelve

Kizárólag a hangjuk alapján az emberek képesek ismerős személyeket felismerni. Ebben a folyamatban számos egyéni ismertetőjegy játszhat szerepet. Ezek a beszéd olyan akusztikai jellemzői közül kerülhetnek ki, amelyek az egyes bemondókra következetesen eltérőek (Laver 1980: 9–10). Beszélőnként változik, hogy a bemondások vége milyen gyakran glottalizált (gyakran, ritkán, vagy néha igen, néha nem). Így ez olyan egyéni tulajdonság, amely segítheti a hallgatót a beszélő felismerésében. Az viszont nyitott kérdés, hogy ennek a jelenségnek tényleg van-e szerepe a felismerési folyamatban, azaz megjegyezzük-e ismerős személyek jellemző fonációs szokásait. Ezt a kérdést az alábbi két hipotézisben pontosítottuk:

1. A végső glottalizáció érzékelhető a hallgatók számára.

2. A végső glottalizáció szerepet játszik az ismert beszélő személy felismerésében – egyes beszélők és egyes hallgatók esetén.

Az első hipotézist két típusú hangzóanyagra már bizonyították: szintetikus kiktartott magánhangzókra és természetes szópárokra (Huber 1992). A jelen tanulmány egyik célja, hogy egy-két szavas szintetikus beszéd alkalmazásával közvetlenebb bizonyítékot nyerjen erre a hipotézisre. Grassegger (2003) megállapításai és személyes tapasztalatok alapján a hipotézis igazolása nagyon valószínű.

A második hipotézist nem jelenthetjük ki minden beszélőre, mert a korábbi tanulmányok szerint nem minden bemondónál jelentkeztek markáns egyéni különbségek. Minden hallgatóra se általánosíthatunk, mert egyrészt az egyes hallgatók által a felismerésre használt ismertetőjegyek eltérőek lehetnek (van Lancker et al. 1985), másrészt az egyes hallgatók különbözőképpen érzékelik az egyes fonációs típusokat (Kreiman et al. 1992).

Percepciós kísérlet segítségével próbáltuk igazolni a hipotéziseket, formánszintézis-technológia alkalmazásával. A módszer hasonlít Allen és Miller (2004) kísérletéhez. Négy beszélő 4, egy-két szavas felvétele alapján elkészült azok szintetikus másolata. Ezekből a másolatokból két változat született: az egyik végén a hangszalagrezgés reguláris, a másik végén irreguláris, azaz glottalizált. Az eredeti bemondások végén két bemondó megbízhatóan glottalizált, kettő pedig ritkán, ezért a két mesterséges változat közül az egyik megfelelt a beszélő szokásos végső hangjellegzetességének, a másik nem. Ezek a párok lehetőséget adtak annak vizsgálatára, hogy a szokásos végződéstől (glottalizált vagy reguláris) való eltérés milyen hatással van a beszélő személy azonosítására.

A végső glottalizáció mellett az alapfrekvencia-görbét is módosítottuk (ennek módját később ismertetjük). Egy felvétel alapfrekvencia-görbéje alapján bizonyos pontossággal megállapítható többek között az F_0 átlaga, amely bizonyítottan a beszélő személy robusztus ismertetőjegye (irodalmi áttekintés: Pardo-Remez 2005). Így ha egy felvételre egy másik beszélő alapfrekvencia-görbét

ültetjük át, akkor – tapasztalataink szerint – a beszélő személye nehezen felismerhető.

Ezzel a drasztikus módosítással több célt is elértünk. Ha mindig ugyanaz a különbség a párok két tagja között (glottalizált – nem glottalizált), akkor a kísérleti személyek hamar rájöhetnek a hipotézisekre és a módszerekre, így a továbbiakban a beszélőfelismerési folyamattól független stratégiát követhetnek. Ezenfelül az egymástól csak az alapfrekvencia-görbében eltérő párok kontrollként szolgáltak: egyrészt igazolták a kísérleti módszer működését, másrészt összehasonlítási alapként szolgáltak. Az F_0 -átlag robusztusságából adódóan az alapfrekvencia-görbére számolt eredmény a glottalizációra kapott eredmény felső korlátja. Így azt feltételeztük, hogy a glottalizáció hatása az alapfrekvencia-görbe hatásánál jóval kisebb, de kimutatható.

Így mindkét szintetikus változathoz (glottalizált és nem glottalizált) további két változat született: egy az eredeti és egy a módosított alapfrekvencia-görbével.

A kísérleti személyek minden lehetséges változatpárt meghallgattak, és eldöntötték, hogy melyik hasonlít jobban a megadott személy hangjára. Az összes kísérleti személy jól ismerte mind a négy beszélőt. Ha a kísérleti személyek gyakran felismerik a beszélő szokásos végső hangminőségét, akkor az alátámasztja a hipotéziseket. Ha a válaszok véletlenszerűek, akkor a hipotézisek valószínűleg tévesek, vagy a kísérlet nem elég érzékeny. A kísérletet amerikai angol nyelvre végeztük, mert a fentebb hivatkozott cikkek többsége ezzel a nyelvvvel dolgozott, valamint a munka jelentős részét a szerző egy amerikai egyetem vendéghallgatójaként végezte el.

Anyag, módszer, kísérleti személyek

Egy saját készítésű korpusz 9 amerikai angol bemondója közül választottuk ki a kísérletben részt vevő 4 személyt. Minden adatközlőre kiszámítottuk a glottalizációs arányt (a végső pozíciók hány százalékában glottalizáltak). Egyes adatközlők szinte mindig glottalizáltak, mások ritkán, és megint mások néha igen, néha nem. A kísérlethez két gyakran és két ritkán glottalizáló személyt választottunk (1. táblázat) – ha igaz a második hipotézis, akkor egy ilyen éles ellentét esetén valószínűleg kimutatható. A két férfi (FR és FG) átlagos alapfrekvenciája jelentős eltérést mutatott. Ez a kísérlet eredményeit nem befolyásolja, mert a hipotézisek vizsgálata során a kísérleti személyek ugyanahhoz a beszélőhöz tartozó, kétféle módon szintetizált hanganyag közül választották ki a beszélőre jellemzőbbet (nem két beszélőt hasonlítottak össze).

A négy adatközlő négy-négy rövid (egy-két szavas) olvasott beszédét használtuk. Az összesen 16 felvételtől másolatszintézist készítettünk egy Klatt-típusú formánszintetizátorral (Klatt–Klatt 1990). A másolatszintetizálás olyan eljárás, amely során olyan szintézisparamétereket keresünk, amelyekkel az eredeti felvételekre minél jobban hasonlító hullámformát kapunk (1. *b* ábra). Az így előállított másolatokon mindig megjelenik az eredeti felvételnek megfelelő glottalizáció is.

1. táblázat: A bemondók adatai

Azonosító	Nem	Végső glottalizáció	Átlagos F_0
NR	nő	ritkán (20%)	232 Hz
FR	férfi	ritkán (9%)	136 Hz
NG	nő	gyakran (93%)	210 Hz
FG	férfi	gyakran (83%)	94 Hz

Minden egyes másolatból a végső glottalizáció és az alapfrekvenciagörbe módosításával több változatot állítottunk elő. Az utóbbi módosítás esetén a másik, azonos nemű beszélő stilizált alapfrekvencia-görbét használtuk a szintézishez. A görbét időben úgy vetemítettük, hogy a csúcsok ugyanazokra az időpontokra essenek, mint az eredeti görbe csúcsai. Így ez a szintetikus változat az alapfrekvencia szempontjából (pl. átlagos F_0) egy másik személyre hasonlított, de minden más akusztikai tulajdonság (pl. beszédtempó, formánsok) változatlan maradt.

A kétféle módosítás alkalmazásával mind a 16 eredeti felvételhez négy szintetikus változat készült (összesen 64 darab): 1. másolatszintézis: minden szempontból hasonlít az eredetire; 2. módosított glottalizáció; 3. módosított alapfrekvencia-görbe; 4. módosított glottalizáció és alapfrekvencia-görbe.

Kilenc kísérleti személy (5 nő és 4 férfi) vett részt a percepciós kísérletben. Mindegyikük vagy angol anyanyelvű volt (5 fő), vagy legalább három éve folyamatosan angol nyelvterületen élt (4 fő, anyanyelvük héber, japán, koreai és mandarin). Életkoruk 23–29 év volt, és nem ismerték a kísérlet hipotéziseit és módszereit. Feltételezhetően jól ismerték a négy adatközlő hangját, mert legalább egy éve ugyanabban a kutatócsoportban dolgoztak vagy évfolyamtársaik voltak.

A kísérletet egy Matlab 7.0-ban írt grafikus program vezérelte. Ez a program játszotta le a hangzó anyagot, jelenítette meg a kérdéseket, és rögzítette a válaszokat. A program mindig minden válaszlehetőséghez egy külön nyomógombot jelenített meg. A válaszokat utólag nem lehetett módosítani, és minden stimulus csak egyszer hangzott el (újrajátszásra nem volt lehetőség). A hanganyagokat a program minden kísérleti személynek külön, véletlenszerű sorrendbe állította.

A kísérlet négy részből állt. Az első ellenőrizte, hogy a hallgatók tényleg jól ismerik-e a beszélők hangját, a második és a harmadik a szintetikus hanganyag minőségét tesztelte, míg a negyedik rész a hipotéziseket vizsgálta. Az első részben csak az eredeti, természetes felvételeket használtuk, míg a többiben csak szintetikusakat.

1. rész. Olyan kísérleti személyekre volt szükségünk, akik jól ismerik a beszélők hangját. Annak érdekében, hogy ezt biztosítsuk, és esetleg kiszűrjük az alkalmatlan hallgatókat, a kísérletet egy beszélőfelismerési feladattal kezdtük. A feladat során egy egy-két szavas felvétel meghallgatása után egy listáról ki kellett választani a hallott beszélő nevét. A négy ismert beszélő mellett két olyan személy felvételeit is lejátszottuk, akiknek a hangját nem ismerték a kísérleti

személyek. Ennek megfelelően a nevek listáján is szerepelt két további tétel „más nő” és „más férfi” felirattal.

2. rész. A cél a formánsszintézissel előállított glottalizált beszéd percepciójának mérése volt: vajon hasonló érzetet vált-e ki a hallgatókból a természetes és a mesterséges glottalizáció. A feladat páronkénti összehasonlítás volt. A párok két tagja mindig ugyanattól a beszélőtől származó, ugyanazt a szöveget realizáló szintetikus hanganyag volt, de az egyik vége glottalizált volt, a másiké nem. A két beszédrészlet lejátszása után a hallgatónak válaszolnia kellett arra a kérdésre, hogy „Melyik felvétel érdekesebb hangzású?” A választ egy 6 pontból álló skálán adhatta meg, ahol a két szélső érték mellett a „határozottan az első” és a „határozottan a második” felirat jelent meg. Ennek a skálának az előnye, hogy a hallgató egyetlen kattintással megválaszolja az eldöntendő kérdést, és egy bizonyosságértéket is megad (3, 4: kicsi; 1, 6: nagy).

3. rész. Ebben a részben azt ellenőriztük, hogy felismerhető-e a beszélő személye a szintetikus hanganyag alapján. A feladat ugyanaz volt, mint az első részben, csak itt nem a természetes, hanem a szintetikus bemondásokat játszottuk le, és nem használtuk az ismeretlen személyek felvételeit.

4. rész. Ennek a résznek a célja a hipotézisek vizsgálata volt. A feladat páronkénti összehasonlítás volt, ahol a párok mindkét tagja ugyanak az eredeti felvételnek két szintetikus változata volt. Így a beszélő és a nyelvi tartalom ugyanaz, de vagy a glottalizáció, vagy az alaphfrekvencia-görbe, vagy mindkettő szempontjából eltér a két felvétel. A képernyőn megjelenő kérdés magyar fordításban: „Melyik XY hangja (vagy melyik áll közelebb hozzá)?” – ahol XY a beszélő személy neve. A válaszokat ismét a hatelemű skálán kellett megadni, mint a második részben. A felvételek 4 szintetikus változatából minden lehetséges párosítást felhasználtunk, mindkét sorrendben.

Az adatok elemzését az SPSS 14.0-ban elérhető leíró statisztikák, egymintás és kétmintás t -próba, valamint ANOVA segítségével végeztük. Minden adatot 3 tizedesjegy pontossággal jelenítettünk meg. Így ha a program számításai szerint $p = 0,000$, akkor azt $p < 0,0005$ -ként értelmezzük és közöljük.

Eredmények

1. rész. Beszélőfelismerés természetes hanganyag alapján

A válaszok 68%-a megegyezett a tényleges beszélő személyével. Amint a módszer leírásánál említettük, ebben a részben két, a hallgatók számára ismeretlen személy hangját is felhasználtuk. A helyes azonosítások aránya a négy ismert személyre 75% volt, a két ismeretlenre pedig jóval alacsonyabb, 53%. Az egyes kísérleti személyekre 54% és 79% közötti felismerési arányokat kaptunk az ismert és ismeretlen beszélőkre együttesen. Eszerint – ahogy az várható volt – nem minden személy ismerte azonos mértékben a beszélőket. Az alábbiakban belátjuk, hogy mindegyikük esetén elvethetjük azt a nullhipotézist, hogy számára ismeretlenek voltak a beszélők.

Ha egy kísérleti személy nem tudta volna felismerni a beszélőket, akkor választai véletlenszerűen oszlottak volna el a lehetséges válaszlehetőségek között. Így 1/3 körüli felismerési arányt érhetett volna el, mert minden kérdésnél 3 férfi vagy 3 női válogomb közül választott volna találgatással (feltételezhetjük, hogy a beszélő nemét meg tudta állapítani). Minden hallgatóra ennél szignifikánsan magasabb értéket kaptunk (egymintás t -próbával minden esetben $t \geq 2,912$ és $p \leq 0,005$), ami kizárja a találgatás lehetőségét. Eszerint minden hallgató számára ismerős volt a négy adatközlő hangja.

A felismerési arány a két ismert férfi esetén sokkal magasabb volt (89% és 100%), mint a két ismert nőnél (54% és 57%). Ennek oka valószínűleg a két férfihang közötti nagy átlagos F_0 -különbség (1. táblázat).

2. rész. Mesterséges glottalizáció és az érdeesség

Az adatok elemzésénél azt vizsgáltuk, hogy vajon a hallgatók a párnak azt a tagját tekintették-e érdekesebbnek, amelyik végét glottalizáltként szintetizáltuk. Az eredmények azt mutatták, hogy 66%-ban igen (standard hiba: 3%), ami szignifikánsan eltér az 50%-os véletlenszerű átlagtól ($t = 5,711$; $p < 0,0005$). Tehát a szintetikus glottalizáció felhasználható a természetes helyett.

3. rész. Beszélőfelismerés szintetikus hanganyag alapján

Szintetikus hanganyag esetén a beszélők felismerési aránya 68% volt, ami nem tér el szignifikánsan az első részben, természetes hanganyagokra számolt átlagtól (kétmintás t -próbával az összes beszélőre $t = 0,157$ és $p = 0,876$; az ismerősökre $t = 1,938$ és $p = 0,053$). Ezért a kísérletben alkalmazhatjuk a szintetikus hangzó anyagot a hipotézisek vizsgálatára.

4. rész. A glottalizáció hatása a beszélőfelismerésre

A felvételek négy szintetikus változatát összesen hatféleképpen párosítottuk össze (2. táblázat). A hanganyagok sorrendje a táblázatban nem lejátszási sorrendet jelent, mert a párokat mindkét sorrendben felhasználtunk. A 6-értékű válaszokat egy kétértékű döntésre (első vagy második hanganyag) és egy háromértékű bizonyosságra választottuk szét. A hipotézisek vizsgálatát a kétértékű döntéseken végeztük.

Egy választ akkor tekintettünk „találatnak”, ha a kísérleti személy a beszélőre jellemzőbb – azaz a 2. táblázat „egyik hanganyag” oszlopában szereplő – változatot választotta. Ezek az alapfrekvencia, a glottalizáció vagy mindkettő szempontjából az eredeti felvételre hasonlítanak.

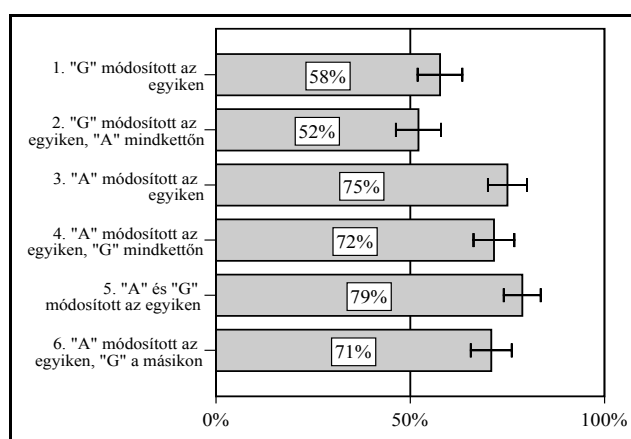
A hat pártípushoz tartozó találati arány (a válaszok hány százaléka találat) a 2. ábrán látható. A hipotézisek szempontjából az első pártípus az érdekes. Ha a végső glottalizáció léte vagy nem léte nem jelent érzékelhető különbséget, vagy ha ez a jelenség nem játszik szerepet a beszélő személy azonosításában, akkor véletlenszerű válaszokat kapnánk (találgatás). Ez azt jelenti, hogy a találati arány körülbelül 1/2 lenne, mert a kísérleti személyek kétértékű döntéseket hoztak. Ha azonban a hipotézisek igazak, akkor az arány ennél magasabb kell, hogy legyen. A mért találati arány 58%. Ez az arány nem sokkal magasabb 50%-nál, de az eltérés szignifikáns (egymintás t -próbával $t = 2,619$; $p = 0,009$). A kis elté-

rés azt jelzi, hogy a glottalizáció megváltoztatásának nincs olyan alapvető hatása a beszélőfelismerésre, mint az alapfrekvenciának. Az alacsony p érték viszont arra utal, hogy – bár kismértékű és nem általános érvényű – létezik ez a hatás, azaz a hipotézisek igazak.

2. táblázat: A hat pártípus

E = eredeti, M = módosított, a beszélőre nem jellemző.

	Egyik hanganyag		Másik hanganyag	
	Glottalizáció	Alapfrekvencia-görbe	Glottalizáció	Alapfrekvencia-görbe
1	E	E	M	E
2	E	M	M	M
3	E	E	E	M
4	M	E	M	M
5	E	E	M	M
6	M	E	E	M



2. ábra

Találati arány a pártípusokra

(G = glottalizáció, A = alapfrekvencia-görbe;
az ábrán a 95%-os konfidenciaintervallumokat is feltüntettük.)

A találat változóra ANOVA-t végeztünk, ahol a hallgató véletlen faktor, a beszélő, a bemondás szövege és a pártípus rögzített faktor volt. A faktorok értékeinek páronkénti összehasonlítását Tukey-féle post hoc próbákkal végeztük 95%-os szinten.

A pártípus faktor szignifikáns volt ($F = 18,327$; $p < 0,0005$), és a post hoc próbák két homogén csoportra bontották értékeit (amelyek tagjai között nem

mutatható ki szignifikáns eltérés, viszont a csoportok között igen): az egyikbe az 1. és 2. pártípus, a másikba a 3–6. pártípusok tartoztak. Ahogy az ANOVA és a 2. ábra is mutatja, az első pártípus találati aránya alacsonyabb a harmadik pártípusénál, ahol az eredeti és módosított alapfrekvencia-görbéjű változatot hasonlítottuk össze. Ez azt jelzi, hogy feltevésünknek megfelelően a beszélőfelismerési folyamatban az alapfrekvencia-görbe által hordozott információ nagyobb súlyú, mint a végső glottalizáció léte vagy nem léte.

Bár az ANOVA nem mutatott ki különbséget az első és a második pártípus találati aránya között, az első szignifikánsan eltér az 50%-tól (lásd fent), a második pedig nem (egymintás t -próba: $t = 0,706$; $p = 0,480$). Azaz ha mindkét felvételen módosítjuk az alapfrekvencia-görbét, akkor valószínűleg már egyik felvétel se hasonlít az eredeti beszélőre, így a válaszadók csak találgatni tudtak.

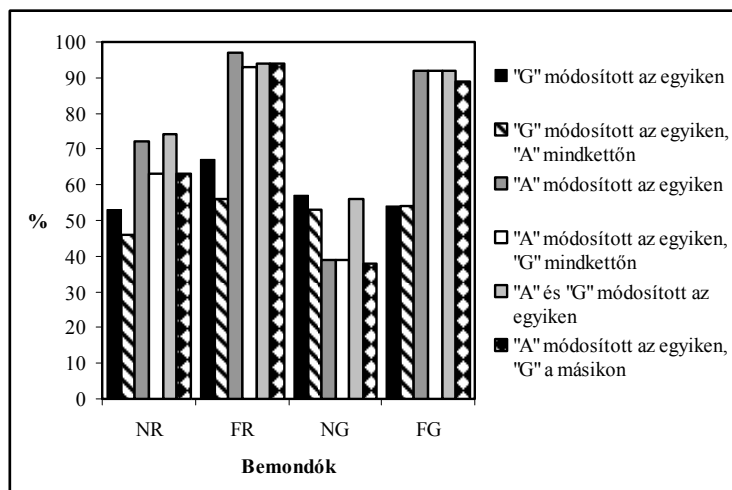
A negyedik pártípusnál is enyhén csökkent a felismerési arány a harmadikhoz képest (75-ről 72%-ra). A legmagasabb arányt akkor kaptuk, amikor a hallgatók az eredeti felvételhez hasonló szintetikus mintát hasonlították össze azzal, aminek mindkét vizsgált tulajdonságát módosítottuk (ötödik pártípus, 79%). A 3–6. pártípusokból álló homogén csoportban akkor kaptuk a legalacsonyabb találati arányt, amikor az egyik felvétel a végső glottalizáció, a másik az alapfrekvencia-menet szempontjából hasonlított az eredetire (hatodik pártípus, 71%). Ezek a kis, nem szignifikáns különbségek a glottalizáció hatásával magyarázhatók.

A pártípuson kívül a beszélő faktor volt szignifikáns ($F = 19,730$; $p < 0,0005$). A két férfi beszélő találati aránya (79% és 84%) magasabb volt, mint a két nő (47% és 62%), hasonlóan a kísérlet harmadik részében tapasztaltakhoz. A jelenleg háttérben valószínűleg az eltérő szintézisminőség áll.

A kísérleti személy faktor különböző szintjeihez tartozó találati arányok között nem volt szignifikáns különbség ($F = 1,150$; $p = 0,372$). Eszerint nem tudunk a hallgatók egyéni jellegzetességeivel vagy eltérő anyanyelvével összefüggésbe hozható különbséget kimutatni az eredményekben.

Az egymásra hatások közül kettő volt szignifikáns: a hallgató-beszélő ($F = 3,576$; $p < 0,0005$) és a pártípus-beszélő ($F = 6,156$; $p < 0,0005$). Az előbbi szerint a beszélőfelismerési folyamat függ a beszélőtől és a hallgatótól is, azaz indokoltan korlátoztuk a második hipotézist egyes hallgatókra és egyes beszélőkre. A pártípus-beszélő párosítások közötti különbségek pedig arra utalnak, hogy különböző beszélők esetén más ismertetőjegyek lehetnek fontosak. A 3. ábra alapján NG-t a glottalizáció, míg FG-t az alapfrekvencia-menet alapján volt könnyebb azonosítani.

A hallgatók az esetek 76%-ában ugyanarra a kérdésre fordított lejátszási sorrendben legfeljebb egy skálaértékkel eltérő választ adtak. Ugyanez a megbízhatósági mérték az első pártípusra 83% volt. Ez azt jelenti, hogy a kísérleti személyek következetesen válaszoltak a kérdésekre, és ha újra elvégeznék a kísérletet, valószínűleg az egyes pártípusokra hasonló eredményeket kapnánk.



3. ábra

Találati arány az egyes bemondókra és pártípusokra

Összefoglalás és kitekintés

Percepció kísérlettel vizsgáltuk meg, hogy vajon a végső glottalizáció érzékelhető-e a hallgatók számára, és hogy szerepet játszik-e a beszélő személy felismerésében. Mivel a beszélőfelismerési folyamatban felhasznált ismertetőjegyek beszélőnként és hallgatónként eltérőek (van Lancker et al. 1985), nem volt célunk e szerep általános érvényű bizonyítása: hipotézisünk szerint ez csak egyes beszélőkre és egyes hallgatókra korlátozódik. Az eredmények alapján ezekkel a megkötésekkel a fenti kérdésekre igen a válasz, de a következtetések hitelességéhez további kísérletek elvégzése szükséges. Az eredmények biztatóak, mert még egy olyan erőteljes beavatkozás mellett, mint az alaphérfrekvencia-görbe átültetése, is kimutatható volt a végső glottalizáció hatása.

Feltételezéseinknek megfelelően a beszélőfelismerési folyamatban az alaphérfrekvencia-menetnél általában kisebb súlyú, de szignifikáns ismertetőjegy a glottalizáció. A négy beszélő közül a gyakran glottalizáló nő esetén azonban a végső glottalizáció több információt hordozott a beszélő személyéről, mint az alaphérfrekvencia-menet.

Az eredményeket szintetikus beszédre kaptuk, így bizonyos szintű fenntartásokkal kell kezelnünk azokat. Bár a kísérlet második és harmadik része alapján ezek minősége (a beszélőfelismerés szempontjából) megközelítette a természetes beszédet, nem biztosított, hogy ez a felvételek minden fontos jellegzetességét utánozni tudta.

Az eredmények egyrészt hozzájárulnak a beszéd paralingvisztikai tulajdonságainak megértéséhez, másrészt alkalmazhatók a beszédtechnológiában.

Irodalom

- Allen, J. Sean – Miller, Joanne L. 2004. Listener sensitivity to individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America* 115. 3171–3183.
- Gordos Géza – Takács György 1983. *Digitális beszédfeldolgozás*. Műszaki Könyvkiadó, Budapest.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Grassegger, Hans 2003. A hangminőség szubjektív és objektív megítélése. *Beszédkutatás* 2003. 44–56.
- Henton, Caroline G. – Bladon, Anthony 1987. Creak as a sociophonetic marker. In Hyman, Larry M. – Li, Charles N. (eds.): *Language, speech and mind: Studies in honour of Victoria A. Fromkin*. Routledge, London, 3–29.
- Huber, Dieter 1992. Perception of aperiodic speech signals. *Proceedings of International Conference on Spoken Language Processing (ICSLP) '92*. Alberta, Canada, 503–506.
- Klatt, Dennis H. – Klatt, Laura C. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America* 87. 820–857.
- Kreiman, Jody – Gerratt, Bruce R. – Precoda, Kristin – Berke, Gerald S. 1992. Individual differences in voice quality perception. *Journal of Speech, Language, and Hearing Research* 35. 512–520.
- Laver, John 1980. *The phonetic description of voice quality*. Cambridge University Press, Cambridge.
- Markó Alexandra 2005. *A spontán beszéd néhány szupraszegmentális jellegzetessége*. PhD-értekezés. ELTE, Budapest.
- Pardo, Jennifer S. – Remez, Robert E. 2005. *The perception of speech*. Technical Report. Speech Perception Laboratory, Barnard College, New York.
- Redi, Laura – Shattuck-Hufnagel, Stefanie 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29. 407–429.
- Slifka, Janet 2000. *Respiratory constraints on speech production at prosodic boundaries*. PhD-disszertáció. MIT, Massachusetts.
- Van Lancker, Diana – Kreiman, Jody – Wickens, Thomas D. 1985. Familiar voice recognition: patterns and parameters. Part II. *Journal of Phonetics* 13. 39–52.

A szerző ezúton fejezi ki köszönetét Kenneth N. Stevens és Stefanie Shattuck-Hufnagel (MIT, USA) szakmai útmutatásáért és Maricza Istvánnak az adatok elemzéséhez nyújtott segítségért. A munka anyagi háttérét a Fulbright-ösztöndíj biztosította.