

# A zöngképzés változásai egy érzelemindukciós számítógépes játék során

BARTÓK MÁRTON

ELTE BTK Fonetikai Tanszék, Budapest

MTA-ELTE „Lendület” Lingvális Artikuláció Kutatócsoport

*bartokmarton@gmail.com*

## Bevezetés

Az érzelmi tartalom meghatározó tényező az emberek közötti interakciókban, és így a beszélt nyelvi interakciókban is. Beszédünket egyrészt befolyásolhatják különböző történések, ingerek által kiváltott spontán érzelmi reakciók, másrészt az érzelmi állapotok tudatos kifejezését gyakran használjuk céljaink elérésére társas interakciókban. Az itt bemutatott kísérlet célja spontán érzelmi reakciók beszédbeli megjelenésének vizsgálata, tehát a beszédtevékenység sokrétű érzelmi vetületén belül a különböző ingerek által kiváltott spontán fiziológiai-biológiai változások megragadására fókuszálunk. Ennek érdekében egy számítógépes játék során manipuláltuk különböző, feltételezhetően érzelmi hatást kiváltó ingerek megjelenését. A játék során megjelenő érzelmi reakciókról, szűkebben véve pedig azok beszédbeli vetületéről egy olyan érzelemelméleti modell segítségével alkottunk hipotéziseket, mely az egyént érő ingerek kognitív kiértékelésének lehetséges kimenetelei alapján jósol különböző idegrendszeri-fiziológiai változásokat (mely változások feltehetően befolyásolják a beszédképzés folyamatát). A beszéd komplex folyamatán belül a zöngképzés érzelemfüggő változatosságát elemeztük, a zöngé változatosságát mind diszkrét fonációtípusok elkülönítésével, mind akusztikai mérőszámokkal jellemezve. A következőkben elsőként az alkalmazott elméleti keretet, majd az abból levezethető, a zöngé érzelemfüggő változásra vonatkozó predikcióit foglaljuk össze.

Bartók Márton 2019. A zöngképzés változásai egy érzelemindukciós számítógépes játék során. *Beszédkutató* 2019. 6–29.

DOI-azonosító: 10.15775/Beszkut.2019.6-29

## A komponens-folyamat modell

Az értékelési (appraisal) modellek (pl. Lazarus 1968; Scherer 1987; Ortony et al. 1990; Scherer 2001) alapfeltevése szerint az érzelmek az őket megelőző szituáció (nem feltétlenül tudatos) értékelése alapján jönnek létre, és az értékelési folyamat egyes lépései határozzák meg a különböző válaszdomének reakció-mintázatait. A dolgozatban ezek közül Scherer (Scherer 1987; Scherer 2001) komponens-folyamat modelljét (Component Process Model) használok elméleti keretként, mivel ez az értékelési modell predikciókat fogalmaz meg az értékelési folyamat különböző lépései (komponensei) által meghatározott beszédbeli változásokra nézve is (Johnstone et al. 2001). A komponens-folyamat modell szerint az ingerek feldolgozása négy fő értékelési szempont köré csoportosul (Scherer 2001), ezek az alábbiak:

**1. Relevancia:** Mennyire releváns az inger az egyén számára? Van-e közvetlen hatása az egyénre vagy a szociális környezetére?

**2. Implikációk:** Milyen implikációkat vagy következményeket rejt magában az inger? Hogyan befolyásolhatják ezek a következmények az egyén aktuális állapotát, rövid és hosszú távú céljait?

**3. Megbirkózási potenciál:** Mennyire tud az egyén megbirkózni ezekkel a következményekkel, vagy milyen mértékben tudja módosítani azokat?

**4. Normatív szignifikancia:** Milyen kapcsolatban van az inger az egyén önképével, szociális normáival és értékeivel?

Ezek az értékelési szempontokon belül Scherer (2001) számos értékelési lépést (Stimulus Evaluation Check) különböztet meg. A továbbiakban ezek közül az értékelési lépések közül csak két, az implikációs értékelési folyamatok alá tartozó komponens mutatunk be részletesebben, azokat, amelyek különböző lehetséges végkimeneteleinek hatását a zöngképésre a kísérlet során vizsgáltuk. Az egyik ilyen komponens az **elvárásoktól való eltérés ellenőrzése** (discrepancy from expectation check), melynek során az egyén azt állapítja meg, hogy a megtapasztalt esemény megegyezik-e az általa alkotott előzetes, az aktuális cselekvéssorozatba illeszkedő várakozásaival, az adott időpillanatra alkotott predikcióival (Scherer 2001: 96). Ezek a várakozások az adott időpillanatig vezető eseménysor vagy az egyén önértékelése tükrében nem feltétlenül pozitívak, Scherer (2001: 96) példájával élve egy diák sikertelen vizsgája után feltehetően várakozásaitól eltérő eseményként értékelheti, hogyha a vizsgaeredményről értesült apja megajándékozza őt.

A másik értékelési lépés a **cél eléréséhez való hozzájárulás ellenőrzése** (*goal/need conduciveness check*). A megtapasztalt események hozzájárulhatnak az egyén különböző céljainak, szükségleteinek teljesüléséhez, vagy egyenesen be is teljesíthetik azokat, de gátolhatják is azok elérését,

hátráltatva az adott célra irányuló tevékenységsorozatot. A fenti példával élve a sikertelen vizsga hátráltatja a diáknak azt a (feltételezett) célját, hogy elvégezze azt a képzést, melyben részt vesz, míg egy sikeres vizsgaeredmény elősegíti a képzés teljesítését. Az elmélet szerint a két ellenőrzési lépés lehetséges kimenetelei ortogonális viszonyban vannak egymással, tehát tetszőlegesen kombinálódhatnak. Az itt bemutatott kutatásban a két lépés lehetséges végkimenetelei által alkotott négy értékelési kombináció mindegyikét vizsgáljuk.

Mielőtt az értékelési folyamatok különböző lehetséges végkimeneteleiből levezethető beszédbeli változásokra vonatkozó predikciókat áttekintenénk, az alábbiakban röviden összefoglaljuk az általunk a beszédképzés folyamatán belül vizsgált részfolyamat, a zöngképzés (fonáció) jellemzőit, és a fonáció változatosságának általunk vizsgált dimenzióit. Az érzelmek vokális kifejezéséért Murray–Arnott (1993) szakirodalmi áttekintése alapján a beszédjel három komponense felelős: a fonáció típusát is magába foglaló hangminőség (voice quality), az intonáció, valamint a beszéd temporális vetülete. A jelen esetben ezek közül a gégeműködés, egészen pontosan a fonáció szerepét vizsgáljuk az érzelemkifejezésben, ezen belül is a fonáció különböző típusainak megjelenését, valamint a fonációtípusokat megragadó akusztikai paraméterek és az alapfrekvencia értékeit elemezzük a különböző érzelmi állapotok függvényében.

A komplex beszédjel forrásának tekinthető zöngé képzése során a hangajkak kváziperiodikus rezgést végeznek. Ez a mozgás egyrészt jellemezhető a zöngeperiódusok gyakoriságával, azaz a zöngé alaphangmagasságával ( $f_0$ ), mely paraméter nagyrészt meghatározza a beszéd észlelt hangmagasságát. Az alaphangmagasság értéke a hangajkak adott hosszra jutó tömegétől függ, így azok hosszanti feszítettségének változtatásával képesek vagyunk az  $f_0$  modulálására: a feszítettebb hangajkak magasabb  $f_0$ -hoz vezetnek, kisebb feszítettséggel pedig alacsonyabb alaphangmagasságat érhetünk el. A gégegében található különböző izmok eltérő feszítettségé következtében nem csak a periódusok gyakorisága változhat, hanem a fonáció típusa is. A fonációtípusok egyfelől megközelíthetők diszkrét kategóriákként, olyan diszkrét kategóriák elkülönítésével mint például modális, leheletes vagy irreguláris fonáció. Az egyes kategóriák közötti eltérések ezen túlmenően mind artikulációs, mind akusztikai mérőszámokkal is megragadhatók. Az alábbiakban az általunk vizsgált három fonációtípus artikulációs és akusztikai sajátosságait ismertetjük, főként Laver (1980) alapján.

A zöngképzés prototipikus módja, a modális zöngé esetén a hangajkak rezgése hatékony, periodikus, és a hangajkak teljes mértékű záródása következtében nem jelenik meg hallható frikció a zöngében (Laver 1980: 111). A hangajkak záródása jóval kevesebb időt vesz igénybe, mint nyitódásuk, hiszen míg a kiáramló levegő csak lassan képes szétfeszíteni a

glottiszt, a Bernoulli-törvényből levezethető szívó hatás gyorsan zárja össze a hangjakakat. A modális zöngé prototipikus jellemzőitől eltérő fonáció-típusokat összefoglaló jelleggel nemmodális fonációtípusokként nevezzük meg, melyek közül az alábbiakban a leheletes, valamint az irreguláris fonáció jellegzetességeit mutatjuk be.

A leheletes zöngé esetében a glottisz záródási és nyitódási ideje hasonló, tehát a glottális áramlási hullámformák szimmetrikusabbak (vö. Huffman 1987), mint modális zöngé esetén. A szimmetrikusabb, a szinuszhullámhoz jobban hasonlító glottális hullámforma az akusztikai jelre is hatással van: az első harmonikus (H1) amplitúdójának relatív erősítéséhez vezethet (Klatt–Klatt 1990). Ezt a relatív erősödést a leggyakrabban a második harmonikus amplitúdójához viszonyítva (H1–H2), decibelben kifejezve számszerűsítik a fonációtípusok vizsgálatában. A H1–H2 mérőszám tehát annál magasabb értéket mutat, minél nagyobb az eltérés – a hangjakak lassabb záródása és ennek betudhatóan a glottális hullámforma szinuszosabb alakjának következtében – az első két harmonikus amplitúdója között a H1 javára. Emellett a leheletes zöngé jellemzően a modálisnál alacsonyabb alapfrekvenciával valósul meg (Fairbanks 1960: 179).

Az irreguláris zöngé képzésekor Laver (1980: 122–126) szerint a hangjakak mediális kompressziója nagy-, míg longitudinális feszítettségük kismértékű. A nagymértékű mediális kompresszió következménye, hogy a tüdőből kiáramló levegő csak ritkábban, és nem feltétlenül szabályos időközönként tudja szétfeszíteni az egymásnak feszülő hangjakakat, tehát a zöngé alapfrekvenciája a modális zöngéjénél alacsonyabb lesz, az egymást követő glottális ciklusok periódusidejében pedig nagymértékű variabilitás figyelhető meg. Az alacsony alapfrekvenciával összefüggésben Wendahl ésmunkatársai (1963: 254) az irreguláris fonáció legfontosabb akusztikai korrelátumaként a periódusokon belüli csillapítást (damping) nevezik meg. A prototipikus irreguláris zöngé tehát hosszú periódusidővel és nagymértékű csillapítással jellemezhető, valamint az egymást követő periódusok hossza variabilis.

### **A stimulus-ellenőrzés kimeneteleinek hatása a fonációra**

Az itt bemutatott kísérletben tehát azt elemezzük, hogy az egyént érő ingerek feldolgozásának a Scherer (2001) komponens-folyamat modellje szerint meghatározott két adott lépése milyen reakciómintázatok megjelenéséhez vezet a beszédben, szűkebben véve pedig a zöngé képzésében. A két vizsgált stimulus-ellenőrzési lépés az elvárásoktól való eltérés, illetve a cél eléréséhez való hozzájárulás ellenőrzése. Az értékelési folyamat lehetséges kimeneteleit az 1. táblázat foglalja össze. A következőkben ismertetjük az egyes kimenetek esetén feltételezetten bekövetkező, a beszédképzésben releváns

pszichofiziológiai változásokat, illetve az ezek következtében potenciálisan megjelenő akusztikai változásokat a beszédjelben. Azt is áttekintjük, hogy a különböző végkimenetek lehetséges kombinációi milyen érzelmeknek feleltethetők meg (főként Johnstone et al. 2001 alapján), amennyiben diszkrét érzelemelméleti nézőpontot alkalmazunk.

1. táblázat: A vizsgált stimulus-ellenőrzési lépések lehetséges kimenetelei, illetve az azokkal kapcsolatba hozható diszkrét érzelmi kategóriák (ezeket szögletes zárójelek között adjuk meg)

|                                  | <b>célhoz való hozzájárulás</b>        |   |
|----------------------------------|--|---|
| <b>elvárásoktól való eltérés</b> | nem várt, hozzájáruló<br>[meglepődés]  | nem várt, hátráltató (mésékelt megbirkózási potenciál)<br>[meglepődés, stressz] |
|                                  | várt, hozzájáruló<br>[megelégedettség] | várt, hátráltató (mésékelt megbirkózási)<br>[stressz]                           |

Pszichofiziológiai szempontból az új, váratlan ingerek észlelése úgynevezett orientációs reakciót (orienting response) vált ki (Sokolov–Cacioppo 1997), melynek főbb pszichofiziológiai jellemzői az izomtónus megnövekedése, valamint a légzés pillanatnyi szünetelése, melyet gyakran fokozott szaporaságú, illetve mélységű légzés követ (Lynn 1966). A megnövekedett izomtónus a hangszalagok nagyobb longitudinális feszítettségéhez, így a fonáció alaphékvenciájának növekedéséhez vezethet (Johnstone et al. 2001: 274–275). A váratlan ingerekre érkező pszichofiziológiai válaszmintázatok között Johnstone és munkatársai (2001) szerint nincs nagy különbség aszerint, hogy az inger hátráltató vagy hozzájáruló hatással van az egyén céljára.

Az egyént hátráltató események kiértékelését tovább árnyalja az értékelési folyamat következő lépése, a megbirkózási potenciál: amennyiben az egyén a célját gátló történést tapasztal, felméri, hogy milyen mértékben áll módjában megbirkózni azzal. Amennyiben a megbirkózás valószínűsége magas, az

egyénnek kismértékű erő kifejtésre van szüksége, ez mérsékelt szimpatikus idegrendszeri aktivitáshoz, és relaxált izomtónushoz vezet. A megbirkózási potenciál kontinuumának másik végén, tehát amennyiben az egyén úgy ítéli meg, a kialakult problémát nem áll módjában elhárítani, a rezignáció állapota következhet be, amely állapot még alacsonyabb szimpatikus idegrendszeri aktivitással, valamint alacsony izomfeszítettséggel jellemezhető (Johnstone et al. 2001: 276). A jelen kísérletben olyan hátráltató események hatását vizsgálom, melyek a fentebb említett két szélsőséges értékelés (magas, illetve minimális megbirkózási potenciál) között húzódó kontinuum közepén elhelyezhető mérsékelt vagy bizonytalan megbirkózási potenciállal jellemezhetők. Ebben az esetben a vegetatív idegrendszer szimpatikus ága aktiválódik, ezzel energiát mozgósítva a megbirkózáshoz szükséges tevékenységek végrehajtására. Ennek következtében a légzés mélysége és szaporasága növekszik (Boiten et al. 1994), valamint feszítettebb lesz az izomtónus (Svebak 1983). Ennek megfelelően az egyén céljait gátló, mérsékelt megbirkózási potenciállal jellemezhető események esetén a beszédet várhatóan magasabb alapfrekvencia, nagyobb intenzitás és gyorsabb artikulációs tempó jellemzi (Johnstone et al. 2001: 277). Johnstone (2001) egy számítógépes játék során manipulálta a kísérleti személyekben a célhoz való hozzájárulás és a megbirkózási potenciál értékelési lépését, és a játékosok beszédprodukciónak audió és elektroglossográfikus jeleken elemezte. A fentiekkel összhangban azt találta, hogy az alapfrekvencia magasabb, a hangjak nyitott állapotának időaránya (OQ) pedig alacsonyabb az egyént hátráltató események hatására, de csak alacsonyabb megbirkózási potenciál (azaz megoldhatónak tűnő probléma) esetén.

A várt, a célhoz hozzájáruló ingerek jellemzően relaxációhoz vezetnek. Amennyiben valamely, az egyén céljához vezető feltétel teljesült, a közelmúltban mozgósított energiaforrások pótlására a vegetatív idegrendszer trofotropikus (energiapótlásra, pihenésre irányuló) aktivitása lesz jellemző (Scherer 2001). A trofotropikus aktivitás a vázizmok kiegyensúlyozott tónusához, valamint relaxált légzéshez vezet. A mérsékelt izomfeszítettség következtében az alapfrekvencia alacsonyabb, míg a relaxáltabb légzés következtében a beszédjel intenzitása szintén alacsonyabb lesz (Johnstone et al. 2001: 275).

Mivel az érzelmeknek a beszédre kifejtett hatását vizsgáló kutatások zöme nem értékelési modelleket alkalmazott elméleti keretként, hanem az érzelmeket diszkrét, egymástól elkülönített kategóriákként kezelte, ezen a ponton célszerű áttekinteni, hogy az általunk vizsgált különböző stimulusértékelési lépések lehetséges végkimeneteleinek kombinációi milyen diszkrét érzelmi kategóriákkal hozhatók összefüggésbe, és eszerint tárgyalni a már ismert akusztikai jellemzőket. Az érzelmek diszkrét kategorizációjának egy lehetséges módját adja Russell (1980) körmodellje, amely az aktiváció és a

valencia dimenziója mentén helyezi el a különböző érzelmeket. A modellben az aktiváció az adott érzelmi állapotban jellemző agyi aktiváció (arousal) szintjére, míg a valencia az adott érzelem negatív vagy pozitív töltetére és ennek mértékére utal. A körmodell origója a semleges valenciájú, közepes aktivációjú érzelmi állapot, az érzelmekek pedig egy origó központú kör pontjaiként helyezkednek el. A különböző értékelési kimeneteket az alábbiakban ezekkel az érzelmi kategóriákkal feleltetjük meg Johnstone és munkatársai (2001) alapján, illetve amennyiben a szakirodalomban nem szerepel ilyen párhuzam, logikai alapon.

A váratlannak értékelt helyzetek által kiváltott érzelmi állapot logikai alapon megfeleltethető a meglepődés diszkrét érzelmi kategóriájának. Ez az érzelem Russell (1980) körmodelljében magas aktivációval és semleges valenciával jellemezhető, éppen annak a következtében, hogy a váratlan esemény pozitív és negatív irányban egyaránt befolyásolhatja az egyén céljainak elérését, tehát az érzelem valenciája nem határozható meg kontextustól függetlenül. Akusztikai jellemzőit tekintve, eljátszott érzelmi állapotokban produkált beszéd alapján a magas aktivációjú érzelmekek általánosságban is jellemző a magas alapfrekvencia (Scherer 1986), a meglepődés pedig más érzelmekekhez képest is kiugró  $f_0$ -értékeket mutat magyar beszélők produkációjában (Bartók 2018). A magas alapfrekvencia mellett a meglepődés kifejezését Laukkanen és munkatársai (1996) inverz szűrővel végzett vizsgálatai alapján a hangajkak érintkezésének a glottális periódusra vetített magas időtartamaránya jellemzi.

A várt, az egyén céljaihoz hozzájáruló események feldolgozása a megelégedettség diszkrét érzelmi kategóriájával hozható kapcsolatba (Johnstone et al. 2001: 275). A megelégedettséget a russelli körmodellben a semlegesnél alacsonyabb aktiváció, valamint pozitív valencia jellemzi. Az alacsony aktivációjú érzelmekek jellemző alacsonyabb  $f_0$  mellett a megelégedettség kifejezése esetén magyar beszélőknél eltérő mintázat figyelhető meg a nemek szerint: női beszélőknél gyakori az irreguláris fonáció, férfiaknál pedig aleheletes zöngé (Bartók 2018). A trofotropikus aktivitásra jellemző relaxáció tehát kétféleképpen is megjelenhet a megelégedettség tudatos kifejezésében: nők esetén a longitudinális feszítettség alacsony, férfiaknál pedig a mediális kompresszió kisebb mértékű. Az irreguláris zöngé megjelenése összefüggésbe hozható továbbá a fent említett fiziológiai alapokon (ti. a longitudinális feszítettség alacsony) nyugvó paranyelvi funkciójával, hogy ez a zöngemínőség gyakran fejez ki inaktivitással, passzivitással összefüggésbe hozható érzelmekek (pl. unalom, szomorúság) (Schuller–Batliner 2014: 89).

A hátráltató, mérsékelt megbirkózási potenciálúnak értékelt események által kiváltott érzelmi állapothoz Johnstone és munkatársai (2001: 277) a stresszt tartják hasonló állapotnak. A stressz állapota két szempontból sem

illeszthető be az eddig tárgyalt diszkrét érzelmi állapotok sorába. Egyrészt a diszkrét érzelemelméleti modellek sok esetben nem tartalmazzák, mint érzelmi állapotot (ahogyan pl. Russellé sem). Másrészt a stressz esetén gyakoribb az indukált stresszes állapotban produkált beszéd jellegzeteségeinek vizsgálata, mint az eljátszott stressz beszédbeli jellemzőinek elemzése. A stressz hatása alatt produkált beszédre a magas alapfrekvencia és a nagy intenzitás jellemző (Hansen–Patil 2007; Sondhi et al. 2015).

A különböző érzelmi állapotok zöngképzésre kifejtett hatását nyelvi tényezők is befolyásolhatják: (Seppi et al. 2010) azt találták, hogy az automatikus érzelemfelismerésben jobb eredmények érhetők el a hangsúlyos szótagok alapján, mint a hangsúlytalan szótagok használatával. Ugyanakkor a hangsúlyos szótagok sokkal rezisztensebbek például a koartikulációs (Fowler 1981) hatásokkal szemben, ami felveti a kérdést, hogy ezzel ellentétben a paralingvisztikai hatások valóban az első szótagon jelennek-e meg inkább. Emellett a különböző változásokat az is befolyásolhatja, hogy az érzelmi hatást követő beszédtevékenységben a szótagok időben eltérő távolságra helyezkednek el az érzelmi hatást kiváltó ingertől, tehát az első, az inger megtapasztalását rövid időkülönbséggel követő szótagon más jellegű lehet az érzelmi reakció, mint az ezt követő szótagokon, melyek produkciójáig több idő telik el. A hangsúly tehát feltételezhetően befolyásolhatja az érzelmi hatások megjelenését, azonban az előbbieken alapján nem egyértelmű, hogy pontosan milyen jellegű hatás várható.

## **Célok és hipotézisek**

A bemutatásra kerülő kísérletben tehát olyan érzelmi állapotok hatását vizsgáljuk, amelyek a megtapasztalt ingerek valószínűsíthető értékelése szempontjából két dimenzió mentén különböznek: elemzem a várt, illetve váratlan ingerek hatásának eltéréseit, illetve az egyén céljaihoz hozzájáruló vagy azt hátráltató események hatása közötti különbségeket. Az értékelési érzelemelméleti modellek predikcióit, valamint a diszkrét érzelemelméleti keretben született paralingvisztikai vizsgálatok eredményeit áttekintve a fonáció szisztematikus variabilitására két szempontból fogalmazunk meg hipotéziseket. Egyrészt olyan feltételezéseket alkotunk, amelyek a különböző fonációtípusok gyakoriságában tapasztalható eltérésekre vonatkoznak a független változók mentén. Másrészt azt is feltételezzük, hogy a különböző kondíciókban megjelenő fonációtípus-gyakoriságokkal összefüggésben a modális zöngén belül akusztikai variabilitás lesz megfigyelhető, tehát amely kondícióban például a leheletes fonáció gyakoriságát feltételezzük, ott azt is várjuk, hogy a modális szakaszokon mért akusztikai és artikulációs paraméterek leheletesebb modális fonációt jeleznek, mint olyan kondíciókban, ahol a leheletes fonáció ritkább. Így az alábbi hipotéziseket



fogalmazzuk meg:

1. Feltételezzük, hogy a fonáció hasonló jellegzetességeket mutat az elvárásoktól való eltérés (váratlanság), az egyén céljainak hátráltatása (sikertelenség) esetén. Amennyiben a beszélő váratlan vagy sikertelen eseményt tapasztal, a gégeizmok nagyobb általános feszítettségének hatására a következő változások jelennek meg: a nagyobb longitudinális feszítettség hatására megnő a zöngé alapfrekvenciája, a nagymértékű mediális kompresszió hatására pedig a hangjakak érintkezésének aránya is magasabb, mint várt, illetve sikeres ingerek megtapasztalása esetén. A kismértékű mediális kompresszióval járó leheletes zöngé gyakorisága alacsony. Szintén ritka a kismértékű longitudinális feszítettséggel járó irreguláris fonáció. Mivel mind a leheletes, mind az irreguláris fonáció esetén azt feltételezzük, hogy ezek a nemmodális fonációtípusok ritkán jelennek meg, a H1–H2 értékei a modális zöngére jellemző (alacsony pozitív) értékeket mutatnak.
2. Szintén feltételezzük, hogy amennyiben a beszélő az elvárásaival konzisztens, a céljához hozzájáruló (tehát várt, sikeres) ingert tapasztal, az alapfrekvencia alacsonyabb, mint a többi kondícióban. Ebben a kondícióban a női beszélők esetében gyakrabban jelenik meg irreguláris fonáció, mint a többi kondícióban, ezzel összefüggésben a H1–H2 értékei alacsonyabbak. Férfi beszélők esetében ezzel szemben a leheletes zöngé gyakori, ami magasabb H1–H2 értékekkel jár, mint amit érzelmi hatás nélkül mérhetünk.
3. Végezetül azt is feltételezzük, hogy a fentebb említett hatások eltérnek hangsúlyos és hangsúlytalan szótagokban.

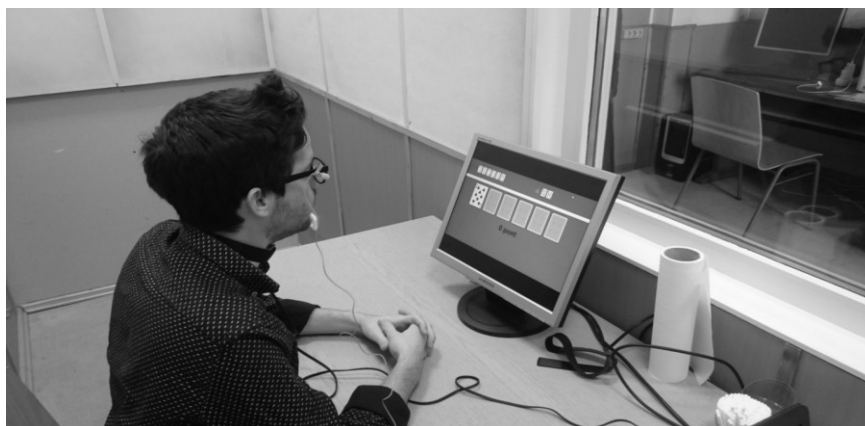
## Módszer

### A vizsgálatban alkalmazott érzelemindukciós technika

Az érzelmi reakciók kiváltására egy egyszerű, számítógépen játszható kártyajátékot fejlesztettünk Shahid és munkatársai (2007) módszerének átdolgozásával, JavaScript programozási nyelvben. A játék alapja az, hogy a játékosoknak kártyák sorozatát kell felfordítaniuk úgy, hogy sikeresen megtippelik, vajon a soron következő kártya magasabb vagy alacsonyabb értékű lesz-e, mint az azt megelőző. Az eredeti módszert kizárólag az érzelmi reakciók pozitív/negatív megoszlásának vizsgálatára használták a sikertelen és sikeres tippelések után rögzített reakciómintázatok alapján, azonban a játék alkalmas a két tárgyalt értékelési lépés különböző végkimeneteleinek elemzésére, amennyiben megvizsgáljuk azt is, hogy a sikeres vagy sikertelen végkimenetelt megelőző döntés milyen nehézségű volt. A módszer lényeges

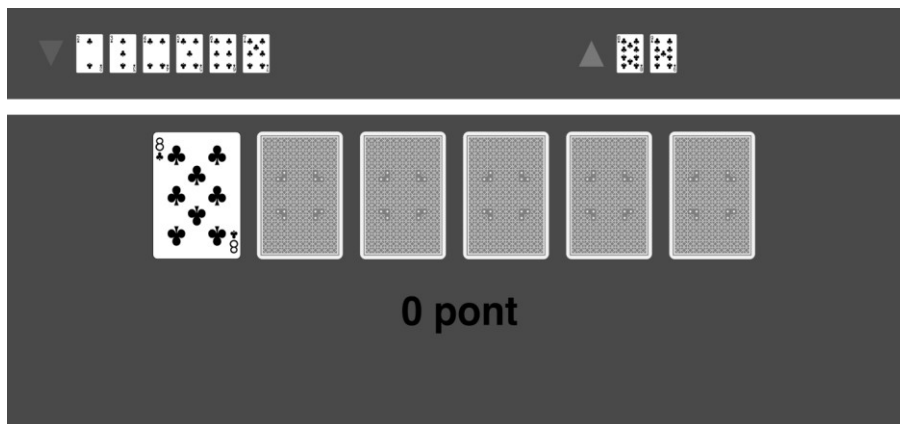
tényezője, hogy a játékot a kísérleti személyek hangutasításokkal vezérlik, ennél fogva tehát a beszédben megjelenő érzelmi reakciókat a játék dinamikájába illeszkedve rögzíthetjük (vö. Aharonson–Amir 2006). A továbbiakban részletesen bemutatjuk a kártyajáték menetét.

A kísérlet során az adatközlők egy monitor előtt ülnek, a csendesített szobában rajtuk kívül a kísérletvezető tartózkodik (vö. 1. ábra).



1. ábra: A kísérleti helyzet

A monitoron (vö. 2. ábra) alaphelyzetben hat kártya látható, az első felfordítva. A játékban a francia kártya 2 és 10 közötti lapjait használjuk.



2. ábra: Az érzelemindukációhoz használt számítógépes játék elrendezése

Az adatközlők azt az utasítást kapják, hogy tippeljék meg, hogy vajon a jobbra következő lefordított kártya értéke alacsonyabb vagy magasabb lesz-e, mint a felfordított kártya. Sikeres tipp esetén a jobbra soron következő kártyáról is el kell dönteni ugyanezt a kérdést az aktuálisan felfordított

kártyához képest, és így tovább. Mivel a kutatásban kizárólag az érzelmi hatást kívántuk vizsgálni, a potenciálisan az eredményeket befolyásoló kognitív terhelés minimalizálására a képernyő tetején megjelenítettük az összes lehetséges kisebb és nagyobb értékű lapot, amelyek számának egymáshoz viszonyított aránya tehát megegyezik a két lehetséges döntés sikerének valószínűségével. A döntés akkor a legnehezebb, amikor a soron következő kártya ugyanolyan valószínűséggel lehet nagyobb, illetve kisebb, mint az épp felfordított lap. A kísérleti személy az *alacsonyabb*, illetve *magasabb* hangutasításokkal irányítja a játékot.

Egy játszma addig tart, míg a kísérleti személy balról jobbra haladva mind az öt lefordított lapot sikeresen fordítja fel, vagy rossz döntést nem hoz. A játszma végén a kártyák feletti vízszintes vonal színe sikertelen játszma esetén piros, sikeres játszma esetén zöld színre vált, és a kísérleti személy a *haladjunk* szó segítségével tud továbblépni a következő játszmahoz. A *haladjunk* szó kiejtése előtt tehát a kísérleti személy vagy sikeres, vagy sikertelen játékeseményt tapasztalt meg. Az elemzés során ennek a szónak az első és második szótagi magánhangzóját vizsgáljuk a játszmák végkimenetelének, illetve a játszmák végét megelőző döntés nehézségének függvényében.

Minden adatközlő négy próbajátszma után harminckét játszmában vett részt, ez összesen körülbelül tíz percet vett igénybe. Az egyes játszmák között az adatközlők a képernyőről semleges érzelmi töltetű, rövid mondatokat olvastak fel annak érdekében, hogy az előző játszma sikeressége vagy sikertelensége ne befolyásolja a következő játszmában megjelenő érzelmi reakcióikat. A kísérleti személyek tetszőleges hosszúságú szünetet tarthattak a tizenhatodik játszma után. Az érzelmi involváltság elősegítése érdekében az adatközlők pontokat szerezhettek a játék során: minden helyes döntés 10 pontot, egy teljes játszma sikeres véghezvitele pedig 100 pontot ért. Az így kapott pontszám alapján minden tíz kísérleti személyből a legjobban teljesítőt tárgyjutalomban részesítettük, és erről a kísérletet megelőzően tájékoztattuk az adatközlőket az érzelmi bevonódás biztosítása érdekében.

Az egyes döntési helyzetekre jellemző megbirkózási potenciál mértékét ( $cp$ ) az alábbi képlettel operacionalizáltuk:

$$cp = 2 * |p_{magasabb} - 0.5|$$

A  $cp$  értéke tehát akkor a legalacsonyabb ( $cp = 0$ ), amikor a két végkimenetelnek ugyanakkora valószínűsége van, hiszen a tippelés ilyenkor teljesen véletlenszerűen történik, így a helyzettel való megbirkózás valószínűsége kisebb. Amennyiben viszont a végkimenetel biztos, tehát vagy csak alacsonyabb, vagy csak magasabb kártyák következhetnek, a  $cp$  értéke 1, hiszen ebben az esetben az adatközlő biztos lehet a döntés meghozatalakor. A harminckét játszma kártyaszekvenciáit úgy alakítottuk ki, hogy vettük a játékban használt 9 lap összes permutációját a képernyőn megjelenő hat

helyre. Ezt a 60480 permutációt sorba rendeztük a szekvenciákban megjelenő döntési helyzetek cp szummája alapján, és minden 1890-edik sorozatot kiválasztva jutottunk a kísérletben használt harminckét kártyasorhoz.

A kutatás során a két vizsgált értékelési lépést a következő két tényező mint független változó hatásaként képeztük le. Mivel a kísérleti személyek célja a játékszituációban feltehetőleg a minél magasabb pontszám elérése a tárgyjutalom megnyerése érdekében, a sikeres végkimenetelű, pontszerzéssel járó játszmák hozzájárulnak az egyén céljának eléréséhez, míg a sikertelen játszmák hátráltatják őt a cél elérésében. A várakozásoktól való eltérést annak függvényében határoztuk meg, hogy a játékos a játszma utolsó döntési helyzetének nehézsége alapján mennyire várhatta az aktuálisan bekövetkezett végkimenetelt (vö. 2. táblázat). Amennyiben a megelőző döntés nehéz volt, a játékos feltehetően a sikertelen végkimenetelt tartotta valószínűbbnek, és a sikeres végeredmény váratlanul érthette, könnyű döntési helyzet esetén azonban a játékos a sikeres eredményt várta, az esetleges sikertelenséget pedig váratlan eseményként értékelhette.

2. táblázat: Az elvárásoktól való eltérés változó származtatása

|                                     |                         | A játszma sikeressége |            |
|-------------------------------------|-------------------------|-----------------------|------------|
|                                     |                         | sikeres               | sikertelen |
| Az utolsó döntési helyzet nehézsége | Nehéz ( $cp \geq 0,5$ ) | nem várt              | várt       |
|                                     | Könnyű ( $cp < 0,5$ )   | várt                  | nem várt   |

Az így kapott két független változó, tehát az elvárásoktól való eltérés (várt vagy nem várt kimenetel), valamint a célhoz való hozzájárulás (sikeres vagy sikertelen játszma) a komponens-folyamat modellnek megfelelően egymással ortogonális viszonyban van (vö. 1. táblázat), ennek megfelelően a két változó interakcióját is vizsgáltuk. Az interakció fontosságára az értékelési modellek értékelési lépéseinek viszonyrendszerével kapcsolatban Johnstone (2001) hívja fel a figyelmet. Az eredmények bemutatása során a két vizsgált faktor megnevezésére a *váratlanság* (elvárásoktól való eltérés), valamint a *sikeresség* (célhoz való hozzájárulás) szavakat használjuk.

Fontos megjegyezni, hogy az adatközlők elméjében zajló értékelési folyamatokról csak hipotéziseket alkothatunk. Az általunk prezentált ingerek (játékesemények) feldolgozának folyamatát számos olyan tényező

befolyásolja, amelyet az értékelési folyamat modellezése során nem tudtunk figyelembe venni: az adatközlő adaptációja a játékmenethez, az egymást követő játzmák eredményeinek sorrendiségi hatása, és még számos, általunk hozzá nem férhető, az adatközlőre jellemző emocionális-kognitív sajátosság is hatással lehet az érzelmi állapot kialakulására.

## Vizsgálati anyag

A felvételeket az ELTE Fonetikai Tanszékén az MTA-ELTE Lendület Lingvális Artikuláció Kutatócsoport műszereivel készítettük, csendesített szobában. Az akusztikus jelet a Laryngograph Ltd. EGG-D200 típusú elektroglossográf készülékének egyik csatornáján rögzítettük. A felvételeken automatikus hangszintű annotálást végeztünk a BAS webszolgáltatás graféma-fonéma konverterének (G2P, Reichel 2012) és a MAUS rendszernek (Schiel 1999) segítségével. Az automatikusan jelölt hanghatárokat manuálisan javítottuk a Praat programban (Boersma–Weenink 2018). Ezen túlmenően annotáltuk a különböző fonációtípussal produkált szakaszokat a szavakon belül. Mind a fonációtípusok gyakoriságát, mind az egyes akusztikai-artikulációs mérőszámok jellemző értékeit a célszó első, szó- és mondathangsúlyt viselő (a továbbiakban hangsúlyos) és második (a továbbiakban hangsúlytalan) /v/ magánhangzóján vizsgáltuk. Nem elemeztük azokat a szavakat, melyek során a beszéd nevetéssel vegyesen valósult meg. A kísérlet végén továbbá minden kísérleti személlyel rögzítettük a célszót ötszöri felolvasásban, mely megvalósulásokat érzelmileg semleges kontrollként kezeltük. Mivel a kontrollanyag rögzítésére a kísérlet végén került sor, feltehetőleg ebben a kondícióban már jóval kisebb a kísérleti személyeknek a kísérleti helyzet szokatlanságából adódó esetleges diszkomfortérzete, illetve a játékhelyzet által keltett esetleges izgalom, így ez a minta közelít leginkább a habituális beszédprodukciónak. A játék során megjelenő különböző helyzetek esetén ezek a hatások (izgalom, diszkomfort) hozzávetőleg változatlanok, ami lehetővé teszi, hogy ezen kondíciók összevetésénél valóban a játékbeli manipulációk hatását érjük tetten, de emellett a kontrollkondíció segítségével arról is képet alkothatunk, hogy a játék során megfigyelt változások hogyan viszonyulnak a habituális közeli beszédprodukciónak.

A fonációtípusok címkézését az auditív információk mellett a beszédjel nyomás-idő függvénye és spektrogramja alapján, a bevezetésben ismertetett akusztikai jellegzetességeket figyelembe véve, a modális, leheletes és irreguláris fonáció megkülönböztetésével végeztük. A hanganyagon akusztikai méréseket is végeztünk, azonban ezek a mérések a következőkben ismertetett módszertani megfontolásokat figyelembe véve csak a modális fonációtípusú szakaszokon történtek. Ahogy azt Keating és munkatársai

(2015) összefoglalják, az irreguláris fonációtípusok bizonyos altípusai valamilyen mértékben nélkülözik a periodicitást. Ilyen altípus a szélsőséges irregularitást mutató aperiodikus irreguláris fonáció, mely esetén nem lehetséges a zöngperiódusok elkülönítése, de problémás a periodicitáson alapuló mérőszámok értelmezése még az úgynevezett diplofónia esetén is, amit a hangjakak különböző területeinek asszimmetrikus rezgéséből levezethető alperiódusok megjelenése jellemez (Aichinger 2015). Ezt figyelembe véve az akusztikai méréseket egységesen kizárólag a modális fonációtípusú szakaszokon végeztük, így a mért értékek a modális fonáción belüli variabilitásról adhatnak képet. Így az akusztikai mérések esetén minden, modális fonációt tartalmazó beszédhangot egyetlen típussal reprezentáltunk. A különböző fonációtípusok gyakoriságát ezzel szemben az összes elkülönített szakaszt figyelembe véve számszerűsítettük, mely beszédhangon belüli fonációtípus-váltás esetén beszédhangonként több adatpontot eredményezett.

### **A vizsgált akusztikai paraméterek**

Az elemzett modális szakaszok teljes időtartamán a következő akusztikai paraméterek átlagát mértük:  $f_0$ , H1–H2. Minden akusztikai mérést a Praat szoftver (Boersma–Weenink 2018) segítségével végeztünk, a következőkben említett beállításoktól eltekintve a szoftver standard beállításait használva. A mérések során az alaphfrekvencia határértékeit nőknél 100–350 Hz-re, férfiaknál 70–250 Hz-re állítottuk (Eager 2015 alapján). Az alaphfrekvencia értékét keresztkorrelációval mértük, a harmonikusok amplitúdóját az így kapott  $f_0$ -értéknél és annak kétszeresénél, illetve az azokhoz viszonyított  $\pm 10\%$ -os frekvenciatartományban megállapított maximális amplitúdóként határoztuk meg.

### **Adatközlők**

Összesen 36 adatközlővel készítettünk felvételt, azonban, amennyiben a kontrollkondícióban rögzített felvételek esetében az elemzett szótagok időtartamának legalább 50%-a nem modális fonációtípusú (2 felvétel) volt, az adatközlő fonációját patológiásnak tekintettük, és adatait kizártuk a további elemzésből. Így végül 34 kísérleti személy adatai elemeztük. Az elemzett minta a nemek szempontjából kiegyenlített volt (17 nő, 17 férfi), az adatközlők életkora 18–35 év közé esett (átlag = 23 év). Az adatközlőknek saját bevallásuk szerint nincs a gégeműködést befolyásoló betegséjük. A felvételek elkészítése előtt az adatközlők írásos és szóbeli tájékoztatást kaptak a kísérlet menetéről, az adatok anonim kezeléséről és arról, hogy részvételi beleegyezésüket bármikor, mindenféle hátrányos következmény

nélkül visszavonhatják. A tájékoztató után az adatközlők beleegyező nyilatkozatot töltöttek ki.

## Statisztikai elemzések

A statisztikai elemzéseket az R szoftver segítségével (R 2018) végeztük. A vizsgált független változók hatását a nemmodális fonációtípusok gyakoriságára  $\chi^2$ -próbával elemeztük. Továbbá lineáris kevert modellekkel elemeztük a sikeresség és a váratlanság hatását a mért akusztikai és artikulációs paraméterekre az lmerTest csomag (Kuznetsova et al. 2015) használatával. A modellekben független változóként szerepelt még a hangsúlyosság és az adatközlő neme, illetve a független változók interakciója is. A modellek random hatásként az adatközlőt tartalmazták. Ennek a random hatásnak a szignifikanciáját az lmerTest csomag ranova funkciójával vizsgáltuk, azonban, amennyiben a random hatás nem volt szignifikáns, nem távolítottuk el a modelltől, hiszen az ismételt méréses kísérleti elrendezést ez a random hatásos modell reprezentálja leginkább.

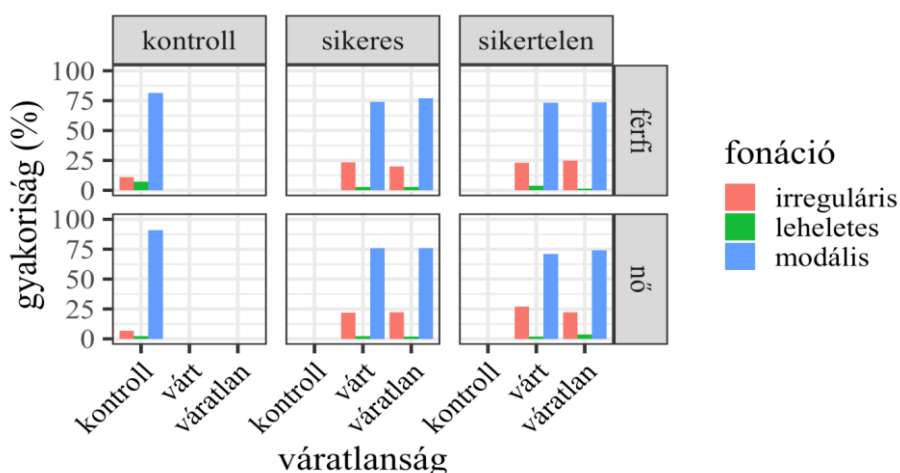
A kevert modellek fix hatásainak szignifikanciatesztelése számos módon történhet. Több százezer szimulált adathalmazon végzett elemzés alapján Luke (2017) megállapítja, hogy a leginkább konzervatív, egyúttal a minta méretére legkevésbé érzékeny metódusok a szignifikancia tesztelésére az úgynevezett restricted maximum likelihood (REML) becsléssel illesztett modellekre alkalmazott Satterthwaite, illetve Kenward-Roger approximáció. A  $p$ -értékek számítása ezt figyelembe véve a modelleken végzett F-próbával, Satterthwaite-approximációt alkalmazva történt, az lmerTest csomag anova funkciójának segítségével.

A kísérleti elrendezésben ugyanazon független változók hatását három különböző függő változó esetén vizsgáltuk. Roettger (2019) szimulációi alapján a mért függő változók számának növelésével (attól függetlenül, hogy a változók korrelálnak-e egymással) nagymértékben megnő az elsőfajú hibák gyakorisága. Ezért a szignifikancia tesztelése során Bonferroni-korrektíót alkalmaztunk ( $\alpha = 0,017$ ). Ennek megfelelően az adatok ábrázolásánál is a korrigált, tehát 98.3%-os konfidenciaintervallumokat alkalmaztunk, melyeket továbbá az ismételt méréses kísérleti elrendezés miatt Morey (2008) módszerével normalizáltunk. A hatások nagyságát  $\chi^2$ -próba esetén a Cramér-féle  $V$  együttható, kevert modellek esetén pedig a parciális eta négyzet ( $\eta_p^2$ ) számításával vizsgáltuk.

## Eredmények

### A fonációtípusok gyakorisága

A sikeresség ( $\chi^2 = 48,05$ ,  $p < 0,001$ ,  $V = 0,091$ ) és a váratlanság ( $\chi^2 = 46,8$ ,  $p < 0,001$ ,  $V = 0,09$ ) változók szignifikáns hatással voltak a fonációtípusok gyakoriságára, azonban, ahogy ez a 3. ábrán is látszik, ez a hatás főként a kontrollkondíció és a kísérleti manipuláció összes szintje közötti általános különbségeknek tudható be: a nemmodális fonáció a kísérleti manipuláció hatására általánosan gyakoribbá vált, a kísérleti manipuláció különböző szintjei között azonban nem találtunk eltéréseket.

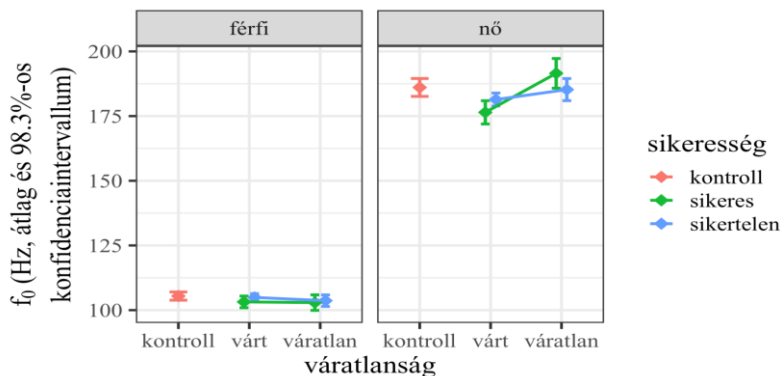


3. ábra: A fonációtípusok gyakorisága a váratlanság, sikeresség és a nem függvényében

### Az alapfrekvencia változásai

Az adatközlő személye nem bírt szignifikáns hatással a változóra: a random hatást tartalmazó modell esetén az AIC értéke magasabb (19092) volt, mint a random hatást nem tartalmazó modell (17874) esetén. Szignifikáns interakciós hatást találtunk a sikeresség, váratlanság és a nem változók között ( $F(1, 2082) = 5,88$ ,  $p = 0,017$ ,  $\eta_p^2 = 0,823$ ) (vö. 4. ábra).





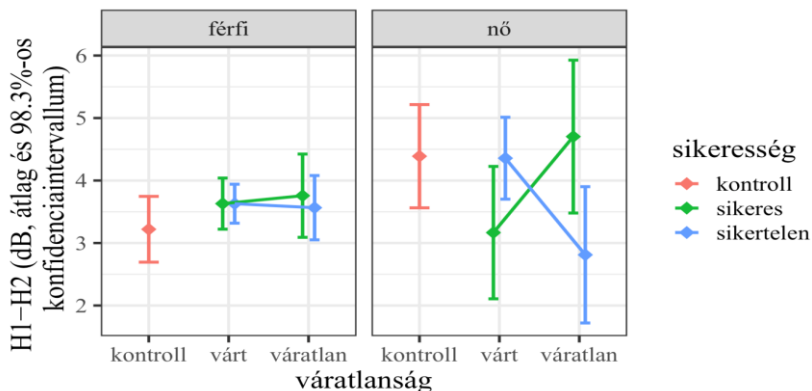
4. ábra: Az alapfrekvencia értékei a váratlanság, sikeresség és a nem függvényében

Míg a férfiak esetén nem volt tapasztalható eltérés az  $f_0$  értékeiben a váratlanság és a sikeresség függvényében, a nők esetében az alapfrekvencia magasabb volt váratlan, sikeres játékeseményeknél, mint a kontrollkondícióban, valamint a várt ingerek az alapfrekvencia csökkenéséhez vezettek. Az utóbbi hatás erősebb volt sikeres játékesemények esetén, mint hátráltató eseményeknél. A hatás mértékét nem befolyásolta, hogy a vizsgált magánhangzó hangsúlyos vagy hangsúlytalan pozícióban volt, ugyanakkor általánosságban az  $f_0$  értékeire szignifikáns hatással volt a hangsúly ( $F(1, 2081) = 348,179, p < 0,001$ ): hangsúlyos pozícióban az  $f_0$  átlagosan 150 Hz, hangsúlytalan pozícióban 135 Hz volt.

### A H1–H2 változásai

Az adatközlő személye nem bírt szignifikáns hatással a változóra: a random hatást tartalmazó modell esetén az AIC értéke magasabb (12849) volt, mint a random hatást nem tartalmazó modell (12247) esetén. Szignifikáns interakciós hatást találtunk a sikeresség, váratlanság és a nem változók között ( $F(1, 2085) = 9,743, p < 0,01, \eta_p^2 = 0,98$ ) (vö. 5. ábra). A férfiak esetében az  $f_0$  értékeihez hasonlóan a H1–H2 esetében sem tapasztaltunk szignifikáns eltéréseket a várt–váratlan, illetve sikeres–sikertelen kondíciók között, csupán a H1–H2 általános növekedése volt megfigyelhető a kísérleti kondíciókban a kontrollkondícióhoz viszonyítva. A nők esetében a váratlanság eltérő hatása volt megfigyelhető a sikeresség függvényében: váratlan, sikertelen, valamint várt, sikeres események során a H1–H2 értékei csökkentek, míg a prediktorok szintjeinek másik két lehetséges kombinációja esetén (tehát váratlan, sikeres és várt, sikertelen eseményeknél) a kontrollkondícióhoz hasonló értékeket mértünk. Ez a hatás szintén nem függött a vizsgált

magánhangzó hangsúlyhelyzetétől, de az alaphangfrekvenciához hasonlóan általánosságban a H1–H2 értékeire szignifikáns hatással volt a hangsúly ( $F(1, 2083) = 11,188, p < 0,001$ ): hangsúlyos pozícióban a H1–H2 értéke átlagosan 4,07 dB, hangsúlytalan pozícióban 3,33 dB volt.



5. ábra: A H1–H2 értékei a váratlanság, sikeresség és a nem függvényében

## Következtetések

Az itt bemutatott kísérlet célja annak megragadása volt, hogy potenciálisan érzelmi reakciót kiváltó ingerek milyen zöngképészesbeli változásokhoz vezethetnek. Két stimulus-ellenőrzési lépés, az egyén elvárásaitól való eltérés (váratlanság), illetve az egyén céljaihoz való hozzájárulás lehetséges kimeneteleit manipuláltuk egy tippelésen alapuló számítógépes játék során. Az érzelemindukció zöngképészesre kifejtett hatásait a játék menetét irányító hangutasítások elemzésével vizsgáltuk. A zöngge jellemzése egyrészt modális és nemmodális fonációtípusok, másrészt két akusztikai mérőszám, a zöngge alaphangfrekvenciája és az első két harmonikus amplitúdójának különbsége (H1–H2) mérésével történt.

Bár a manipulált stimulus-ellenőrzési lépések különböző kimenetelei között nem találtunk különbséget a nemmodális fonációtípusok gyakoriságában, a hangutasítások modális részén mért akusztikai paraméterek esetében szignifikáns interakciós hatás volt a váratlanság, hozzájárulás, illetve a beszélő neme között. Ugyanakkor a nemmodális fonációtípusok gyakorisága általánosan, a konkrét játék helyzetétől függetlenül nagyobb volt az érzelemindukciós játék során, mint a játék után rögzített kontroll-kondícióban. Ez azt jelentheti, hogy bár a játék helyzetben való részvétel általánosságban nagymértékben befolyásolja a zöngképészes, az erősen kontrollált stimulusellenőrzési kimenetek között csak kisebb mértékű akusztikai változásokat tapasztalhatunk.

A két akusztikai paraméter esetén megfigyelhető interakciós hatások a

manipulált stimulus-ellenőrzési lépések és a beszélő neme között arra utalnak, hogy a kiváltani tervezett érzelmi reakciók kizárólag nők beszédében jelennek meg. Erre az érzelmi reaktivitás és érzelemszabályozás mértékének esetleges nemi eltérései adhatnak magyarázatot. Azon a laikus megfigyelésen túl, miszerint a nők általában érzelmesebbek, mint a férfiak (Grossman–Wood 1993), számos, az érzelmi aktiváció (arousal) szintjét különböző fiziológiai mérőszámokkal megragadó kísérlet eredményei támasztják alá a nők nagyobb mértékű érzelmi reaktivitását (Grossman–Wood 1993; Orozco–Ehlers 1998; Bradley et al. 2001; Kemp et al. 2004). Az itt bemutatott kísérlet eredményei tudomásunk szerint a fonáció érzelemfüggő változásainak vizsgálatában elsőként tárnak fel az érzelmi reaktivitás nemek közötti különbségeiből levezethető tendenciákat. Összességében az itt bemutatott eredmények arra irányíthatják a figyelmünket, hogy mind a fonáció, mind általánosságban a beszédképzés és az érzelmi állapotok összefüggéseinek vizsgálatát árnyalhatja az adatközlők érzelmi reaktivitásának, illetve más pszichológiai sajátosságainak figyelembevétele.

Felmerülhet továbbá, hogy az érzelmi reaktivitás mértékét jelen esetben a módszertani megoldás is befolyásolta: a számítógépes játékban való részvétel gyakorlata sztereotipikusan közelebb áll a férfi adatközlőhöz, ami esetleg ahhoz vezethetne, hogy a számítógépes játékokban nagyobb rutinnal rendelkező férfiak kisebb érzelmi bevonódással képesek részt venni egy ilyen feladatban. Ugyanakkor Eklund (2016) eredményei alapján a számítógépes kártyajátékokat magukba tömörítő úgynevezett “tradicionális” digitális játékokat ugyanolyan arányban játsszák nők és férfiak, tehát valószínű, hogy a kísérletben alkalmazott metódusnál ilyen hatás nem jelenik meg, azonban a jövőben mindenképpen érdemes lehet más, nem feltétlenül számítógép használatán alapuló metódusokkal is vizsgálni a stimulusellenőrzés lehetséges kimeneteleinek hatását.

Eredményeink azt mutatják, hogy a nők beszédében az alapfrekvencia magasabb, amennyiben váratlan és céljaikhoz hozzájáruló eseményeket tapasztalnak meg, míg az elvárásoknak megfelelő játékesemények esetén az alapfrekvencia csökkenése figyelhető meg. Ezek a hatások feltételezhetően a váratlan ingerek következtében beálló feszítettebb izomtónus, valamint a várt, a célhoz hozzájáruló ingerek által kiváltott relaxált állapot következményei (Johnstone et al. 2001). A fenti két hatás közül a relaxáció irányába való elmozdulás a nagyobb mértékű.

A H1–H2 értékei a nők produkciójában csökkenést mutattak váratlan, hátráltató, illetve várt, hozzájáruló események esetén. A H1–H2 értékeinek csökkenése váratlan, hátráltató kondícióban azt jelezheti, hogy a nők habituálisan leheletesebb zöngképzése modálisabbá válik (Hanson–Chuang 1999), feltehetően az ebből az érzelmi állapotból levezethető feszítettebb izomtónus következtében (Johnstone et al. 2001). Várt, az egyén céljaihoz

hozzájáruló ingerek esetében érdemes a H1–H2 csökkenését az alapfrekvencia változásaival együtt értelmezni: ebben a kondícióban az alacsony H1–H2 értékek alacsony  $f_0$ -értékekkel társulnak, amely jellemzők inkább az irreguláris fonáció megjelenésére utalhatnak. Az irregularitás megjelenése ebben a kondícióban feltehetőleg a nagymértékű relaxáció eredménye, hasonlóan a magyar nők eljátszott érzelmeinél, a megelégedettség kifejezésében megjelenő gyakori irreguláris fonációhoz (Bartók 2018).

Az előbbieken tárgyalt hatások függetlenek voltak attól, hogy a hangutasítások első, hangsúlyos, vagy második, hangsúlytalan magánhangzóját vizsgáltuk, annak ellenére, hogy az automatikus érzelemfelismerés hatékonysága különbségeket mutat az érzelmi tartalmat hordozó szótag hangsúlyhelyzete alapján (Seppi et al. 2010). Ezt az eredményt az is magyarázhatja, hogy mind az  $f_0$ , mind a H1–H2 esetén a hangsúlyhelyzet szignifikáns főhatását figyelhettük meg, tehát ezek a paraméterekben alapvetően részt vesznek a hangsúly jelölésében. Ennek a nyelvi funkciónak a betöltése ezeknél a paramétereknél feltételezhetően gátolja a paranyelvi tartalom változatossága következtében kialakuló akusztikai variabilitást.

Az itt bemutatott kísérletben a játékesemények manipulációja tehát több olyan akusztikai változást eredményezett a beszédjelben, amelyek feltételezhetően a különböző játékesemények nemtudatos értékelésének hatására bekövetkező spontán fiziológiai változások akusztikai következményei voltak. Az a megfigyelés, hogy az emberi beszédben az érzelmi tartalom megjelenése a beszélőt érő ingereket feldolgozó nemtudatos értékelési lépésekből levezethető fiziológiai-biológiai változások által lehet meghatározva, megerősíti azokat a megállapításokat, amelyek szerint az érzelmi tartalom kifejezése hasonlóságokat mutat különböző állatfajok (Morton 1977), és feltehetően az emberi és állati érzelemkifejezés között is (Darwin 1872; Xu et al. 2010). Ezt támasztják alá azok az eredmények is, amelyek szerint az emberi és állati hangadásokban megjelenő érzelmi tartalom észlelése hasonló mechanizmusokat követ emberek és kutyák percepciójában (Andics et al. 2014).

## Irodalom

- Aharonson, V. – Amir, N. 2006. Emotion elicitation in a computerized gambling game. In: *International Conference on Information Technology: Research and Education*. Tel-Aviv, Israel. 179–183.
- Aichinger, P. 2015. *Diphthongic voice - Definitions, models, and detection*. Thesis. Medical University of Vienna, Austria.
- Andics, A. – Gácsi, M. – Faragó, T. – Kis, A. – Miklósi, Á. 2014. Voice-

- sensitive regions in the dog and human brain are revealed by comparative fMRI. *Current biology* 24. 574–578.
- Bartók M. 2018. A gégeműködés variabilitása az érzelemkifejezés függvényében. *Beszédkutatás* 2018. 30–62.
- Boersma, P. – Weenink, D. 2018. PRAAT: Doing phonetics by computer. 6.0.31 (Letöltve: 2017. augusztus 22.)
- Boiten, F. A. – Frijda, N. H. – Wientjes, C. J. 1994. Emotions and respiratory patterns: review and critical analysis. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology* 17. 103–128.
- Bradley, M. M. – Codispoti, M. – Sabatinelli, D. – Lang, P. J. 2001. Emotion and motivation II: sex differences in picture processing. *Emotion (Washington, D.C.)* 1. 300–319.
- Darwin, C. 1872. *The expression of the emotions in man and animals*. John Murray, London, England.
- Eager, C. 2015. Automated voicing analysis in Praat: statistically equivalent to manual segmentation. In: *The Scottish Consortium for ICPHS*. 551–585.
- Eklund, L. 2016. Who are the casual gamers? Gender tropes and tokenism in game culture. In: Leaver, T. – Willson, M. (eds.), *Social, Casual and Mobile Games: The Changing Gaming Landscape*. Bloomsbury Academic. 15–29.
- Fairbanks, G. 1960. *Voice and articulation drillbook*. Harper, New York.
- Fowler, C. A. 1981. Production and perception of coarticulation among stressed and unstressed vowels. *Journal of Speech, Language, and Hearing Research* 24. 127–139.
- Grossman, M. – Wood, W. 1993. Sex differences in intensity of emotional experience: A social role interpretation. *Journal of personality and social psychology* 65. 1010–1022.
- Hansen, J. H. L. – Patil, S. A. 2007. Speech under stress: Analysis, modeling and recognition. In: Müller, C. (ed.) *Speaker Classification*. Lecture Notes in Computer Science. 108–137.
- Hanson, H. M. – Chuang, E. S. 1999. Glottal characteristics of male speakers: acoustic correlates and comparison with female data. *The Journal of the Acoustical Society of America* 106. 1064–1077.
- Huffman, M. K. 1987. Measures of phonation type in Hmong. *The Journal of the Acoustical Society of America* 81. 495–504.
- Johnstone, T. 2001. *The effect of emotion on voice production and speech acoustics*. Thesis. Crawley, University of Western Australia.
- Johnstone, T. – Van Reekum, C. M. – Scherer, K. R. 2001. Vocal expression correlates of appraisal processes. In: *Appraisal processes in emotion: Theory, methods, research*. Series in affective science. Oxford University Press, New York, NY, US. 271–284.

- Keating, P. A. – Garellek, M. – Kreiman, J. 2015. Acoustic properties of different kinds of creaky voice. In: *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, Scotland.
- Kemp, A. H. – Silberstein, R. B. – Armstrong, S. M. – Nathan, P. J. 2004. Gender differences in the cortical electrophysiological processing of visual emotional stimuli. *NeuroImage* 21. 632–646.
- Klatt, D. H. – Klatt, L. 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.
- Kuznetsova, A. – Brockhoff, P. – Rune, H. B. C. 2015. LmerTest: Tests in linear mixed effects models. *R Package Version 2*.
- Laukkanen, A.-M. – Vilkmán, E. – Alku, P. – Oksanen, H. 1996. Physical variations related to stress and emotional state: a preliminary study. *Journal of Phonetics* 24. 313–335.
- Laver, J. 1980. *The Phonetic Description of Voice Quality*. Cambridge University Press, Cambridge.
- Lazarus, R. S. 1968. Emotions and adaptation: Conceptual and empirical relations. *Nebraska Symposium on Motivation* 16. 175–266.
- Luke, S. G. 2017. Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods* 49. 1494–1502.
- Lynn, R. 1966. The orientation reaction. In: Lynn, R. (ed.) *Attention, Arousal and the Orientation Reaction*. International Series of Monographs in Experimental Psychology. Pergamon. 1–13.
- Morey, R. D. 2008. Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology* 4. 61–64.
- Morton, E. S. 1977. On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist* 111. 855–869.
- Murray, I. R. – Arnott, J. L. 1993. Toward the simulation of emotion in synthetic speech: a review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America* 93. 1097–1108.
- Orozco, S. – Ehlers, C. L. 1998. Gender differences in electrophysiological responses to facial stimuli. *Biological Psychiatry* 44. 281–289.
- Ortony, A. – Clore, G. L. – Collins, A. 1990. *The cognitive structure of emotions*. Cambridge University Press.
- R, Core Team 2018. R: A language and environment for statistical computing.
- Reichel, U. 2012. PerMA and Balloon: Tools for string alignment and text processing. In: *Proceedings of INTERSPEECH* 346.
- Roettger, T. B. 2019. Researcher degrees of freedom in phonetic research. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 10. 1.

- Russell, J. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39. 1161–1178.
- Scherer, K. R. 1986. Vocal affect expression: a review and a model for future research. *Psychological Bulletin* 99. 143–165.
- Scherer, K. R. 1987. Toward a dynamic theory of emotion: The component process model of affective states. In: *Geneva Studies in Emotion and Communication*. 1–98.
- Scherer, K. R. 2001. Appraisal considered as a process of multilevel sequential checking. In: *Appraisal processes in emotion: Theory, methods, research*. Series in affective science. Oxford University Press, New York, NY, US. 92–120.
- Schiel, F. 1999. Automatic phonetic transcription of nonprompted speech. In: *Proceedings of the International Congress of Phonetic Sciences*. 607–610.
- Schuller, B. – Batliner, A. 2014. *Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing*. Chichester, UK, Wiley.
- Seppi, D. – Batliner, A. – Steidl, S. – Schuller, B. W. – Nöth, E. 2010. Word accent and emotion. In: *Proceedings of Speech Prosody 2010*.
- Shahid, S. – Kraemer, E. – Swerts, M. 2007. Audiovisual emotional speech of game playing children: Effects of age and culture. In: *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. 2681–2684.
- Sokolov, E. N. – Cacioppo, J. T. 1997. Orienting and defense reflexes: Vector coding the cardiac response. In: *Attention and orienting: Sensory and motivational processes*. Lawrence Erlbaum Associates Publishers Mahwah, NJ, US. 1–22.
- Sondhi, S. – Khan, M. – Vijay, R. – Salhan, A. 2015. Vocal indicators of emotional stress. *International Journal of Computer Applications* 122. 38–43.
- Svebak, S. 1983. The effect of information load, emotional load and motivational state upon tonic physiological activation. In: *Biological and psychological basis of psychosomatic disease*. Advances in the biosciences, Vol. 42. Pergamon Press, Elmsford, NY, US. 61–73.
- Wendahl, R. W. – Moore, P. – Hollien, H. 1963. Comments on vocal fry. *Folia Phoniatrica* 15. 251–255.
- Xu, Y. T. – Kelly, A. O. – Smillie, C. 2010. Emotional expressions as communicative signals. In: Hancil, S. – Hirst, D. (eds.) *Prosody and Iconicity*. 33–60.

### **Köszönetnyilvánítás**

Az itt bemutatott munka nem jöhetett volna létre Deme Andrea és Markó Alexandra áldozatos témavezetői segítségével.  
A kutatás folyamán a szerző az Emberi Erőforrások Minisztériuma által adományozott NTP-NFTÖ-17 Nemzet Fiatal Tehetségeiért Ösztöndíjban részesült.

### **Phonatory changes during emotion-inducing game events**

Phonatory changes during emotion-inducing game events: the effect of discrepancy from expectations and goal conduciveness

This paper aims to describe how phonation changes during emotion-inducing stimuli. 34 Hungarian speakers (17 female, 17 male) were asked to compete in a simple computerised guessing game, using voice commands to proceed after faced with the result of their guess. We expected that acoustic measures taken on these voice commands differ based on two affective components: goal conduciveness (successful or unsuccessful result) and discrepancy from expectations (unexpected or expected result based on the uncertainty of the guess).

According to the results, only female subjects show phonatory variation as a result of varying emotional states: their fundamental frequency is higher at unexpected game events and their phonation is less breathy (lower H1-H2) at unexpected, unsuccessful events. Both of these changes can be caused by higher muscle tension in unexpected, unsuccessful situations and lower tension when experiencing goal conducive, expected events.