

Budapesti Műszaki és Gazdaságtudományi Egyetem, Természettudományi Kar

Pszichológia Doktori Iskola

Tisljár-Szabó Eszter

Érzelmi állapotok és módosult tudatállapotok beszédjellemzőinek vizsgálata

Doktori értekezés

Témavezető: Dr. Pléh Csaba

2014

Tartalom

Köszönetnyilvánítás	3
Bevezetés.....	4
Új tudományos eredmények, tézisek.....	7
Tudományos közlemények.....	10
Tanulmány [1]	11
Tanulmány [2]	57
Tanulmány [3]	93
Tanulmány [4]	111
Diskusszió.....	139
Alkalmazhatóság, konklúzió	141
Összesített irodalomjegyzék.....	143

Köszönetnyilvánítás

Először is szeretném megköszönni Pléh Csabának, témavezetőmnek, hogy bízott bennem és a témámban, és lehetőséget biztosított számomra a doktori programban való részvételre. Az itt elsajátított tudás alapvetően határozta meg kutatásokkal kapcsolatos szemléletemet.

Köszönet illeti a Kognitív Tudományi Tanszék azon munkatársait, akik oktatóként, kutatóként vagy PhD hallgatóként jelen voltak 2008 és 2011 között a tanszék életében, és megtanítottak valamire, felhívták a figyelmemet dolgokra, vagy szakmai jótanácsokkal láttak el. Enélkül az inspiráló környezet nélkül nem jött volna létre ez a munka. Külön köszönet Lukács Áginak, akihez mindig fordulhattam kérdéseimmel. Köszönet Forgács Bálintnak, Pajkossy Petinek, Keresztes Attilának, Tauzin Tibinek, akik tanácsaikkal, megjegyzéseikkel, kérdéseikkel sokat segítettek a kutatások formálódásában.

Külön köszönet Németh Dezsőnek és Janacsek Karolinának, akik készséggel segítettek cikkem tökéletesítésében.

Külföldi tartózkodásom során az Osztrák-Magyar Akció Alapítvány támogatott, Bécsben pedig Florian Menz volt, aki biztosította számomra a kísérletezés feltételeit. A Debreceni Egyetem Magatartástudományi Intézetében köszönöm Bugán Antal segítségét, aki türelemmel szemlélte a fokozatszerzéssel kapcsolatos történéseket, és támogatta a konferenciákon való részvételeimet.

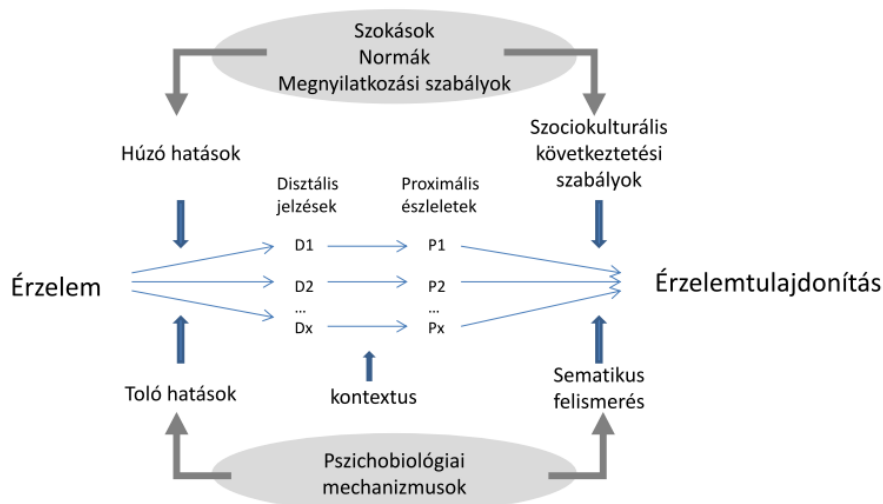
Köszönöm családomnak kitartó bizalmukat és támogatásukat, és külön köszönet apukámnak, aki átadta kutatásokkal kapcsolatos szemléletét, és aki sokat segített a módszertani kérdések megválaszolásában.

Köszönöm férjemnek, Rolinak, hogy négy éve töretlenül mindenben mellettem van.

Bevezetés

Számos modell és kísérleti bizonyíték áll rendelkezésünkre azzal kapcsolatban, hogy a beszédprodukción és a beszédészlelés hogyan működik (pl. Garrett, 1988; Levelt, 1989; Marslen-Wilson, Tyler, 1980). Kevesebbet tudunk azonban arról, hogy a normálistól vagy semlegestől eltérő érzelmi állapotokban vagy módosult tudatállapotokban hogyan változik meg a beszéd, illetve hogy milyen beszédjegyek segítik a hallgatót a beszélgetés során a beszélő érzelmi vagy egyéb, a „semlegestől” eltérő állapotának felismerésében, beazonosításában. Hétköznapi tapasztalataink is vannak arról, hogy hogyan beszél az, aki vidám, mert valami jó dolog történt vele, vagy izgul egy vizsgán, amelynek nagy tétje van, és azt is tudjuk, hogy barátaink, jó ismerőseink hangjából elég nagy pontossággal felismerjük, hogy milyen érzelmi állapotban vannak, vagy akár hogy mennyi alkoholt ittak a megelőző órákban. Mégis fontos kutatási kérdés, hogy a különböző pszichológiai állapotváltozások során pontosan mi változik meg a beszédben, és hogy mi az, ami ténylegesen befolyásolja a hallgatót, amikor a beszélő érzelmi vagy mentális állapotára, személyiségére vonatkozóan következtetéseket von le.

Elméleti keretként Scherer (2003, 2013) a vokális érzelemkifejezés kontextusának modelljét alkalmazzuk (ld. 1. ábra). A modell szerint a folyamat első állomása során a beszélő különböző – objektív módszerekkel is megfigyelhető – hang- és beszédjegyek segítségével kifejezi (kódolja) érzelmi állapotát: az érzelmi arousal megváltozása olyan fiziológiai változásokkal jár együtt, amelyek a légzés, a hangképzés és az artikulációs folyamatok tekintetében érzelemspecifikus prozódiai mintázatokat eredményeznek. Ezek a kombinációk az ún. disztális (ti. a megfigyelő szempontjából távoli) jelzések, amelyek a beszéd révén eljutnak a hallgatóhoz. Ez utóbbi, a hallgatóhoz eljutott, észlelt jelzések pedig az ún. proximális (ti. a megfigyelőhöz közelebbi) jelzések vagy észleletek. Felmerül, hogy pontosan mit is tekinthetünk proximális észleletnek: az alaphártya rezgéseit, a hallókéreg neuronjainak tüzelési mintázatát vagy a hallókérgi reprezentációt? Scherer (2013) amellett érvel, hogy a proximális jelzés a hallókérgi reprezentációnak felel meg, így pl. az alaphérfrekvenciához tartozó proximális jelzés a hangmagasság lenne. Ha elfogadjuk ezt a feltevést, akkor lehetővé válik a proximális jelzések, a reprezentáció (tudatos részének) mérése önbeszámoló segítségével; például egy vizsgálati személy be tud róla számolni, hogy azért gondolja, hogy a beszélő vidám, mert magas a hangja és gyorsan beszél.



1. ábra Scherer a vokális érzelemkifejezés kontextusának modellje (Scherer, 2003, 2013 és Grandjean, 2006 nyomán)

A modell rámutat arra a tényre, hogy az objektíven mérhető disztális jelzések nem feltétlenül azonosak azokkal a proximális jelzésekkel, amelyeket „okoznak”. Bár tény, hogy a proximális jelzések a disztálisakon alapulnak, ez utóbbiakat módosíthatják, torzíthatják az átvivő csatorna (távolság, zaj) vagy az észlelő szerv sajátosságai (pl. bizonyos frekvenciák felerősítése).

A dekódolás végeredménye tulajdonképpen egy, a beszélő (érzelmi) állapotára vagy személyiségére vonatkozó következtetés, amely a beszéd megváltozásának, a proximális jelzéseknek az internalizálásán alapszik. Azonban számos helyen hiba csúszhat a folyamatba. Lehetséges például, hogy egy meglévő érzelmi állapot nem okoz nagy változásokat a disztális jelzésekben (például mert el tudjuk rejteni valódi érzéseinket). Az is előfordulhat, hogy a disztális jegyek ugyan megfelelően kifejezik az adott állapotot, de az átvitel vagy az észlelés során olyan módon torzulnak vagy módosulnak (például a zaj miatt), hogy proximálisan, a hallgató számára már nem hordozzák a lényegi információkat. Valamint az is lehetséges, hogy bár a disztális és a proximális jellemzők megegyeznek, a hallgató következtetési folyamata, a kettő közötti kapcsolat reprezentációja „romlik el”, például idegen nyelvet hallgatva másféle érzelmeket tulajdonítunk a beszélőnek, mint amit ő valójában kifejezni szándékolt. Az ábrán látjuk, hogy modell szerint mind az érzelemkifejezést (kódolást), mind az azonosítási (dekódolási) folyamatot pszichobiológiai mechanizmusok (toló hatások) és szociális normák, megnyilatkozási szabályok (húzó hatások) befolyásolják.

Míg az említett jelenségek a természetes szociális interakciókban együtt vannak jelen, a kutatás szempontjából elválhatnak egymástól. Ha arra vagyunk kíváncsiak, hogy bizonyos állapotok milyen vokális jegyekkel járnak együtt (hogyan melyek annak disztális jelzései), akkor először is pontosan definiálnunk kell az állapotokat. Tehát tudnunk kell, hogy mi alapján állítjuk, hogy egy beszélő szomorú, depressziós vagy extrovertált. Ellenben ha arra vagyunk kíváncsiak, hogy bizonyos paralingvisztikai jegyek kombinációja hogyan működik proximális jelzésként (ti. hogy beszédjegyek bizonyos kombinációi milyen érzelmítéletet váltanak ki), akkor az a fontos, hogy a hallgatók ítéletei megbízhatóak legyenek – például magas fokú egyetértés legyen abban, hogy milyen vonásokat vagy állapotokat tulajdonítanak a beszélőnek. Tudnunk kell tehát, hogy mi a kutatásunk célja, mert ez fogja meghatározni a kutatási paradigmát. Ha az a cél, hogy megértsük, hogy a beszélő aktuális (érzelmi) állapota hogyan befolyásolja a vokális kifejezést, akkor ún. produkciós kutatásra van szükség, ahol eleve ismertek és definiáltak az (érzelmi) állapotok, és az ezekben az állapotokban elmondott beszédmintákat vizsgáljuk artikulációs vagy akusztikai szempontból. Ha a cél annak megismerése, hogy a disztális jegyek egyes konfigurációja milyen érzelm- vagy állapottulajdonításhoz vezet, akkor percepciós kutatást kell végeznünk. Vizsgálhatjuk csupán egy-egy beszédparaméter érzelmtulajdonításban betöltött szerepét is, ilyenkor szisztematikusan manipuláljuk a disztális jegyeket, hogy pontosan felmérjük a következtetés ok-okozati szerkezetét. (A módszerekről részletesebben ld. az [1] tanulmányt).

Hangsúlyozzuk, hogy a modell általában az érzelmek kifejezésével kapcsolatos, tehát nemcsak a vokális, hanem a testbeszéddel vagy az arcon történő érzelmkifejezés folyamatát is magyarázhatja. Mivel azonban jelen dolgozat a belső állapotok és a beszédbeli változások kapcsolatait vizsgálja, így a folyamatnak ezt a részét hangsúlyoztuk.

Bár az általunk elemzett anyagok nyelvi természetűek (megnyilatkozások, beszédrészek), jelen dolgozat nem foglalkozik a beszédprodukció és a beszédészlelés klasszikus pszicholingvisztikai kérdéseivel. Így nem írunk arról, hogy hogyan zajlanak az egyes nyelvi szinteken végbemenő produkciós és megértési folyamatok, nem részletezzük az ezeket tárgyaló modelleket. Ugyanígy nem írunk részletesen az érzelemelméletekről, az érzelmkifejezés és -észlelés fiziológiai és neurológiai hátteréről. Tudomásul vesszük a jelenségek komplexitását, de a vizsgálhatóság érdekében, egy nagyon szűk terület néhány problematikus kérdésével foglalkozunk.

Új tudományos eredmények, tézisek

A beszédszünetek szerepe az érzelemazonosításban anyanyelvű és idegen nyelvű szövegben [2]

1. tézis: Pusztán egy beszédparaméter, a beszédszünetek hosszának megváltozása következtében megváltozik a hallgatók által a beszélőnek tulajdonított érzelmek intenzitása. A beszéd egyéb paramétereit felhasználva a hallgatók automatikusan érzelmi állapotokat tulajdonítanak a beszélőknek tartalmilag semleges szövegek esetében is.
2. tézis: A szünetek hatása az érzelemtulajdonításra megmutatkozik német anyanyelvűeknél is magyar nyelvű szöveget hallgatva. Magyar nyelvű szöveget mind a magyar mind a német anyanyelvű hallgatók szomorúbbnak, kevésbé vidámnak, ijedtebbnek és kevésbé pozitívnak ítélik, ha a szünetek hosszabbak.

Két kísérletünk (Tisljár-Szabó és Pléh, 2014) arra a kérdésre kereste a választ, hogy csupán egy beszédparaméter – a néma szünet hosszának – megváltozása hogyan befolyásolja a megítélőket az érzelemtulajdonításban. Egy további célja volt a kutatásnak annak vizsgálata, hogy idegen nyelvű beszédet hallgatva is hatással van-e a szünethossz az érzelemazonosításra. Egy perc hosszú magyar nyelvű beszédrészleteket használtunk fel, amelyekben a beszélők egy érzelmileg semleges témáról beszéltek (a munkájukról, egy átlagos napjukról). Az ezekben természetes módon előforduló beszédszünetek hosszát manipuláltuk négyféleképpen: minden szünet hosszát megnöveltük vagy lecsökkentettük két-kétféle módon. Az eredeti és a manipulált magyar hanganyagot magyar és német anyanyelvű vizsgálati személyeknek mutattuk be, akiknek az volt a feladatuk, hogy egy 1-6-ig skála segítségével megállapítsák, hogy a hallott beszéd mennyire hangzott szomorúnak, vidámnak, dühösnek stb.. Az eredmények azt mutatták, hogy mind a magyar, mind a német anyanyelvűek ítéleteit befolyásolta a szünetek hossza. A két csoport értékeléseit egybevéve azt mondhatjuk, hogy ugyanazokat a beszédmintákat a vizsgálati személyek szomorúbbnak, kevésbé vidámnak, kevésbé pozitívnak és ijedtebbnek ítélik, ha a szünetek hosszabbak. A szünetek hossza a szomorú-vidám ítéletekre volt a legnagyobb befolyással. Az eredmények alapját tehát úgy tűnik, hogy a szünetek hossza, és így a beszéd-szünet arány a szomorú-vidám dimenzióhoz és

a félelem állapotához kapcsolódik, más érzelmek észlelésében (és valószínűleg a kifejezésében is), valószínűleg más paraméterek megváltozása fontosabb. A szünetek ilyen funkciója a magyar és a német nyelvben azonos.

Az eredmények azt is mutatták, hogy ugyanazoknak a beszédmintáknak a német anyanyelvűek több érzelmet tulajdonítottak, mint a magyar anyanyelvűek. Az összes értékelést és beszélőt figyelembe véve a német anyanyelvűek mérgesebbnek, szomorúbbnak, meglepettebbnek, ijedtebbnek, kevésbé pozitívnak valamint hevesebbnek ítélték a beszélőket, mint a magyarok. Tehát a beszéd paraméterei, mint pl. a hangmagasság, dallam, hangerő, hangszín és ezek együttes előfordulása inkább keltette a német anyanyelvűekben azt a benyomást, hogy a beszélő mérges, szomorú, meglepett, fél, negatív vagy heves, mint a magyarokban.

A kutatás újszerűnek számít a felhasznált módszertana tekintetében. Azok a kutatások, amelyek azt tűzik ki célul, hogy egy-egy beszédparamétert szisztematikusan manipuláljanak és vizsgálják ennek hatását az érzelemítéletekre (pl. Ladd et al., 1985; Bergmann et al., 1988; Breitenstein et al., 2001; Scherer et al., 1984; Cahn, 1990; Carlson, 1992; Burkhardt et al., 2000) színészek által elmondott mondatokat vagy szövegrészeket használnak fel. Bár ezáltal jó minőségű, jól kontrollált ingeranyaghoz jutnak, elvész a beszéd természetessége. Mi olyan beszédreszletek használtunk fel, amelyekben önként résztvevők (nem színészek) beszéltek magukról hosszabb időn keresztül, így a beszéd természetes módon hangzott, és tartalmazta a spontán beszédre jellemző kifejezésmódokat, szüneteket, megakadásokat. A kutatás másik újszerűsége, hogy korábban nem vizsgált módon elemzi a tempó hatását az érzelemítéletekre. Míg a korábbi kutatások az artikulációs tempót és a (szüneteket is magában foglaló) beszédtempót együtt változtatták, addig mi csak a szüneteket manipuláltuk, az artikulációs tempót változatlanul hagyva. Így végül a randomizálást kiváltó latin négyzet (latin square) kísérleti elrendezés lehetővé tette, hogy a kapott eredményeket kizárólag a beszédszünetek hosszának tulajdonítsuk.

Az alkohol hatása a beszédre és a munkamemóriára [3-4]

3. tézis: Az alkohol hatással van a beszédbeli megakadásokra: az összes típusú hibázást (megakadások, elakadások, rontások, kihagyások) figyelembe véve alkoholos állapotban több hibát vétenek a személyek, mint józan állapotban. Ez a fonológiai, a fonetikai, a morfológiai és a szintaktikai típusú nyelvtörők estében is igaz.
4. tézis: A beszéd lassulását alkoholos állapotban a néma szünetek hosszának megnövekedése és nem az artikuláció lassulása okozza.
5. tézis: A mérsékelt mennyiségű alkoholnak a hangmagasságra gyakorolt hatásával kapcsolatos korábbi kutatási eredmények nem egységesek. A saját eredményeink szerint a hangmagasság nem változik 0,08%-os véralkoholszint hatására
6. tézis: Az alkoholnak nincs hatása a számterjedelemteszten nyújtott teljesítményre, azonban hat a nyelvtörők visszamondásához szükséges munkamemória-teljesítményre: Alkoholos állapotban ugyanannyi számot tudnak felidézni a személyek, mint józan állapotban, azonban többször fordul elő, hogy nem tudják befejezni a mondatot (elakadás), hogy más szóval helyettesítik a szót, amit nem tudnak felidézni (rontás) vagy hogy kihagynak egy szót (kihagyás).

A harmadik és a negyedik tanulmány a disztális jelzések vizsgálatának körébe tartozik: azt vizsgáltuk, hogy alkoholos állapot hatására hogyan változnak meg a beszéd bizonyos paraméterei. Kitértünk a rövid távú emlékezetre gyakorolt hatásra is. Mivel az alkohol hatással van a hangulatra, valamint a kognitív és a motoros folyamatokra is, így feltételeztük, hogy alkohol hatására változások következnek be a beszédben. Vizsgálati személyeink feladata az volt, hogy 3x5 cl tömény szeszes ital elfogyasztása után valamint józan állapotban nyelvtöröket mondjanak vissza, ezenkívül mindkét alkalommal felvettük velük a számterjedelemtesztet. Mivel a nyelvtörök visszamondása beszédprodukciós és emlékezeti kapacitást is igényel, ezért az alkohol hatását a továbbiakban külön vizsgáltuk a) a különböző beszédjegyekre, b) az emlékezeti teljesítményre vonatkozóan. A negyedik cikk tartalmazza az

emlékezeti teljesítményre és a pontatlan artikulációra vonatkozó elemzéseket, míg a harmadik cikk részletesebben vizsgálja a hangmagassággal kapcsolatos adatokat.

Az eredmények szerint alkoholos állapotban a személyek több hibával beszélnek és több szünetet tartanak; a hangmagasság és az artikulációs tempó nem változik (Szabó et al., 2012, Tisljár-Szabó et al., 2013). Az alkohol emlékezeti teljesítményre gyakorolt hatásával kapcsolatban azt találtuk, hogy mérsékelt mennyiségű, 0,08%-os véralkoholszintet eredményező alkoholnak nincs hatása a számterjedelemteszten nyújtott teljesítményre, ugyanakkor hat a nyelvtörök visszamondásához szükséges emlékezeti teljesítményre. A vizsgálat egyik alapfeltevése az, hogy a verbális hibák egy részéért az emlékezet megváltozott működése a felelős, háttérükben felidézési problémák állnak. A feltevésünkkel összhangban a megváltozott emlékezeti működést a következő hibatípusokban tapasztaltuk: elakadás (amikor a személyek nem tudják befejezni a mondatot), rontás (amikor más szóval helyettesítik a szót, amit nem tudnak felidézni) és kihagyás (amikor kihagynak egy szót). Alkoholos állapotban tehát a személyek ugyanannyi számot tudtak felidézni, mint józan állapotban, de több elakadás, rontás vagy kihagyás típusú hiba fordult elő a visszamondáskor. (Szabó et al., 2012).

Tudományos közlemények

- [1] Tisljár-Szabó, E. (2014) Érzelem és beszéd. In: Pléh Cs. és Lukács Á. (szerk.) Pszicholingvisztika, 951-986. Budapest, Akadémiai Kiadó.
- [2] Tisljár-Szabó, E., Pléh, C. (2014) Ascribing emotions depending on pause length in native and foreign language speech. *Speech Communication*, 56, 35–48. DOI 10.1016/j.specom.2013.07.009 (IF: 1,283)
- [3] Tisljár-Szabó, E., Rossu, R., Varga, V., Pléh, Cs. (2014) The effect of alcohol on speech production. *Journal of Psycholinguistic Research*. (Online first, DOI 10.1007/s10936-013-9278-y) (IF: 0,642)
- [4] Szabó, E., Rossu, R., Varga, V., Imréczi, K. (2012) Az alkohol hatása a rövid távú emlékezetre és a beszédprodukciós folyamatokra. *Magyar Pszichológiai Szemle*, 67, 3, 409-430. DOI: 10.1556/MPSzle.67.2012.3.1.

Tanulmány [1]

Tisljár-Szabó, E. (2014) Érzelem és beszéd. In: Pléh Cs. és Lukács Á. (szerk.) Pszicholingvisztika, 951-986. Budapest, Akadémiai Kiadó.

(a megjelent változattal tartalmilag megegyező kézirat változat)

Beszédünk, hangunk segítségével sokféleképpen kifejezhetjük érzelmeinket. Gyakori, hogy szavak segítségével fogalmazzuk meg, hogy milyen érzelmi állapotban vagyunk. Amikor azt mondjuk, *boldog vagyok, ideges vagyok* stb., érzelemszavainkat használjuk. Máskor csak körülírjuk hangulatunkat (*azt kívántam, bárcsak örökké tartana az a pillanat*), vagy elmeséljük, mi történt velünk (*hirtelen megjelent vicsorogva egy kóbor kutya*). Azzal hogy milyen szavakat (*a titkárnő* versus *az a kis csitri nem volt hajlandó beengedni*), vagy milyen mondattípusokat használunk (*Bárcsak mehetnék én is! – Megyek én is!*), jelezzük beszélgetőtársunknak lelkiállapotunkat, a témával kapcsolatos érzéseinket. Az a társalgási stratégia, amit folytatunk (pl. demonstráló hallgatás), vagy a hangnem, amiben beszélünk (pl. ironikus) szintén megmutathatja érzelmi állapotunkat (Fiehler, 2001 nyomán).

Érzelmeinket nem csak érzelemkifejező, érzelmi állapotokra utaló szavak segítségével tudjuk kifejezni. Vokális, tehát hanggal együtt járó, de nem verbális kifejezőmód például a nevetés, amely általában örömet fejez ki. A beszédhez szorosan kapcsolódik, de nem a közlés maga a fontos, hanem a mikéntje, amikor érzelmeinket a hangszín, a beszédtempó vagy a hangmagasság változtatásával fejezzük ki (a későbbiekben erről lesz szó bővebben, ld. 3. alfejezet).

Fiehler (2001) kiemeli, hogy az érzelemszabályozás rendszere szituációnként, szerepenként, nemenként, kultúránként változhat, a szabályok semmiféleképpen nem univerzálisak. Az érzelmek vizsgálatakor figyelembe kell venni, hogy az, hogy mi minősül érzelemkifejezésnek, helyzetenként és személyenként változhat. Bármely létező viselkedésmódot és reakciót értelmezhet a befogadó érzelemnek, függetlenül a cselekvő személy szándékától.

A következőkben olyan nyelvészeti és/vagy pszichológiai kutatásokat mutatunk be, amelyek témája az érzelmek és a nyelv, a beszéd, a hang¹. E területek mindegyike abból indul ki, hogy az érzelmek tükrözik a személy lelkiállapotát, és megjelennek a beszédben, a nyelvhasználatban. A munkákat a szerint csoportosítottuk, hogy milyen iskola vagy megközelítés áll háttérükben. Az egyes megközelítések legfőképpen módszertanukban különböznek, ami pedig általában meghatározza, hogy melyek azok a jelenségek, amelyeket vizsgálatuk tárgyává tesznek. Röviden írunk a nyelvészeti társalgáselemzésről (1. alfejezet), amely a hétköznapi párbeszédekben létrejött rendszerességet kutatja. A fejezet második nagyobb témaköre a szavak, a nyilatkozatok jelentésének szerepével foglalkozik és bemutatja a tartalomelemzés módszerét (2. alfejezet). A fejezet középpontjában a vokális, azaz hanggal történő érzelmkifejezés áll (3. alfejezet).²

Scherer (2000) hatféle érzelmi állapotot különböztet meg: preferenciák (kedvel – nem kedvel), érzelmek (szomorú, vidám), hangulatok (derűs, kedvetlen), interperszonális beállítódás (távolságtartó, támogató), attitűd (kedvel, utál) és személyiségvonás (szorongó, féltékeny). Oatley és Jenkins (2001) is különbséget tesz az érzelmek (érzelemepizódok), a hangulatok, az érzelmi zavarok és a személyiségvonások között. Ezek a kategóriák több szempontból is különböznek egymástól, például intenzitásukban, időtartamukban, a kiváltott esemény fókuszában vagy a bekövetkező változás sebességében. Mivel az eddigi kutatások, amelyek az érzelmek és a beszéd kapcsolatát vizsgálják, egyelőre nem terjednek ki mindegyik állapot széleskörű jellemzésére, jelen fejezetben nem vesszük figyelembe ezeket a csoportosításokat. Egy-egy alfejezeten belül egyszerre szerepelhetnek érzelmek, hangulatok vagy személyiségvonások.

¹ Különbséget teszünk a nyelv, a beszéd és a hang között. Létezik például nyelvi közlés beszéd nélkül az írás esetében, vagy nyelvi közlés hang nélkül de beszéddel a jelbeszéd esetében. Beszéd nélkül, csupán hangokkal is sok mindent kifejezhetünk, pl. az említett nevetés vagy a *mmm* hangsor „elégedettséget kifejező” intonációban egy finom étel megkóstolásakor. Azt mondhatjuk, hogy a beszéd esetében felismerhetőek a szavak, a tartalom, míg a hang magára a hangzó jelre vonatkozik, tartalom nélkül. Amikor nyelvről beszélünk, általában a nyelvi rendszer szintjeiről teszünk megállapításokat. A legtöbb esetben ez a három szint egyszerre van jelen.

² Jelen fejezet nem elemzi a szépirodalom, a költészet érzelmkifejező eszközeit. Egy vers formája, verssorainak hossza, lüktetése már önmagában érzelmkifejező, de beszélhetünk a hangok szimbolikájáról, a szavak metaforikus használatáról, a metonímiáról, az ellentétek használatáról. Ezekről ld. bővebben Fónagy (1998) művét.

1. A nyelvészeti társalgáselemzés

A nyelvészeti társalgáselemzés (angolul conversation analysis, németül Konversationsanalyse) elméletét és módszertanát Iványi Zsuzsanna (2001) mutatta be először magyarul. Ez a nyelvészeti terület Amerikából indult el az 1970-es években, onnan került át Európába, különösen angol és német nyelvterületre. Fontos jellemzője, hogy hétköznapi társalgási helyzeteket vizsgál. A szituáció szereplői hétköznapi emberek (foglalkozásukat, szerepüket, szándékaikat tekintve természetesen lehetnek különbözőek), a helyzet maga pedig egy természetes helyzet, nem beállított, nem eltervezett, hanem olyan, mint amelyet nap mint nap átélünk. A kutató a beszélgetéseket felveszi magnóra, újabban digitális diktafonra, majd lejegyzí és elemzi azokat (a társalgáselemzés módszertanáról lásd részletesebben az 1. keretes részt).

A megközelítés feltételezi, hogy a beszélgetés valamilyen szabályrendszer alapján jön létre, amely azonban nem előre meghatározott, hanem a társalgó felek interakciójából lépésről lépésre alakul ki. Ennek próbálja meg leírni törvényszerűségeit, szerkezeti mechanizmusait. Az elemzés alapvetően kvalitatív jellegű. (A 2. keretes részben arra mutatunk röviden példát, hogy hogyan zajlik egy párbeszéd kvalitatív elemzése.)

A társalgáselemzés módszertana

A társalgáselemzés módszertana speciális. Alapvető célkitűzése, hogy a hétköznapi életben előforduló párbeszédet vizsgáljon, egy szociális eseményt akar folyamatában regisztrálni és elemezni. A helyzetek sohasem kísérletiek vagy beállítottak, próbál azonban olyan eseményt keresni, amely viszonylag egész, hiánytalan, mivel az elemzés szempontjából fontos például az is, hogy egy beszélgetés hogyan kezdődik vagy hogyan ér véget.

A felvett anyagot a kutató leírja, ami hosszú és alapos munkát igényel. Nem csupán a személyek által elmondottak tartalma a fontos, hanem egy sor speciális jellel jegyzik le a hangsúlyra, a hangerőre, a tempóra vonatkozó információkat (ezek természetesen némileg szubjektívek), vagy az egymást átfedő megnyilatkozás-részleteket. A legtöbb leírásban helyet kap a körülmények, külső események lejegyzése is (pl. „papírokat lapozgat”, „csörög a telefon”).

Az elemzés során először is a rendszerességet keresik az anyagban. Megfigyelik tehát, hogy milyen jelenségek járnak együtt. Aztán feltárják a problémákat (például az orvos nem reagál a beteg betegséggel kapcsolatos érzelmi kijelentéseire), és azt, hogy az interakciós partnerek hogy próbálják ezeket megoldani (a beteg újból és újból visszatér ugyanahhoz a témához, az orvos pedig újból és újból félbeszakítja a beteget). Végül le kell írni, hogy hogyan hozták létre a beszélgetőpartnerek az új rendszert, és ennek mik a sajátosságai (nem hatékony anamnézis felvétel, rossz érzésekkel távozó beteg). Ezeknek az apró, nehezen észrevehető, ismétlődő jegyeknek a gyűjteménye adja a vizsgálat tulajdonképpeni adatbázisát (Lalouschek, 2002ab; Lalouschek és Menz 2002).

A beszélgetések lejegyzésekor a társalgáselemzés kutatói leggyakrabban a GAT (Gesprächsanalytisches Transkriptionssystem) rendszerét használják, amelynek kidolgozása Selting és munkatársai (1998) nevéhez fűződik. A következőkben egy német nyelvből magyarított példán keresztül mutatjuk be az egyes jelenségekre használt rövidítéseket, jeleket, a teljesség igénye nélkül. Az beszélgetésben az A-val jelölt férfi az autóbalesetét meséli el az interjúkészítőnek (I).

- 001 A <<len>> és öö- .hh (---) AKKO::R (---)
 002 akkor éppen: (1.0) öö abban az IDŐBEN amikor aztán- (1.3)
 003 a MÁSODIK (-) sorscsapás történt- (-)
 004 akkor öö po éppen a::
 005 Próbaidőn voltam- (1.0)
 006 IGEN és akkor jött öö hh (2.2)
 007 ez a NAP amikor a balesetem> volt-
 008 I <<pp>=mhm.> (1.3)
 009 A igen: ez volt a próbaidő utolsó hete- (2.4)
 010 és igen attól (--) attól a pillanattól kezdve
 011 [azt mondhatom] minden más lett>
 012 I [mhm]
 013 I Mhm (--)
 014 A Tehát a- (1.6) baleset az, (2,6)
 015 MEGtörtént
 016 <<all> tehát a:: úgy volt, hogy valaki (.) egy másik Autó (.) Vezető (-)
 frontálisan-
 017 Egy TEHERgépkocsi
 018 Szemből belémjött; (--)
 019 .hh teát- (.) a kocsimnak-> (2.2)
 020 Öö nagyon nagy; (1.4)
 021 <<pp> szóval azt a:> (.) <<f> BORzalmas FÁJdalmat>>öö (.)
 022 azt tehát azt SOHA Életemben nem fogom elfelejteni;

<<len>> lento, lassan

<<pp>> pianissimo, nagyon halkán

<<al>> allegro, gyorsan

<<f>> forte, hangosan

, kissé emelkedő hangsúly

; kissé ereszkedő hangsúly

. eső hangsúly

- állandó hangsúly

(.) (-) (--) (---) szünet

(1.3) 1,3 másodperces szünet

;; ::, ::: nyújtás

[] egymást átfedő beszédrészek, amikor a két fél egyszerre beszél

HANGsúly

= gyors, közvetlen folytatás

.h, .hh, .hhh levegővétel

h, hh, hhh kilégzés

Forrás: Deppermann, A., Lucius-Hoene, G. (2005) Trauma erzählen – kommunikative, sprachliche und stimmliche Verfahren der Darstellung traumatischer Erlebnisse. *Psychotherapie und Sozialwissenschaft. Zeitschrift für Qualitative Forschung und klinische Praxis*. 1. 35–73.

1.1.A társalgáselemzés módszere az érzelemkutatásokban

A társalgáselemzés módszerével végzett elemzésben két központi általános kérdés van (Fussell, 2002): (a) milyen érzelmek vannak jelen a társalgásban? (2) ezek hogyan mutatkoznak meg a társalgó felek viselkedésében? Első lépésként meghatározzák az ún. tematizált érzelmeket, tehát amelyeket a beszélők expliciten kifejeznek (pl. „*Nagyon IDEGESÍT, hogy mindig elkésel!*”). Második lépésként megfigyelik, hogy melyek azok az érzelmek, amelyek a helyzetből adódnának (pl. műtét előtti konzultáció az orvossal – félelem). Majd megnézik, hogy melyek azok a kifejezések, jelek, amelyek érzelmekre utalhatnak (krákogás vagy „*Néha azért érkezhetsz pontosan*”), és mondatról mondatra elemzik a párbeszédet. Végül arra tér ki az elemzés, hogy utal-e valamilyen érzelemre a két fél beszédének mintázata (például a gyakori félbeszakítások agresszív vagy domináns kommunikációra utalnak).

Német nyelvterületen, különösen Bécsben és Bielefeldben számos kutatás született az érzelmek kifejezésével kapcsolatban (Fiehler, 1990; Lalouschek és mtsai, 1990; Lalouschek, 2002a; Menz és mtsai, 2008; Sator, 2003). Az elemzések nagy része olyan érzelmeteli szituációkat vizsgál, mint amilyen az orvos-beteg kommunikáció vagy a pszichoterápiás ülés. A kórházi helyzetek elemzésekor abból a kérdésből indultak ki, hogy miért van az, hogy a

betegek szinte minden egyes alkalommal rossz érzésekkel távoznak a rendelőből, még akkor is, ha kiderül, hogy semmi bajuk nincs. Bár természetesen maga a környezet, az orvosi műszerek és más betegek látványa, a gondolat, hogy az ember beteg, szorongást váltanak ki, a kutatók szerint a rossz érzésekhez nagyban hozzájárul az orvos és a beteg közötti párbeszéd zavara is. Számos orvos-beteg interakciót elemezve kiderült, hogy az orvosok például gyakran megsértik a mindennapi beszédhelyzetek szokásos szabályait, így elmaradnak az üdvözlés és az elköszönés vagy a bemutatkozás formulái. A vizsgálatok többségén megfigyelhető, hogy a beszélgetés úgy zajlik, mintha a felek elbeszélnének egymás mellett, mintha meg se hallanák a másikat, amit az okoz, hogy különbözik az orvos és a beteg célja. Az orvos célja, hogy a tünetek alapján minél hatékonyabban, minél gyorsabban (idő- és energiahiány miatt) megállapítsa a diagnózist, és további kezelésre küldje a beteget vagy gyógyszert írjon fel neki. A betegben sokszor nem embert, hanem tünetegyüttest lát, a testi dolgok érdeklik, a beteg lelkivilágát holmi zavaró tényezőnek tekinti. Ezzel szemben a beteg mit sem tud saját testéről, gyakran még a tüneteit sem tudja pontosan megfogalmazni (*amikor felálltam, úgy éreztem, mintha...* Ahelyett, hogy azt mondaná: *ekkor meg ekkor pontosan itt ilyen és ilyen típusú fájdalmat éreztem*). Számára a legfontosabb kérdés, hogy súlyos-e a betegsége, elsősorban azt szeretné, ha megnyugtatón, hogy minden rendben lesz. Úgy érzi, hogy a betegség pontos megállapításához minden részletet el kell mesélnie (*a nagybátyám unokatestvérének is voltak hasonló problémái...*). A nagy különbség, hogy a való életben az élmények, érzések kifejezésekor számíthatunk beszélgetőpartnerünk reakciójára, ami lehet valamilyen együttérzés, érdeklődés, esetleg tanácsadás, míg az orvosi rendelőben ezekre csak ritkán kerül sor. A társalgáselemzés fogalmaival azt mondhatjuk, hogy sérül a szekvencia. Ennek természetesen számos oka van, azonban néhány apró kommunikációs technikával nagyon sokat lehet javítani a betegek közérzetén (ld. Lalouschek, 2002b).

2. Keretes rész: egy rövid párbeszéd kvalitatív elemzése a társalgáselemzés módszerével (Lalouschek 2002b, 71. o. – <i>und jetzt bin ich ganz nervös...</i>)		
Szituáció: A 62 éves páciens, aki nőgyógyászati műtét előtt áll, azt meséli, hogy hogyan, mikor került a kórházba. Kiderül, hogy épp most tudta meg, hogy a műtétet elhalasztották.		
Résztevők: női beteg (B), Orvostanhallgató, diák (D)		
027	B	Kedd óta vagyok itt. És azt gondoltam, hogy holnap
028	B	kerül rám a sor – hogy holnap csinálják, és épp az előbb

		mondta	
	D		mhm
029	B	a doktornő – hogy csak hétfőn csinálják meg. És most	
	D	Ez érthető	Lelkileg
		már	
030	B	tisztára /- ideges/ vagyok és nyugtalan. /nevet/	
	D	felkészítette magát	
031	B	Hát igen, az ember lelkiileg valahogy	
	D	mhm	
032	B	felkészíti magát. – (...) ahh, már kedd óta -	
033	B	És már megvolt minden vizsgálat. Azt hittem	
034	B	hogy most már holnap meglesz. Mert az ember	
	D	Ez teljesen érthető. Egyedül van	
035	B	csak fél egy kicsikét.	
036	D	egész hétvégén vagy – jön valaki meglátogatni?	
	D		Hány
037	B	Igen-igen. Remélem a gyerekek. Ők öö jönnek.	
038	D	gyereke van?	

Lalouschek (2002b, 71.) a következőképpen elemzi a leírt szituációt: A páciens expliciten kifejezi erős érzéseit a pszichésen megterhelő operációval kapcsolatban („ideges”, „nyugtalan” (030)). A diák helyénvalónak nyilvánítja az érzelmet („ez érthető” (030)): azoknak a betegeknek, akiknek eltolják a műtétjük, joguk van erős érzelmeket kifejezni. Ezzel az érzelemre úgy tekint, mint egy általános, mindenkit és nem csak jelen beteget érintő jelenségre. A diák továbbra is marad a kézenfekvő, általános-hétköznapi szinten: a páciens érzelmeit a műtét elhalasztásának technikai okával magyarázza, de nem kérdez rá a háttérben húzódó személyes jelentésre.

Bár a páciens átveszi ezt az értelmezést, átfogalmazza egy személytelenebb és enyhébb formába („az ember”, „valahogy” (031)). Ez az átfogalmazás utalás lehet arra, hogy a beteg más tényezőknek tulajdonítja az érzéseit, mint az orvostanhallgató. Az utalást a diák ismét csak jóváhagyja („mhm” (032)), de nem értelmezi.

A páciens továbbra is ragaszkodik ahhoz, hogy megpróbálja kifejezni érzéseit (032-035). Ahogy a (027) sorban elkezdte, itt újra megfogalmazza és mintegy végszóként, ha személytelenül és enyhébb formában is, de kimondja („*mert az ember csak fél egy kicsikét*” (034/5)), hogy mi az ő félelmének alapvető oka: nem a műtét időbeli eltolása, hanem maga az operáció az, amitől fél. A diák újból nem lép tovább annál, hogy jóváhagyja a közlést („*Ez teljesen érthető*” (035)), lezárja a témát, és „áttolja” a problémát a *hétvégére* (036).

A párbeszéd egy későbbi szakaszából kiderül, hogy a beteg férje másfél évvel ezelőtt, egy műtétet követően veszítette életét, de a diák erre sem reagál személyes kérdéssel vagy megjegyzéssel.

Összességében azt mutatja az elemzés, hogy a diák nem ragadja meg a páciens élményeinek és érzelmeinek személyes jelentését. A tények szintjén marad, és az orvos-beteg beszélgetésekből jól ismert élményszerű-érzelmi jelentéseket figyelembe nem vevő, detematizáló stratégiát alkalmazza.

Egy vizsgálatban Deppermann és Lucius-Hoene (2005) traumaelbeszéléseket elemeztek. Vizsgálták az elbeszélés szerkezetét, dinamikáját, a beszélő nyelvi és kommunikációs stratégiáit, valamint a prozódiai, vokális, nonverbális kifejezésjegyeket. A GAT (Gesprächsanalytisches Transkriptionssystem) módszerét (Selting és mtsai 1998, ld. keretes rész) használták a felvételek írásbeli lejegyzésére. A vizsgálat anyagát olyan televízióban, rádióban elhangzott elbeszélések adták, amelyek 8-16 évvel korábbi eseményeket írtak le, egyik esetben például egy autóbalesetet. Az interjúk eredeti célja az élettörténet és nem a trauma elmesélése volt. Az érzelmi érintettséget több vonás is kifejezte. A beszélők explicit módon kiemelték, hogy a trauma még a jelenben is fontos része az életüknek, a mesélés alatt megfogalmazták átélt érzelmeiket (pl. *nagyon szégyellem magam emiatt*). A nonverbális eszközökkel kifejezett érzelmeiket két csoportra lehetett osztani: félelem és szorongás, valamint szomorúság és reménytelenség érzése. A félelemteli élmény leírása közben a beszélők hangja magasabb és hangosabb lett, gyorsult a beszédtempójuk, jellemző volt a nyugtalan vagy elakadó légzés az ún. levegős beszédmód, a mondathatárokat nem követő intonáció. A szomorúság és reménytelenség érzésének kifejezésekor az előbbiekkal ellentétes jelenségeket figyeltek meg a kutatók: mélyebb, halkabb, erőtlenebb, monoton vagy intonáció nélküli beszédet, sóhajtozásokat, a hang kontrollálásának problémáját (pl. sírás). Általánosságban, mindkét érzelemfajta kifejezésekor jellemző volt a gyakori nyelés, a

köhögés, krákogás, a nehéz ki- vagy belégzés. Gyakoribbá váltak a szünetek³, – különösen a hosszú (egy másodpercnél hosszabb) szünetek, akár egy szintaktikai egységen belül is. Gyakori volt még a megfogalmazás megszakítása, amely folytatódhatott újrafogalmazással, tartalmi javítással, de volt, hogy félbeszakadt a gondolatmenet. Nonverbális jegyek közül sokszor megjelent az ajkak összepréselése, harapdálása, a vállrángatás, a fejrázás. Az is előfordult, hogy a személyek hosszabb ideig lehunyva tartották szemüket, valamilyen más irányba tekintettek vagy eltakarták kezükkel az arcukat.

Az érzelmek és az érzelmeteli események verbális megjelenése nincs teljesen összhangban: a beszéd nonverbális jegyeinek megváltozása mindig megelőzi a tematizált, szavakkal kifejezett érzelmi élményt. Bár a folyamat maga egyáltalán nem tudatos, és valószínűleg részben spontán arousalfolyamatok befolyásolják, azt a célt szolgálja, hogy a beszélő ezáltal jelezze a hallgató számára, hogy valamilyen érzelmeteli élmény elmesélése következik. Másrészt az érzelmekkel telített esemény elbeszélése után is észlelhetők még a megváltozott vokális és nem nyelvi jegyek. Tehát bár a téma már más, a hangban még mindig megvannak az idegesség, lehangoltság stb. jegyei, ami valószínűleg a levezetést szolgálja az elbeszélő számára. Ez a jelenség is jól mutatja a beszélgető felek egymásrahangelődésének és a kommunikáció koordinációjának bonyolult folyamatát.

Az itt bemutatott tanulmányok olyan helyzeteket elemeztek, amelyek mindegyikében jelen vannak az érzelmek. Az egyes helyzetek különbözhetnek egymástól (csoportos ülés, orvos-beteg interakció, monologikus szövegek, narrációk), de a társalgáselemzés módszerével sokféle beszélgetéstípus és sokféle jelenség (mondatok típusa, témaváltások módja, szünetkezelés stb.) elemezhető (ld. Wodak, 1981).

³ Goldman-Eisler (1968) vizsgálatai alapján arra a következtetésre jutott, hogy a hezitációs szünetek a kognitív folyamatokkal, a beszédbeli tervezéssel vannak kapcsolatban, és a kognitív terhelés növekedésével nő a szünetek száma. Bár egyes érzelmekkel kapcsolatos vizsgálatok kimutatták az érzelmeknek a szünetekre gyakorolt hatását, tulajdonképpen a mai napig nincs pontos ismeretünk arról, hogy az érzelmi és a kognitív folyamatok hogyan befolyásolják a beszédtervezés és -kivitelezés folyamatát.

2. A szavak jelentésének szerepe az érzelemkifejezésben

2.1. A szavak érzelmi jelentése

Az egyes szavak érzelmi jelentésének első vizsgálata Osgood és munkatársai (1957) nevéhez fűződik. Osgoodék alapfeltevése az volt, hogy minden fogalom értékelhető, leírható előre meghatározott mellékpárok segítségével, amelyek három dimenzióba: az érték, az erő és az aktivitás dimenziójába tartoznak. Tehát a három dimenzió mindegyikét melléknévpárok írják le, így az érték dimenziót például a jó-rossz, az erő dimenziót a kemény-puha, az aktivitás dimenziót a meleg-hideg párok. Az Osgood által kidolgozott módszer alapján a megítélők feladata az, hogy a melléknévpárokból álló hétfokú skálákon értékeljenek bizonyos fogalmakat. Magyarországon úttörőként először Putnoky (1978) kezdte el alkalmazni a szemantikus differenciál módszerét. Konkrét, közepesen konkrét és elvont szavakat osztályoztatott hétfokú skálákon, valamint értékeltette ezeket képkiváltó, hangkiváltó és mozgáskiváltó hatásuk szempontjából is. Ezzel egy időben Pléh és Czigler (1979) politikai kifejezések megítélésére dolgozott ki szemantikus differenciált, amely lehetővé tette az ideológiamentes és objektív megítélést.

2.2. A narratív pszichológiai tartalomelemzés

A tartalomelemzés módszerét Gottschalk és Gleser (1969, ld. Karczag, 1988) dolgozta ki az érzelemkifejezés egyéni vizsgálatára epilepsziás, majd pszichotikus betegek beszédének elemzése alapján, később kiterjesztve a vizsgálatot a szuicid, a szorongásos és az agresszív jegyekre is.

A narratív pszichológiai tartalomelemzés alapfeltevése, hogy az elmondott vagy leírt történetek a személy belső állapotaira utalnak, az a mód, ahogyan az emberek magukról beszélnek, azok a szavak, amelyeket használnak, árulkodnak mindarról, amit a személy gondol vagy érez, a szövegek elemzésével pedig közelebb juthatunk a személy (vagy a nemzet) lelki tartalmaihoz. A tartalomelemzés az elbeszélés nyelvi alakzataiból pszichológiai jelenségekre mutat rá. Bizonyos szó kategóriák – például pozitív érzelmi töltetű melléknevek

vagy névmások – előfordulását, szókategóriák együttjárását, szókategóriák és életesemények együttes előfordulását vizsgálja egy adott korpuszban, amely lehetővé teszi például gondolkodásmódok, érzések, egyes tünetegyüttesek jelenlétének, betegségből való gyógyulás bekövetkeztének, bizonyos életesemények feldolgozottságának mértékének meghatározását.

Kezdetben az elemzés kizárólag a szavak szintjén folyt. A szövegben előforduló szavakat két szempontból osztályozták: pszichológiai és nyelvtani szempontból. A pszichológiai szempontú osztályzás során megkülönböztették a fogalmakat és az absztrakciókat, a környezeti objektumokra utaló szavakat, ezen belül az énré és a nem-énre vonatkozó szavakat (más emberek, állatok, növények, tárgyak), valamint a mozgásokra, a motivációkra, az észlelésre és a gondolkodásra vonatkozó kifejezéseket. A másik csoportosítás alapján pedig nyelvtani kategóriákba sorolták a szavakat (igék, főnevek, névmások, határozószók stb.) (ld. Gottschalk és mtsai, 1969).

A 20. század utolsó évtizedeiben olyan számítógépes programokat kezdtek el kifejleszteni, amelyek alkalmasak a tartalomelemzésre (pl. az Atlas.ti). Ezek a programok lehetővé teszik a szavak automatikus valamint manuális kódolását is. Az automatikus kódolás során a program megkeresi és kijelöli az általunk megadott szót vagy szókategóriát az összes kijelölt szövegben, a manuális kódolás pedig lehetővé teszi a szavak, szószerkezetek, esetleg hosszabb szakaszok egyenkénti felcímkézését, kategóriákba sorolását (pl.: „pozitív érzelm”). A kutató egyszerre látja a kezelőfelületen az eredeti szöveget, a kódokat és a kódokhoz írt megjegyzéseit, a program pedig egy gombnyomásra kilistázza az egy kategóriába tartozó elemeket, kiszámolja ezek előfordulási gyakoriságát és az együttjárásokat (a tartalomelemzésről bővebben ld. Ehmann, 2002).

Hasonlóan számítógépesített tartalomelemzési módszert kínálnak azok a programok, amelyek egy vagy több előre meghatározott szógyűjtemény szavait keresik egy adott szöveggörpuszban. A programok különböznek abból a szempontból, hogy a szótárjaik milyen típusú szavakat tartalmaznak. A TAS/C (Mergenthaler, 1996) például két kategóriára fókuszál: az érzelmi töltetre és az absztraktságra. Így az egyik szógyűjtemény mintegy 200 érzelmi töltetű szót tartalmaz, míg a másik olyan angol képzőket, mint az *-ity*, a *-ness* vagy az *-ing*, amelyek az absztraktság jelölői. A DICTION (Hart, 2001) elsősorban politikai beszédek elemzésére készült, és az aktivitás, az optimizmus, a biztosság, a valóságosság és a hasonlóság dimenziókhöz tartozó szavakat elemzi. A LIWC (Pennebaker és mtsai 2001) eredetileg azért készült, hogy segítségével meghatározhatóvá váljon, hogy a negatív életesemények leírása

(önmagunkból „kiírása”) milyen módon járul hozzá a gyógyuláshoz. Az elemzés számos kategóriát érint, amelyek között vannak általános nyelvészeti kategóriák (névmások – egyes vagy többes számúak, névelők, prepozíciók), pszichológiai folyamatokra utaló szavak (érzelmeket vagy kognitív folyamatokat jelölő szavak, pl. műveltető igék), a viszonylagossággal kapcsolatos szavak (igeidő, mozgás, tér) vagy hagyományos tartalmi kategóriák (otthon, foglalkozás).

A Gottschalk és Gleser által kifejlesztett szempontokat a legkülönbözőbb területeken felhasználják, többek között a pszichoterápia-kutatásban, az álomelemzésekben. Több kutatás alapján azt mondhatjuk, hogy a tartalomelemzés alkalmas bizonyos pszichiátriai megbetegedések, például a skizofrénia, a depresszió, a paranoia vagy a szomatizációs zavarok elkülönítésére is (áttekintést ld. László, 2005a). A depresszióra például jellemző, hogy a személy egyes szám első személyű névmásokat (én, engem) sokkal gyakrabban használ, mint a második vagy a harmadik személyű névmásokat, ami a kutatók szerint a másokkal való kapcsolatteremtés nehézségére utal (Bucci és Freedman, 1981; Weintraub, 1981). Más érzelmi zavarral rendelkező csoportoknál, mániában és öngyilkosságot elkövetőknél is megfigyelték ugyanígy az egyes szám első személy előtérbe helyezését (Stirman és Pennebaker, 2001; Lorenz és Cobb, 1952). Ugyanakkor ennek fordítottja, az „én” elhagyása, és a „mi” megnövekedett használata jellemzi azokat, akik valamilyen nagy traumán vannak túl (Stone és Pennebaker, 2002). Számos kutatás bizonyította, hogy ha a személyek írnak arról a negatív érzelmi állapotról, amiben vannak, akkor nő a pszichológiai és a fizikai jóllétük. Azoknak a személyeknek az írásaiban, akiknél javulás következett be, megfigyelték, hogy napról napra több lett a pozitív érzelmekkel és a kognitív folyamatokkal (belátás, ok-okozat) kapcsolatos szó (Pennebaker és Francis, 1996), illetve a névmás (Campbell és Pennebaker, 2002).

Élettörténeti elbeszélésekben mutatták ki, hogy a téri viszonyok, a közeledés-távolodás dinamikája az elbeszélő és a vele kapcsolatban lévő személyek közötti kapcsolat- és érzelemszabályozásra utal (Frye, 1998). Pohárnok (2004) borderline szindrómával diagnosztizált nők élettörténetét hasonlította össze depresszióban szenvedő nők élettörténetével. Alapfeltevése az volt, hogy az elbeszélésekben megjelenő közelítésre illetve távolodásra vonatkozó kifejezések (pl. *anyához mindig odabújtam*) mintázata rámutat a személy kapcsolati és érzelmi működésére. A közös teret, illetve a közeledés-távolodás dimenzióit a középpontba állítva a tartalomelemzés módszerével végzett vizsgálat eredményei a szerzők szerint jól összeegyeztethetők korábbi elméletek állításaival: a borderline zavart

mutató betegek számára az anyától való leválás krízis volt, amely során a távolodás a végleges elhagyatással, míg a közeledés az önálló identitás elvesztésével fenyegetett. A depressziósokhoz képest a közeledések-távolodások abszolút számában nem volt különbség, azonban a borderline személyek beszámolóiban jóval több közeledés-távolodás váltás volt megfigyelhető, ami jól mutatja a borderline személyiségzavarra jellemző ingadozást.

Pólya és munkatársai (Pólya, 2004, Pólya és mtsai, 2005, Pólya és mtsai, 2007) kutatásainak középpontjában az elbeszélői perspektíva és az események érzelmi feldolgozásának mélysége állt. Pólya (2004) a múltban történt érzelmetli események leírását kérte homoszexuális férfiktól illetve lombikbébi programban résztvevő nőktől. A személyeknek azt az eseményt kellett elmesélniük a vizsgálatban, amikor elmondták családjuknak, hogy homoszexuálisok, illetve a nők esetében, amikor megtudták, hogy természetes úton nem lehet gyerekük. Pólya és munkatársai (2007) a vizsgálati személyeket olyan múltbeli események felidézésére kérték, amikor valamit megvalósítottak, veszteséget vagy félelmet éltek át, illetve két olyan esemény elmesélésére kérték őket, amikor pozitív illetve negatív érzelmeket éltek át számukra fontos személyekkel kapcsolatban. Az érzelmi feldolgozottságot, az érzelmek szabályozásának képességét különböző kérdőívek segítségével mérték. Az eredmények mindkét kutatás esetében azt mutatták, hogy a visszatekintő perspektíva (pl. múlt idejű igék, múltra vagy távolra utaló határozószók) használata az események feldolgozottságát és a kiegyensúlyozott identitásállapotot jelzi, míg az átélő elbeszélő perspektíva (pl. jelen idejű igék, jelenre és közelre mutató határozószók) érzelmi kiegyensúlyozatlanságra utal (példákat lásd az

1. táblázatban).

Nyelvi jegy	Visszatekintő perspektívaforma	Átélő perspektívaforma
Idői rámutatás Igeidő Idői határozószavak	Múlt Pl. akkor	Jelen Pl. most
Helyi rámutatás Helyhatározó szavak Mutató névmások	Pl. ott Pl. az	Pl. itt Pl. ez
Személy rámutatás Személyes névmások	Én: elbeszélő Te: hallgató	Én: szereplő Te: szereplő
Specifikus kifejezések	Dátumok, pl. 1990 nyár	Indulatszavak, pl. hűha, hopp
Mondat módja	Csak kijelentő	Nincs megkötés

1. táblázat Példák a visszatekintő és az átélő perspektíva nyelvi megjelenésére (forrás: Pólya és mtsai, 2005)

Más típusú vizsgálatok a magyar történelem fontos eseményeit és ezek érzelmi vonatkozásait elemzik történelemkönyvek vagy történelmi regények segítségével. László (2005b) kutatásában általános iskolás, középiskolás és egyetemista diákokat kért meg egy pozitív és egy negatív esemény elbeszélésére a magyar történelemből. Az elbeszélések alapján kirajzolódik, hogy a személyek a magyar történelemre úgy tekintenek, mint amely sok szép és pozitív eseménnyel kezdődött, amiket aztán több százados balsors követett, ahol a negatív események közül a '48-as forradalom emelkedik ki pozitívként. A rendszerváltás utáni történelem megítélése enyhén pozitív. A pozitívnek címkézett események két fontos téma köré csoportosulnak: a békében élés és a szabadságért való harc, míg a negatív események az életek elvesztése, az értékek elvesztése, a földterület elvesztése és az idegen uralom köré.

Fülöp Éva kutatásában (2008) arra volt kíváncsi, hogy magyar történelmi regényekben és történelemkönyvekben hogyan jelennek meg az érzelmek. Különbségeket leginkább a történelemkönyvek szövegei alapján talált. Az eredmények szerint több az érzelemtulajdonítás a saját csoportra, mint a más nemzetre vonatkozóan, ami az empátia jelenlétére utal a saját nemzetre vonatkozóan. A történelemkönyvekben az öröm gyakrabban fordul elő külső csoportnál, mint sajátnál, míg a félelem a saját csoportra volt a jellemzőbb. A leggyakoribb érzelm a magyarokra vonatkoztatva a remény érzése volt. Az is megfigyelhető volt, hogy a történelemkönyvekben a szerzők a magyarokra több távolító elemet használtak, mint más nemzetekre, a magyarok tehát hajlamosabbak érzelmi távolságot tartani másoktól, amely távolságtartás egyben növelheti a szelfkategorizációt.

A tartalomelemzés nemcsak egy bizonyos érzelm jelenlétére, hanem annak nagyságára, mélységére is rámutathat. Ehhez figyelembe veszik a meghatározott tartalmi kategória előfordulásának gyakoriságát (minél gyakrabban fordul elő egy kategória, annál erősebb az érzelm), a közvetlenség, a direktség mértékét (a *bárcsak meghalna* kifejezés sokkal direktebb, mint a *nem szeretem* kifejezés), valamint a személyes érintettség mértékét, amivel a személy a gondolatait, cselekvéseit, érzéseit leírja (pl. *A háborúban féltem, hogy meghalok.* vagy *A háborúban emberek halnak meg.*) (Battacchi és mtsai, 1996, 107–113).

A tartalomelemzés tehát azt vizsgálja, hogy a beszédben, a történetben közvetlenül megfigyelhető nyelvi, lexikális tartalmi jegyek hogyan használhatók fel pszichés állapotok meghatározására. Kiindulási pontja, hogy beszédünkben mi magunk is benne vagyunk: az a mód, ahogyan mondanivalónkat megfogalmazzuk, azok a szavak és kifejezések, amelyeket

használunk, mind-mind rólunk árulnak el valami fontosat, utalhatnak nemünkre, társadalmi helyzetünkre, aktuális lelkiállapotunkra, személyiségünkre, attitűdjeinkre.

3. Vokális érzelm kifejezés

A vokális érzelm kifejezések a beszéd akusztikai mintázatában megjelenő, érzelmi állapotokra utaló jelzések. Általában az, amit mondunk, és az, ahogyan az üzenetet mondjuk, összhangban van. Ha boldogok vagyunk, mert egy rég nem látott ismerőssel találkozunk, és felkiáltunk, hogy „*de örülök, hogy látlak!*”, nem csak szavaink (*örülök*) fejezik ki érzelmi állapotunkat, hanem a hangsúly, a hanglejtés, a beszédtempó is mind jelzi, mit érzünk.

Ugyanakkor előfordulhat az is, hogy a kétféle üzenet elválik egymástól: amit mondunk és az, ahogy azt mondjuk, nem egyezik egymással, inkongruens. Például ha leplezni szeretnénk izgalmunkat, és azt mondjuk, hogy „*nyugodt vagyok*”, ugyanakkor hangunk remeg, szánk kiszárad, ezért berekedünk, a szavakkal kifejezett érzelmi állapot (nyugodtság) és a vokális jegyek (idegesség) egymásnak ellentmondanak. Ilyen esetben egy jó megfigyelő inkább az utóbbi, nonverbális jegyeknek fog hinni.

A vokálisérzelem-kutatás területén belül az egyik elképzelés abból az alapfeltevésből indul ki, hogy meghatározott számú alapérzelem létezik (Ekman, 1999, Izard, 1993). Eszerint az ún. „diszkrét érzelmek” megközelítés szerint – ahogy az arckifejezések esetében: egy-egy érzelmehez egy-egy specifikus arckifejezés tartozik – úgy a hang esetében is: egy-egy alapérzelemnek egy-egy akusztikus mintázat felel meg (ld. Juslin és Scherer, 2005). Tehát míg az öröme például a hangmagasság és a hangerő, illetve ezek variabilitásának megemelkedése, a beszédtempó gyorsulása, a szünetek rövidülése jellemző, addig a szomorúságra pontosan ezek ellenkezője igaz.

A másik elképzelés szerint az érzelmek és az őket jellemző beszédparaméterek dimenziók mentén helyezhetők el (ld. Juslin és Scherer, 2005; Laukka, 2004). A főbb összefüggéseket az arousallel kapcsolatban mutatták ki (pl. Breitenstein és mtsai, 2001; Laukka és mtsai, 2005): ha nő az aktivációs szint (pl. a szorongás esetében), ezzel együtt nő az alaphangmagasság, az alaphangmagasság-variabilitás, gyorsul a beszéd, rövidebbek lesznek a szünetek és nő a hangerő. A kellemesség-kellemetlenség (valence) dimenzióval kapcsolatos eredmények már

sokkal inkonzisztensebbek. Néhány kutatás kimutatta, hogy a pozitív érzelmek esetében csökken a hangmagasság, nő a hangmagasság-variancia, gyorsabb és halkabb lesz a beszéd, rövidülnek a szünetek (Laukka és mtsai, 2005). Ugyanakkor más kutatásoknak nem sikerült ezeket az eredményeket kimutatniuk (pl. Pereira, 2000). A legkevesebb kutatás a befolyásolhatóság/kontrollálhatóság (potency) dimenzióval kapcsolatban készült⁴, és az eredmények itt a legkevésbé egybehangzóak (ld. Apple és mtsai, 1979; Laukka és mtsai, 2005). A befolyásolhatóság növekedésével (pl. utálat) nő a hangmagasság variabilitása, a hangerő, a hangerő variabilitása és a magas frekvenciákon az energia. Ugyanakkor a hangmagasság és a beszédtempó esetében egymásnak ellentmondó eredményeket kaptak az egyes kutatások. Juslin és Scherer (2005) amellett érvelnek, hogy – bár bizonyos beszédparaméterek összekapcsolhatók egy-egy dimenzióval – minden érzelmet egyedi akusztikai mintázat jellemez, amelyet nem csak a kiváltó helyzet, hanem számtalan egyéb tényező (a helyzet értékelése, a beszélő neme, a szituáció stb.) is meghatároz.

3.1.A vokális érzelm kifejezés kutatásának módszerei

A kutatásokat módszertanilag három nagy csoportba sorolhatjuk: a természetes helyzetekben előforduló érzelmeket felhasználó (3.1.1), az érzelmi állapotokat kiváltó (3.1.2) és az eljátszott érzelmeket vizsgáló kutatások (3.1.3). Az itt felsorolt módszereket bizonyos korlátozásokkal vagy kiegészítésekkel más típusú (pl. szociálpszichológiai) kutatásokban is alkalmazzák. A vokális érzelm kifejezés kutatásakor a legfontosabb szempont, hogy keletkezzen olyan hanganyag, amely jól vizsgálható és jól összehasonlítható, más szempontokat, például az érzelmekkel való megküzdést vagy a beszélő arckifejezéseit általában nem veszik bele az elemzésbe és az értelmezésbe.

A leggyakrabban megfigyelt beszédparaméterek minden esetben az F0 vagy alaphangmagasság (Hz), az intenzitás vagy hangerő (dB), ezek variabilitása, valamint az idői jellemzők, úgymint szünetek (ezredmásodperc), beszédtempó (hang/másodperc, nagyobb

⁴ A legújabb érzelmkutatások eredményei alapján még egy további dimenziót (bejósolhatóság, unpredictability) feltételezhetünk (Fontaine és mtsai, 2007).

egységek esetében szó/perc). Ezek mérésére a leggyakrabban használt és legnépszerűbb szoftver a Praat (Boersma és Weenick, 2009), amely ingyenesen letölthető (www.praat.org).

3.1.1. Természetes helyzetekben előforduló vagy valós érzelmek

Az ezen a területen készült kutatások az érzelmeket természetes helyzetekben vizsgálják, illetve feltételezik, hogy a beszédhelyzet alatt, amit megfigyelnek, a beszélő valamilyen valós érzelmi állapotban van.

Roessler és Lester (1976) vagy Frolov és munkatársai (1999) pszichoterápiás ülések hanganyagát elemezték. Johannes és munkatársai (2000) asztronauták kiképzésén, egy veszélyes és izgalmas feladat végrehajtása közben készült hanganyagokat vizsgálták. Jól használhatók még talk showkban elhangzott beszélgetések, riportok, interjúk, amelyek valamilyen érzelmeteli eseménnyel kapcsolatban, például természeti katasztrófa során vagy egy sportesemény alatt készültek (áttekintés ld. Scherer 2003).

Ezeknek a kutatásoknak magas az ökológiai validitása, hiszen az egyes érzelmeket azok természetes megjelenésekor próbálják megragadni, és ez nagy előnyt jelent más módszerekkel szemben. Módszertanilag azonban problematikusak, hiszen általában csak egy, vagy nagyon kevés beszélő hangját vizsgálják, és a körülmények miatt a felvétel minősége sem mindig tökéletes, például nagy a nehezen kiszűrhető háttérzaj. Ezen kívül gyakran nehéz meghatározni, hogy a beszélő pontosan milyen érzelmi állapotban volt, vagy hogy mi váltotta ki nála az érzelmet. Mivel ezek a kutatások általában már meglévő, kész interjúkat, beszédeket elemeznek, ezért arra sincs lehetőség, hogy a beszélgetés végén a kutató megtudja, amire még kíváncsi. A különböző személyek és helyzetek összehasonlítása szintén problematikus. Ennél a módszernél szembesülünk a természetes helyzet és a reprezentativitás közötti feszültséggel. Tudjuk, hogy a helyzet, amelyet elemzünk, nem áll a mi befolyásunk alatt, nem a hangelemzés céljából állt elő, hogy minden helyzet és minden beszélő nagyon különbözhet egymástól, és így nehéz általános, mindenkire igaz következtetéseket levonni. Cserébe azonban a létező legtermészetesebb érzelmeket figyelhetjük meg benne. (A természetes helyzetekben kifejezett érzelmeket tartalmazó adatbázisokról ld. Douglas-Cowie és munkatársai (2003).

3.1.2. Kiváltott érzelmek

Az érzelmek vokális megjelenése úgy is vizsgálható, hogy bizonyos érzelmi állapotokat kísérletileg váltunk ki, és ebben az állapotban rögzítjük a vizsgálati személyek beszédét.

A Velten-módszert (Velten, 1968) kifejezetten abból a célból fejlesztették ki, hogy az érzelmeknek a beszédre gyakorolt hatását vizsgálják a segítségével. A kísérleti személyek 60 állítást kapnak, amelyek egyes szám első személyben vannak megfogalmazva, és egy adott érzelmi állapottal kapcsolatosak (ld. 2. táblázat). A résztvevőket arra kérik, hogy olvassák fel ezeket az állításokat, és próbálják meg magukat beleélni annak a személynek az állapotába, aki ezeket a sorokat írhatta. Bár maga a módszer rendkívül egyszerű, mégis hatékonyan bizonyult mind a boldogság, mind a lehangoltság, mind a harag kiváltására (ld. Cowie és Cornelius, 2003).

Öröm:

- Szinte mindenben megtalálom a jót.
- Lelkesnek és bizakodónak érzem magam.

Lehangoltság:

- Annyira egyedül éreztem magam, hogy sírni tudtam volna.
- Túl sok rossz dolog történik az életemben.

2. táblázat Példák a Velten-féle hangulatkiváltó módszer mondataiból

Az autobiografikus emlékezés módszerét Brewer és munkatársai (1980) fejlesztették ki. A vizsgálatban arra kérik a személyeket, hogy idézzenek fel egy érzelmmel teli eseményt a saját életükből, például egy olyat, amikor egyedül érezték magukat, vagy amikor nagyon boldogok vagy épp frusztráltak voltak.

Gyakran alkalmaznak zenét érzelmek kiváltására (ld. Westermann és mtsai, 1996). A zenét használó kísérletek legfőbb tapasztalata, hogy a negatív érzelmek kiváltása sikereesebb, mint a pozitívaké, és mind pozitív mind negatív érzelmek esetében a zene által kiváltott érzelem

minden esetben kisebb mértékű, mint a film által kiváltotté. A hatás akkor a legnagyobb, ha nemcsak lejátsszák a zenét, hanem valamilyen instrukcióval meg is kéri a kísérleti személyeket, hogy próbáljanak meg az adott érzelmi állapotba kerülni (a zene érzelemkiváltó hatásáról lásd Scherer és Zentner, 2001).

A zene és az önéletrajzi emlékezés módszereinek kombinációját használta egy nemrégiben végzett magyar kutatás is (Szabó, 2008). A kísérleti személyeket arra kérték, hogy idézzenek fel magukban velük megtörtént vidám illetve szomorú eseményeket, és amíg átgondolták a történetet, vidám illetve szomorú zenét hallgattak. A semleges beszédmintákat egy átlagos nap elmesélése adta. Korábbi kutatásokkal ellentétben nem csak egy-egy mondatról, hanem félperces szakaszokról tett megállapításokat a szerző, és az egy személyhez tartozó szomorú, vidám és semleges beszédrészletek különböző beszédjegyeit hasonlította össze. Az eredmények azt mutatták, hogy a mért paraméterek alapján különbség van a szomorú, a semleges és a vidám érzelmi állapot között a hosszabb, félperces szakaszokon is. A legnagyobb különbségek a szomorú és a vidám érzelmek között voltak megfigyelhetők: szomorú hangulatban lassabban artikuláltak a személyek, hosszabb szüneteket tartottak, nagyobb volt beszédükben a szünetek aránya a teljes beszédidőhöz viszonyítva, és halkabb volt a hangjuk, mint vidám érzelmi állapotban.

Cohen és munkatársai (2010) az IAPS-t (International Affective Picture System, Lang és mtsai. 2005) képeit használták érzelem kiváltására, azonban két kísérletük eredménye azt mutatta, hogy bár az önbeszámolók alapján változás történt a személyek érzelmi állapotában, a prozódiai jegyek nem mutattak változást. A szerzők a harmadik kísérlet tanulsága alapján az önéletrajzi emlékezést javasolják érzelemkiváltó módszernek.

Számos módja létezik még az érzelemkiváltásnak (ld. Banos és mtsai. 2003, Cowie és Cornelius, 2003, Scherer, 2003, Oatley és Jenkins, 2001, 56. o.). Előfordul, hogy nyugtató hatású szert adnak a kísérleti személyeknek, vagy filmeket, hipnózist, imaginációt, szerepjátékokat, virtuális valóságot, számítógépes játékokat alkalmaznak. Más pszichológiai területekről is számos egyéb módszer ismert: lehetséges negatív érzelmek kiváltása például kellemetlen szagokkal vagy akár elektrosokkal. A kísérleti személyeknek esetleg mások előtt kell szerepelniük, vagy nehéz feladatot kell elvégezniük rövid idő alatt, esetleg felbosszantja őket a kísérletvezető. Pozitív érzelem kiváltásának a módszerei közé tartozik például, amikor a kísérleti személyek ajándékot kapnak, vagy megdicséri őket a kísérletvezető (lásd ezekről Forgács, 2003).

A kiváltott érzelmek módszerét azért részesítik előnyben, mert a kontroll mértéke megfelelően nagy. Kísérletünkhöz akár előre kiválaszthatunk bizonyos személyiségű embereket, és minden résztvevő ugyanazon procedúrán megy keresztül. Gyakori, hogy a személyeknek a kiváltott érzelmi állapotban fel kell olvasniuk egy szöveget, vagy el kell mondaniuk egy célmondatot, ezáltal a kapott hanganyagok is nagyon jól összehasonlíthatóvá válnak.

A módszer hátránya maga a kísérleti helyzet. A kísérlet következtében általában gyengébb érzelm (inkább hangulat) jön létre, mint a természetes helyzetekben történő érzelmkifejezés esetében, hiszen az emberek nem szívesen adják át magukat érzelmeiknek egy számukra ismeretlen személy előtt. Az is gyakori, hogy bár ugyanazokat a feladatokat végzi el minden kísérleti személyünk, nem ugyanolyan hatást érzünk el náluk, hiszen a kísérleti helyzet másként hathat rájuk, máshogy értékelhetik a szituációt. Az elrendezés azonban például lehetővé teszi ennek korrigálását is: utólag megkérdezhetjük a résztvevőktől, hogy milyen érzelmet, érzelmeket éltek át, és mennyire érezték erősnek ezeket.

3.1.3. Eljátszott érzelmek

Az érzelmek beszédben való megjelenésének talán legkedveltebb módja az, amikor hivatásos vagy laikus színészeket kérnek meg arra, hogy játszanak el bizonyos érzelmeket, gyakran egy előre meghatározott semleges tartalmú, esetleg értelmetlen mondattal (hogy kizárják a szavak érzelmi jelentésének hatását). Így az érzelmek egymással és a semleges érzelmi töltettel is jól összehasonlíthatók.

Banse és Scherer (1996) kutatásukban 12 hivatásos színészt, 6 férfit és 6 nőt kértek meg arra, hogy játszanak el érzelmeket, köztük olyan ritkábban kutatottakat is, mint például az érdeklődés, a büszkeség vagy a szégyen. Az érzelmek kiválasztásakor azt is figyelembe vették, hogy ezek ne csak arousal, de kellemesség tekintetében is a dimenziók különböző pontjain helyezkedjenek el, valamint egy érzelmcsaládba tartozó, egymáshoz hasonló érzelmeket is belevettek az elemzésbe. Így összesen 14 érzelmet választottak ki: hűvös harag és forró düh, szorongásos félelem és pánikszerű félelem, depresszív szomorúság és kétségbeesés, boldogság és túláradó öröm, érdeklődés és unalom, szégyen és büszkeség, undor és megvetés. Célmondatként két értelmetlen, külföldinek hangzó mondatot használtak.

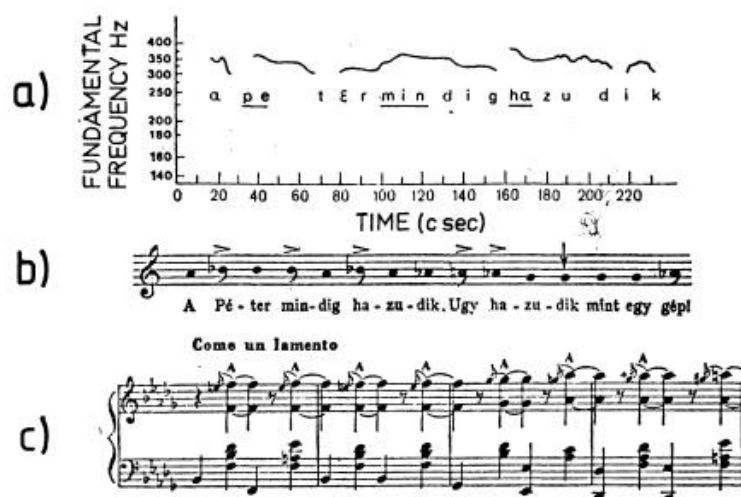
A végső elemzésbe 280 mondatot vontak be. Főbb eredményeik szerint az intenzívebb, magas arousalas érzelmek – például a kétségbeesés, a forró düh vagy a pánikszzerű félelem – esetében nőtt az alaphangmagasság és a hangerő, míg a kevésbé intenzívek – például az unalom vagy a megvetés – esetében csökkent ennek a két paraméternek az értéke. A beszédtempó növekedése a kétségbeesés kivételével szintén az arousalal volt összefüggésben.

Bänzinger és Scherer (2005) ez előbbi kutatás 9 színészének mondatait választotta ki, az érdeklődésük középpontjában pedig az alaphangmagasság vizsgálata állt. Részletesen elemezték a mondatok átlagos alaphangmagasságán kívül az intonációs mintázatot, a hangmagasság kontúrját is. Eredményeik azt mutatják, hogy az érzellemmel együtt járó arousal egyértelműen meghatározza a közlés alaphangmagasságát: a magasabb arousal szint magasabb alaphangmagassággal jár együtt. Ugyanakkor a hangmagasság kontúrja, az, hogy a mondat intonációja szökő, emelkedő, lebegő vagy ereszkedő, érzelmenként változik, és nem kapcsolható egy dimenzióhoz sem. A közlés második hangsúlyos helyének hangmagassága például nem függött össze az arousalal: a magas arousalas forró düh vagy az áradó öröm és az alacsony arousalas hideg harag esetében is nagy volt a második hang kitérése az első hanghoz képest, míg a szomorúság vagy a boldogság esetében kicsi volt a különbség.

3.1.4. Magyar kutatások

A magyar érzelmes beszéd vokális jegyeinek vizsgálatával kevés kutatás foglalkozott, ezek nagy része Fónagy Iván nevéhez köthető (ld. Fónagy, 1981; 2001). Az első ilyen – mai napig a külföldi irodalomban is sokat idézett – kutatás Fónagy és Magdics (1967) nevéhez fűződik, akik a magyar beszéd dallamát elemezték. Vizsgálati módszerük meglehetősen aprólékos munkát igényelt. 82 különböző érzelmet, illetve attitűdöt kottáztak le úgy, hogy egy-egy szótaghoz egy-egy hangjegyet rendeltek hozzá (a 2. ábra mutat erre egy példát). Legtöbb esetben a megnyilatkozások tartalmilag is érzelmet fejeztek ki. A vizsgált anyagot részben magnetofonra vett spontán beszélgetések, részben színészek által eljátszott megnyilatkozások adták. Egy mondat egy adott érzelmet hordozott. Egyrészt tehát lekottázták az érzelmet, attitűdöt kifejező mondatot, másrészt szöveggel, zenei fogalmakkal is leírták a mondat intonációját és hangszínezetét.

A vágyakozásnak megfelelő intonációt például a következőképpen írják le: „az első két vagy több szótag felgyorsul. Az első és általában hangsúlyos szótag hangmagassága és hangereje lassan, kis mértékben emelkedik, amit ezen paraméterek fokozatos csökkenése követ. Az intenzitás a közlés végéig csökken, a dallam enyhe emelkedéssel végződik. A beszédtempó lassú, a beszédmód „legato” (egyenletes, megszakítás nélküli)” (ld. Fónagy, 2001, 104.).



2. ábra „A Péter mindig hazudik, úgy hazudik, mint egy gép” mondat panaszkodó hanglejtésben. Az (a) ábra az alapfrekvencia változását mutatja, a (b) ábra az alapfrekvencia átírását hangjegyekké, a (c) ábra részlet Monteverdi Orfeo c. operájának panaszdalából (Fónagy 1981, 68.)

Bár a dallammenet leírása nagyon részletes, a vizsgálat hátránya, hogy nem ad pontosabb leírást egyéb prozódiai jegyekről, például a hangidőtartamokról, valamint ha szótagonként egy hangjegyet használunk, az nem képes a valódi dallammenet gazdagságának tükrözésére.

Egy másik magyar kutatásban Bolla (1992) a semleges tartalmú *Nyolc óra* kifejezés 28-féleképpen kiejtett változatát elemezte műszerrel. A 28 kifejezést különböző beszédhelyzetekhez kötötte, hogy megkönnyítse a színészek feladatát. A boldog izgatottság például a következőket jelentette: *Boldog felkiáltás, végre elérkezett a várva várt nyolc óra.* Bolla megmérte a közlések időtartamát, tempóját, hangmagasságát, hangterjedelmét, dallammenetét, hangerejét az összes megnyilatkozásra. A munka leíró jellegű, a szerző nem von le semmilyen következtetést, nem végez összehasonlításokat. Az adatok

összehasonlításával azonban megtudhatjuk, hogy pl. a semleges kijelentéshez képest a „boldog izgatottság” állapotában kimondott *Nyolc óra* lassabb beszédtempójú, kisebb hangterjedelmű és halkabb, míg a „lemondás, szomorúság” állapotában kimondott még lassabb, még kisebb hangterjedelmű és még halkabb.

Az eljátszott érzelmek vizsgálatának előnye, hogy nagyon jól kontrollált, stúdióminőségű felvételekhez juthatunk. Előnye még hogy az eljátszott érzelmek általában intenzívebbek és nagyobb hangbeli változásokat mutatnak, mint a kiváltott vagy akár a valós érzelmek, így jól megfigyelhetők, jól mérhetők. Ez ugyanakkor a hátránya is a módszernek, hiszen a színészek rendszerint túljátsszák az érzelmeket, előadásmódjukban az elvárásoknak megfelelő változások felnagyítva jelennek meg. Ugyanakkor az olyan apró hangbeli változások, amelyeket a testünk vegetatív változásai okoznak (a gyorsabb szívverés vagy a száj kiszáradása) a színészek beszédmintáiban általában nincsenek benne, hiszen a vegetatív változások manipulálására a színészek nem képesek. A hallgatók azonban különbséget tudnak tenni például a valódi, átélt mosolygás, illetve a mechanikus mosolygás közben mondott beszédminták között (Aubergé és Lemaître, 2000). Így bár a színészek által eljátszott érzelmek vizsgálata jól kontrollált módszer, és használható eredményeket ad például a szintetizált érzelmes beszéd előállításához, az említettek miatt nem nélkülözhető a természetes helyzetekben megfigyelhető vagy a kiváltott érzelmek vizsgálata sem.

3.2.A produkciós kutatások eredményei

Az alábbiakban több összefoglaló cikk (Pittam és Scherer, 1993; Johnstone és Scherer, 2000; Douglas-Cowie és Cowie, 2002; Scherer, 2003; Banse és Scherer, 1996) alapján áttekintjük, hogy milyen főbb beszédjegyek jellemzik a főbb érzelmeket. A 3. táblázat a három fenti módszerrel végzett kutatások egy irányba mutató eredményeit mutatja. A táblázatban a leggyakrabban vizsgált akusztikus jegyek és érzelmek összefüggéseit foglaltuk össze. Az adatok elsősorban angol, német és francia beszélőkkel végzett kísérletek eredményeire támaszkodnak.

	Harag, Düh	Félelem	Szomorúság	Stressz, szorongás	Öröm
Intenzitás	Növekszik	Növekszik	Csökken	Növekszik	Növekszik
F0 átlag	Növekszik	Növekszik	Csökken	Növekszik	Növekszik
F0 variabilitás	Növekszik		Csökken		Növekszik
Hangterjedelem	Növekszik	Növekszik (csökken Banse és Scherer (1996))	Csökken		Növekszik
Dallamvonal	Ereszkedő	Ereszkedő	Ereszkedő	Ereszkedő	Emelkedő
Beszéd-, artikulációs tempó	Növekszik	Növekszik	Csökken		Növekszik

3. táblázat: Érzelmi állapotok és akusztikus jegyek és összefüggései

Az undor érzéséhez kapcsolódó kutatások eredményei nem konzisztensek. Némely kutatás növekvő F0 átlagot talált, míg mások pont csökkenőt. Egyrészt kevés ilyen kutatás volt, másrészt módszertanilag is nagyon különböztek abból a szempontból, hogy hogyan próbálták meg az érzelmet kiváltani (pl. egy kellemetlen film megtekintésével, vagy színészeket kértek fel az undor eljátszására).

3.2.1. Az érzelmek hatása a kisebb beszédegységekre

Az elmúlt években elindult egy olyan kutatási irány, amely kisebb egységek, egyes beszédhangok, a hangkapcsolatok, rövidebb hangsorok változását vizsgálja az érzelmek függvényében. Bár ezeknek a kutatásoknak az előzményeit megtaláljuk például a magyar kutatásokban is (ld. Fónagy, 1981), a technikai fejlődés jelentősen módosította a vizsgálati módszereket.

A

4. táblázatban Fónagy (1981, 76. o.) kutatási eredményeit láthatjuk. A kutatásban négy magyar színésznő mondta el a /kisera me:ra ba:vataɡ/ jelentés nélküli mondatot 12 attitűdöt (félelem, magyarázkodás, öröm, csodálkozás stb.) eljátszva. A táblázat a közlés mássalhangzó-időtartamhoz képesti relatív magánhangzó-időtartamait (t mgh/msh) mutatja a különböző állapotokban. Az adatok szerint az ún. gyengéd érzelmek vagy attitűdök esetében (panaszkodás, szomorúság, öröm, magyarázkodás) hosszabbak a magánhangzók, mint az ún. támadó, veszekedő érzelmek (gyűlölet, harag, irritáció, félelem) esetében. Az első esetben dallamosabbak és szabályosabbak, egyenletesebbek is a közlések.

panaszkodás	2,06
szomorúság	2,01
öröm	1,85
magyarázkodás	1,80
csodálkozás	1,78
fenyegetés	1,72
nyilatkozat	1,61
félelem	1,59
irritáció	1,58
harag	1,50
gyűlölet	1,35

4. táblázat A magánhangzók relatív időtartama (t mgh/msh) a /kisera me:ra ba:vataɡ/ jelentés nélküli mondatban különböző érzelmi állapotokban, attitűdökben (Fónagy, 1981, 76. nyomán)

Mitchell és munkatársai (2000) szótagok hosszának változását vizsgálták ingerült (és egyre ingerültebb) állapotokban. A kísérleti helyzetben a kísérletvezető sokszor egymás után úgy tesz, mintha félreértené a kísérleti személy által elmondott három számjegyű szám első vagy utolsó jegyét, így a kísérleti személy folyton kijavítja a kísérletvezetőt, megismétli a helyes számot, egyre idegesebbé válva. Az eredmények azt mutatták, hogy a javítás és az ingerültség hatására megváltozik a számok kimondásának ritmusa. Ha az első számjegyet kell kijavítani a kísérleti személynek (KV: *Tehát 559?* KSZ: *Nem, KILENC59!*), akkor nő az első és a második tag között eltelt idő, míg ha az utolsó számjegyet (KV: *Tehát azt mondja 555.* KSZ:

Nem, 55KILENC!), akkor a második és az utolsó szótag közötti idő nő. Minél többször kellett kijavítania a számot a kísérleti személynek, annál hosszabbak lettek a kritikus szótagok.

Egy másik kutatásban (Kienast és Sendlmeier, 2000) színészek által eljátszott mondatok magán- és mássalhangzóit elemezték. Vizsgálták az előre- és a visszaható hasonulásokat, a zöngésedést és a nazálissá válást a képzés helye és módja függvényében, valamint a nyitottság fokát. Elemezték a magánhangzók első és második formánsainak helyeit, valamint bizonyos zöngétlen réshangokat (pl. f, s). A harag kifejezése esetében mérték a legpontosabb artikulációt. A szerzők szerint ennek oka, hogy a harag esetében a beszélők általában megnyújtják a magánhangzókat és jobban hangsúlyozzák a szótagokat, ami pedig jobban artikulált mássalhangzóképzést eredményez. A félelmet, az unalmat és a szomorúságot kifejező mondatokra a pontatlan magánhangzóképzés volt jellemző. A félelem esetében ezt a felgyorsult artikulációs tempó, a magánhangzó-rövidülések magyarázhatják, míg az unalom és a szomorúság esetében a kisebb izomfeszültség. A vidám mondatok esetében mind az első mind a második formáns helyei feljebb tolódtak, ami a magánhangzók rövidülésének, illetve a gége megemelkedésének a következménye. A réshangok képzése szűkebb lett.

Az itt bemutatott kutatások és eredmények igazolják, hogy az intenzitás, a hangmagasság és az idői jegyek alkalmasak az egyes érzelmek vokális megjelenésének leírására. Ugyanakkor, amint a 3. táblázatban is láthatjuk, bizonyos érzelmek gyakran nagyon hasonlítanak a velük együttjáró beszédjegyek tekintetében. Bár a valóságban tudjuk, hogy például az öröm és a harag vokális kifejezése nagyon is különböző, a fenti táblázat minden mutatója egy irányba mutat a két érzelemnél. A hasonlóság oka ebben az esetben az, hogy mindkettő arousalemelkedéssel együttjáró érzelem, a fent felsorolt mutatók pedig – úgy tűnik – ennek változására érzékenyek. Az eredmények finomításához, egy-egy érzelem pontosabb leírásához és más érzelmektől való megkülönböztetéséhez más mutatók mérésére lenne szükség, ami eddig egyelőre csak kevés kutatásban valósult meg.

Az első nagy kérdés, hogy egyáltalán milyen fizikai paraméterek írják le, ha valakinek „rekedtes”, „levegős”, „feszés” vagy „elhaló” a hangja. Ma már léteznek olyan kutatások, amelyek ezeket a finom különbségeket próbálják megragadni. Az artikuláció pontosságára utalhat a második formáns helye és sáv szélessége (Kienast és Sendlmeier, 2000; Drahotová és Mtsai, 2008), a rekedtességre az irreguláris fonáció vagy glottalizáció (Böhm, 2009), a hangszínre az energia frekvenciák szerinti eloszlása (Kienast és Sendlmeier, 2000), a légzési,

hangképzési és artikulációs működés megváltozására a hosszú idejű átlagos spektrum (pl. Banse és Scherer, 1996). Ezen mutatók vizsgálata már a számítógépes nyelvészet irányába mutat (ld. 3.4. fejezet lentebb).

3.3. Percepciós kutatások

A percepciós kutatások arra keresik a választ, hogy miként képesek a hallgatók csupán a beszédhangból, tehát tartalomtól függetlenül megállapítani a beszélők érzelmeit, hogy melyek azok a paraméterek, amelyek a legfontosabbak egy-egy érzelem felismerésében. A legtöbb percepciós kutatásban először arra kérnek hivatásos vagy laikus színészeket, hogy egy bizonyos standard tartalmú vagy értelmetlen szót, kifejezést vagy mondatot adjanak elő több különböző érzelmi töltettel (Johnson és mtsai, 1986; Banse és Scherer, 1996). Hogy megkönnyítsék a színészek dolgát és hozzááruljanak a hiteles előadásmódhoz, a mondatokat gyakran egy-egy szituációba ágyazzák. A végső elemzésbe azonban csak a célmondatot veszik bele, a többi szövegrészt figyelmen kívül hagyják. A színészek által elmondott mondatok listáját random módon játsszák le a vizsgálati személyeknek.

A megítélőknek általában egy megadott listáról kell kiválasztaniuk, hogy melyik érzelmeket hallották, a kutatók pedig azt nézik, hogy a hallgatóknak milyen mértékben sikerült az egyes érzelmeket felismerni. A különböző áttekintések (Scherer és mtsai, 1991; Scherer, 2003) átlagosan 55-65 %-os felismerési arányról számolnak be, ami jóval nagyobb a véletlenszerű találgatásnál.

Ezeknél a kutatásoknál az egyik sarkalatos pont az, hogy hogyan kérdezik meg a személyeket a véleményükről. Az egyik lehetséges módszer, hogy listát adnak a megítélőknek, nekik pedig ki kell választaniuk, hogy éppen melyik érzelemnek megfeleltethető mondatot hallották (kényszerválasztásos módszerek). A lista állhat annyi érzelmmegnevezésből, ahány érzelmeket a kísérlet során bemutatnak, de olykor bővítik ezt a listát, és beletesznek olyan érzelmmegnevezéseket is, amelyeket a bemutatásban nem hallanak a személyek. Más esetben minden bemutatáskor két-két érzelem közül kell egyet kiválasztani. Vannak olyan kutatások is, amelyek nem adnak listát a megítélőknek, hanem arra kéri őket, nevezzék meg az érzelmeket, ami eszükbe jut a kijelentést meghallgatva, és esetleg megkérdezik még tőlük, hogy miért ezt választották, milyen beszédjegyek voltak számukra a döntőek. Az első

esetekben a véletlenszerű találgatáshoz képest érdekesesek a kapott eredmények, a nyitott kérdés esetben pedig az a fontos, hogy a rengeteg választ, amit kaptunk, valahogy rendszerezni tudjuk – például független megítélők vagy korábbi érzelencsoportosításos vizsgálatok alapján.

A produkciós kutatásoknál már említettük, hogy Banse és Scherer (1996) kutatásukban színészeket kértek meg érzelmi töltetű közlések elmondására. A mondatokat nemcsak akusztikailag elemezték, hanem felismeréses tesztelésnek is alávetették. Az eredeti több mint 1300 közlést először is megszűrték, független megítélők kiválasztották azokat, amelyek a legjobban reprezentálták az egyes érzelmeket. A végső 280 mondatot 12 egyetemista ítélte meg, akiknek az volt a feladata, hogy a 14 felsorolt érzelm közül kiválasszák, hogy melyik felelt meg az éppen hallott mondatnak. Az eredmények azt mutatták, hogy a személyek legnagyobb pontossággal a forró dühöt (78 %), az unalmat (76 %) és az érdeklődést (75 %) ismerték fel. Alacsonyabb volt a felismerési aránya a hideg harag (34 %) és a pánikszerű félelem esetében (36 %), valószínűleg azért, mert szerepelt a választható érzelmek listáján az ezekkel egy érzelencsaládba tartozó forró düh és a szorongás is. A legkevésbé pontosan az undort (15%) és a szégyent (22 %) ismerték fel a megítélők.

Földi Éva (1996) vizsgálatában egy semleges tartalmú mondatrészlet (*még mindig*) 25 féle különböző prozódiajú változatát használta fel, amelyekben különböző „közléstartalmak” fejeződtek ki (pl. semleges kijelentés, lemondó szomorúság, rettegés, bosszankodás, gyanakvó kérdés stb.⁵). A kapott mintákat egyrészt számítógéppel elemezte, megmérte az alapfrekvencia- és az intenzitásmenetüket, illetve a teljes időtartamukat. Másrészt a percepció tesztelés módszerével is vizsgálta ezeket. A 89 első éves magyar szakos egyetemistának a hangfelvétel háromszori meghallgatása után kellett a következő két kérdésre válaszolnia: „*Milyen közléstartalmat tulajdonít az elhangzott hangsornak?*” és „*Mely akusztikus sajátosságok fejezik ki a közléstartalmat?*”. A résztvevők nem kapták meg sem a lehetséges érzelmek, sem a lehetséges akusztikus sajátosságok (beszéddallam, tempó, hangszínezet...) listáját. A teljes minta érzelmi tartalmát a megítélők 17 %-a ismerte fel pontosan, 15 %-uk pedig megközelítően jól. A találati arány a semleges kijelentés, a kétségbeesés és a bosszankodás esetében volt a legmagasabb. Ugyanakkor ezeknél a közléseknél is megfigyelhetők voltak rossz válaszok. Így a semleges kijelentésre gyakran mondták, hogy fáradt, unott vagy haragos; a kétségbeesésre hogy félénk vagy csalódott; a

⁵ vegyük észre, hogy itt az érzelmi töltet és a mondat típusa (kérdés, felkiáltás) kombinálódik

bosszankodásra hogy zaklatott vagy vidám. A szörnyülködés közléstartalom felismerése volt a legkevésbé sikeres: pontos megnevezést senki sem tudott adni, és megközelítően jó megoldást is csak hárman.

A szupraszegmentális paraméterek megnevezésénél a válaszok 12,1 %-a értékelhetetlen volt. Az értékelhető válaszokban a legtöbben a hanglejtésformát, a hangszínezetet, az emfatikus nyújtást és a beszédtempót tartották meghatározónak. A viszonylag alacsony felismerési arány hátterében az állhat, hogy a résztvevőknek nagyon sokféle mondatot kellett megítélniük, amelyek nem csak érzelmi töltetben, de modalitásban is különböztek egymástól. A modalitások és az érzelmi töltés együttes variálása nem biztos, hogy szerencsés. Ráadásul nem egy lista alapján, hanem maguktól, előzetes gyakorlás és útmutatás nélkül nevezték meg a megítélők az érzelmeket, ami – mint az eredmények is mutatják – nem volt egyszerű feladat az elsőéves egyetemisták számára.

Markó Alexandra (1999) kutatása részben hasonló ehhez. A *Minden így marad* megnyilatkozást vizsgálta harminchatféle intonációs megformáltságban, amelyeket egy diák mondott fel hangszalagra, majd ezeket vetette alá percepció próbának. A résztvevőknek választással kellett meghatározniuk a „modalitást és a közléstartalmat”⁶, az első esetben öt (kijelentő, kérdő stb.), a másodikban 31 megadott kategória – érzelem vagy attitűd – közül kellett választaniuk (pl. boldog, dacos, szemrehányó, türelmetlen stb.). A résztvevők második feladata az volt, hogy a hallott közlést illesszék be egy leírt szituációba – szintén választással. A modalitás esetében a személyek csak egy minta, a szomorú közlés típusát ismerték fel 100 %-osan. Más esetekben például gyakran összecserélték a felkiáltó és a kijelentő mondatokat. A közléstartalom tekintetében az eredmények azt mutatták, hogy 31-ből 10 közléstartalmat ismertek fel a megítélők 50 %-nál nagyobb arányban. Legjobban a kárörvendő, gúnyos kijelentést, a boldog felkiáltást, a szomorú kijelentést, a nyafogást és a durcás sértődöttséget tudták beazonosítani⁷. A mondatokat a szerző akusztikai elemzésnek is alávetette. Így például a gúnyos és a semleges mondatok összehasonlításakor azt kapta, hogy míg a percepció teszten nagy különbség volt a felismerés aránya között (a gúnyos kijelentést sokkal többen felismerték), a két közlés dallamvonala illetve intenzitása nem különbözött ilyen nagy

⁶ A vizsgálat célja az intonáció „értelem- és érzelemkifejező szerepének” egyidejű vizsgálata volt, mivel „e kettő gyakran nem válik élesen szét” (i.m. 30. o.).

⁷ A kísérletnek a szokásostól eltérő módjára mutathat rá, hogy míg Scherer (2003) összefoglalója alapján (ld. lentebb) a szomorúságot sokkal pontosabban felismerik a személyek, mint az örömet, itt a sorrend megcserélődik.

mértékben. A különbséget ebben az esetben az idői jegyek és a hangszínezet különbözősége adja. A boldog felkiáltás esetében viszont ennek ellenkezője történik: a dallamvonal és az intenzitás szerkezet nagyon eltér a semleges kijelentéséről, az idői vonások viszont kevésbé különböznek.

A magyar kutatásokról összességében azt mondhatjuk, hogy inkább nyelvészeti, mint pszichológiai irányultságúak. Módszertanukban, precizitásukban, az eredmények értelmezésében is nagyban eltérnek az európai pszichológiai, pszicholingvisztikai kutatási gyakorlattól, nem használják azok jól bevált kutatási és elemzési módszereit. Általában megelégszenek az eredmények kvalitatív bemutatásával, nem végeznek statisztikai számításokat, magyarázataik pedig gyakran inkább minősítők, mintsem korábbi eredményekkel összevetett kritikai elemzések.

Vokálisérzelem-felismerés látássérült személyeknél

A kérdésfeltevés, hogy siket illetve vakon született személyek érzelm kifejezése illetve érzelm felismerése más-e mint az egészségeseké, Eibl-Eibesfeldt-hez vezethető vissza. Eibl-Eibesfeldt (1973) siketen és vakon született csecsemők érzelm kifejezéseit vizsgálta. Kutatásai szerint az érzelmek kifejezése univerzális és velünk született, mivel ezeknél a siket és vak csecsemőknél is megfigyelhetők voltak olyan érzelm kifejezések, mint a nevetés, ajaklebiggyesztés, sírás és egyéb érzelmi hangadások.

Bár érdekes kérdés lenne, hogy vajon a látássérült személyek pontosabban felismerik-e a hang alapján az érzelmeket, mint az egészségesek – mivel a vizuális csatorna nem áll rendelkezésükre –, ezzel a témával szinte egyáltalán nem foglalkoznak kutatások. A kevés vizsgálat egyike Dyck és munkatársai (2004) nevéhez fűződik, akik hallássérült és látássérült gyerekeket és serdülőket vizsgáltak olyan általuk kifejlesztett tesztek segítségével (Emotion Recognition Scales), amelyek az érzelmek felismerésének és megértésének képességét mérik. Az érzelmek *megértését* olyan tesztek vizsgálták, amelyekben például a gyerekeknek meg kellett nevezniük a hallott történetek szereplőinek érzelmi állapotát (*Susan kapott egy biciklit*

szülinapjára. Mit érez most Susan?) vagy érzelemszavakat kellett definiálniuk (pl.: „*mit jelent az a szó, hogy dühös?*”). Az érzelmek *felismerésének* képességét érzelmi arckifejezések tesztel és egy vokálisérzelem-felismeréses tesztel mérték (a Vocal Cues Test), amely különböző érzelmi intonációval elmondott kifejezéseket tartalmazott. A kifejezések egy része valós mondat volt („*Nem hiszem el*”), míg másik részük jelentés nélküli szótagból vagy számokból összeállított hangsor volt. A gyerekek feladata az volt, hogy kitalálják: milyen érzelemre utalhat az intonáció.

Az eredmények azt mutatták, hogy mind a látássérült mind a hallássérült gyerekek és serdülők elmaradást mutatnak tipikus fejlődésű társaikhoz képest. De míg a hallássérültek esetében ez a lemaradás mind az érzelfelismerésre, mind az érzelfmegértésre érvényes, addig a látássérültek esetében csak az érzelfelismerésre szorítkozik. A látássérült gyerekek és serdülők ugyanúgy teljesítenek az érzelmi helyzetek megértését vizsgáló teszteken, mint egészséges társaik. A kutatók megvizsgálták azt is, hogy a késés összefügg-e azzal, hogy ezek a gyerekek általában verbális képességek tekintetében is elmaradást mutatnak egészséges társaikhoz képest. A csoportokat a Wechsler-féle intelligenciateszt három, verbális képességet mérő skáláján elért pontszám alapján illesztették, és az eredmények azt mutatták, hogy a hallássérült személyek ugyanúgy vagy még jobban is teljesítenek az összes teszten, mint az egészségesek. A látássérültek azonban éppen a Vokális Érzelfkifejezés alteszten teljesítettek rosszabbul. A látássérültek tehát nehezebben ismerték fel az érzelmeket a hang információk alapján, mint az egészségesek.

Ennek az elmaradásnak az oka egyelőre nem világos. A kérdés azért is érdekes, mert a kapott eredmény ellentmond elvárásainknak, hiszen azt feltételeznénk, hogy a látássérültek jobban fognak teljesíteni ezen a teszten: mivel érzelmekkel kapcsolatos fontos vizuális információk nem állnak rendelkezésükre, jobban figyelembe veszik a hallott információkat.

Az egyik lehetséges magyarázat korábbi eredményekre épít, amelyek szerint a látássérültek inkább szemantikus struktúrák segítségével képezik le a körülöttük lévő világot, és kevésbé figyelnek a jelentést módosító vokális jegyekre. Ezt támasztaná alá az az eredmény, hogy a látássérült személyek rosszabbul teljesítenek a *valós szavakat* tartalmazó vokális érzelfelismerő teszten, mint az egészséges személyek, azonban a *jelentéssel nem bíró kifejezéseket* tartalmazó alteszten nem. Tehát a jelentés elvonja a figyelmüket a vokális jegyekről.

Az az eredmény, hogy a látássérült személyek jól teljesítenek az érzelmek megértésében, de az érzelmek felismerésében nem, azt mutatja, hogy két egymástól elkülöníthető képességről van szó. Így az érzelmek észlelésének és megértésének képessége is máshogy viszonyul a verbális képességekhez. Bár a látássérült személyek rosszabbul teljesítettek az érzelmek felismerésében, mint azt nyelvi képességeikben hasonló egészséges személyek, az érzelmi szavak teszten – ahol érzelemszavakat kellett definiálniuk – mind a látássérültek, mind a hallássérültek jobban teljesítettek, mint az egészségesek.

A kérdés tehát, hogy pontosan milyen területeken hogyan teljesítenek a látássérült személyek az érzelemfelismerésben illetve érzelemmegértésben, nem teljesen tisztázott. Jelen ismereteink szerint nem készült vizsgálat a témában felnőttekkel, így nem tudjuk, hogy felnőtt látássérültek például hogyan teljesítenek hasonló teszteken, illetve milyen felismerési arányt érnének el egy a fenti módszerrel (pl. Banse és Scherer, 1996) végzett kutatásban. A látás- illetve hallássérült gyerekekkel és serdülőkkel végzett kutatások általában késésről, elmaradásról számolnak be, ugyanakkor nem tisztázott, hogy felnőtt korban változik-e ez a folyamat, és ha igen, milyen irányba.

20 európai és amerikai percepciók kutatás eredményei összegzése alapján (Scherer, 2003) azt mondhatjuk, hogy a haragot átlagosan 77, a szomorúságot 71, a félelmet 61, az örömet 57, az undort 31 %-os pontossággal ismerik fel a hallgatók a beszédhangból. Az egyes kutatások eredményei között azonban nagy különbségek vannak az alkalmazott módszertől függően.

Érdekes különbség, hogy bizonyos érzelmeket a hang alapján, míg másokat az arc kifejezés alapján ismerünk fel jobban. Míg az öröm arc kifejezését szinte tökéletes biztonsággal képesek vagyunk felismerni, addig, mint láttuk, a hang alapján a felismerés aránya csupán 57 %-os. Az undor esetében szintén hasonló a helyzet. Johnstone és Scherer (2000) szerint ennek evolúciós okai vannak. Hiszen ha például egy étel romlott, akkor ezt tudatni egymással, az undor érzését kifejezni főleg akkor van értelme, ha a két élőlény közel van egymáshoz, ha látják egymást. Míg ha egy idegen betolakodóval vagy ellenséggel találjuk magunkat szembe (félelem vagy a harag), akkor nagyon is hasznos, ha ezt hangunkkal nagyobb távolságokon keresztül is tudatni tudjuk társainkkal.

Érzelemítélet						
Eljátszott érzelem	Harag	Undor	Félelem	Öröm	Szomorúság	Nincs érzelem
Harag	58	18	6	4	2	12
Undor	31	40	4	4	4	17
Félelem	0	2	60	3	27	8
Öröm	6	3	21	51	6	12
Szomorúság	1	2	24	1	63	8
Nincs érzelem	4	4	4	5	11	72

5. táblázat Juslin és Laukka (2001) kutatásában vizsgálati személyek színészek által játszott közlések érzelmi töltetét ítélték meg. A táblázat mutatja (félkövérrel kiemelve) a helyes találatok arányát (%-ban), valamint azt, hogy milyen más kategóriákat milyen arányban neveztek meg helytelenül a megítélők.

A felismeréses kutatások egyik érdekes kérdése, hogy az egyes érzelmeket a hallgatók milyen más érzelmekkel tévesztik össze. Mint Banse és Scherer (1996) kutatása ismertetésénél említettük, az egy érzelemcsaládba tartozó érzelmek például könnyen összetéveszthetők. Ugyanígy a kellemesség dimenzió mentén közel álló érzelmeket, például az érdeklődést, a boldogságot és a büszkéséget is nehéz egymástól elkülöníteni. A hibázási mintázatok vizsgálata (ld. pl. 5. táblázat) azonban közelebb vihet minket ahhoz, hogy különböző dimenziók mentén el tudjuk egymástól különíteni az egy érzelemkategóriába tartozó érzelmeket (Johnstone és Scherer, 2000).

3.4. Érzelmek és beszédtechnológia

A beszédszintézis célja, hogy számítógépes programok segítségével mesterségesen állítson elő emberi beszédet. Mivel az emberi beszéd természetes jellemzője a prozódia, ezért az informatikusoknak az is a feladatuk, hogy dallammal, hangsúlyokkal lássák el a mesterséges beszédet. A legújabb kutatások pedig arra is keresik a választ, hogy hogyan lehet mesterséges érzelmes beszédet előállítani. Egyrészt megpróbálnak az eddigi produkciós kísérletek eredményei alapján érzelmetli megnyilatkozásokat előállítani, másrészt ezeket a gépi beszéd-megnyilatkozásokat percepciók próbáknak vetik alá, hogy teszteljék, mennyire hasonlítanak

az emberi beszédhez. A gépi érzelemfelismerés során olyan programokat próbálnak meg előállítani, amelyek a természetes emberi beszédben előforduló érzelmeket automatikusan felismerik.

3.4.1. Érzelmesbeszéd-előállítás

Yanushevskaya és munkatársai (2005) egy adott, formánsszintetizátorral előállított szócsoporthoz szavait módosították különbözőképpen úgy, hogy olyan szavakat kaptak, amelyek levegős, suttogó, laza-rekedtes, feszes és köznap hangszínezetűek voltak. Az alaphangfrekvencia változtatásával pedig öt különböző érzelmeket: szomorúságot, félelmet, unalmat, örömet és a felháborodást kifejező szavakat állítottak elő. A harmadik esetben kombinálták a hangszínezet- és az alaphangfrekvencia-változtatásokat, és a következő párokat kapták: levegős hangszínezet – szomorúságot kifejező F0-menet, suttogó hangszínezet – a félelem F0-menete, laza-rekedtes hangszínezet – az unalom F0-menete, feszes hangszínezet – az öröm F0-menete, feszes hangszínezet – a felháborodás F0-menete. A meghallgatásos tesztek során a résztvevők ellentétpárok segítségével értékelték a szavakat. A kutatók eredményül azt kapták, hogy azokat az érzelmi tónusú szavakat, amelyeket az alaphangfrekvencia változtatásával hoztak létre, rosszabbul ismerték fel a tesztelők, mint azokat, amelyeknél csak a hangszínezetet változtatták. A változtatások kombinációja esetében sem volt nagyobb a felismerési arány, egyedül a félelem esetében volt kiugróan magasabb a felismerés. A szerzők arra a következtetésre jutottak, hogy az alaphangfrekvencia változtatása önmagában nincs meghatározó befolyással az érzelemfelismerésre.

Audibert és munkatársai (2005) hasonló módon vizsgálták nyolc érzelmeket: az aggodást, a csalódottságot, az undort, a nyugtalanságot, a boldogságot, a lemondást, a megelégedettséget, a szomorúságot és a semleges kifejezést. Nagyobb különbségnek mondható, hogy itt egy színész mondta el a később számítógéppel módosított alapszavakat. Módosították az alaphangfrekvenciát, az intenzitást, a hangszínezetet és a hangidőtartamot. A meghallgatásos tesztek végső következtetése az volt, hogy nincs olyan paraméter, amely önmagában, egyedül teljes mértékben kifejezné az érzelmeket. Fontos eredményeik, hogy az alaphangfrekvencia-menetnek nagyobb a jelentősége a pozitív érzelmek kifejezésében, mint a negatívakéban, a hangszínezet és a hangidőtartamok nagyobb jelentőséggel bírnak a negatív érzelmek

kifejezésében, mint a pozitívokéban. Az intenzitásmenet nem hordozott jelentős információt az érzelemről abban az esetben, ha kizárólag azt változtatták.

Fék és munkatársai (2005) kutatásukban az alaphfrekvencia-menet és a hangidőtartam szerepét vizsgálták. Semleges tartalmú mondatokat (pl. *Ő egy egyszerű tanársegéd vagy Délután felhívnak, hogy megbeszéljünk egy találkozót.*) kilencféle érzelmi hangsúllyal (félelem, meglepődés, harag, szomorúság, öröm, lekicsinylés, aggodás, értetlenség, megkönnyebbülés) olvasott föl három színész. A semleges érzelmi töltetű mondatokat számítógép segítségével módosították úgy, hogy a színészek által felolvasott érzelmi töltetű mondatokból kinyert alaphfrekvenciát és hangidőtartamot ráültették ezekre a semleges mondatokra. A kapott mondatok összes változatát (a három színész bemondásában, az összes érzelmi töltettel, a természetes ejtésű és a számítógéppel módosított változatban) megítéléses vizsgálatnak vetették alá. A meghallgatásos tesztek során arra a kérdésre keresték a választ, hogy mely érzelmek fejezhetők ki elfogadhatóan pusztán az F0-menet és a hangidőtartamok átültetésével. A megítélőknek egyrészt arra kellett választ adniuk, hogy a meghallgatott mondatoknak milyen az érzelmi modalitásuk (az összes lehetséges modalitás listájából kellett választaniuk), másrészt meg kellett ítélniük, hogy milyen volt a hang minősége. A természetes ejtésű mondatok közül legkönnyebben a haragot, a szomorúságot és a lekicsinylést ismerték fel a megítélők. Közepesen jól ismerték fel a meglepődést, az értetlenséget és a megkönnyebbülést, és legrosszabbul a félelmet és az aggodást. E két utóbbi érzelmestében a leggyakoribb tévesztés az egymással való felcserélés volt, ami abból eredhet, hogy ez a két érzelmest elég hasonló egymáshoz. A számítógéppel módosított minták esetében változott a felismerés sorrendje. Legjobban a meglepődést, a lekicsinylést és az értetlenséget ismerték fel a személyek, közepesen jól a szomorúságot, a többi érzelmest pedig ennél rosszabban. Nem volt azonban összefüggés a szintetizált felvételek minősége, és a kifejezett érzelmest felismerése között. A legjobb minőséget a haragot hordozó felvételek érték el, míg ezek felismerési százaléka igen alacsony volt. A fő kérdésekre azt a választ kapták, hogy az alaphfrekvencia-menet és a hangidőtartam a meglepődés és a lekicsinylés esetében a legfontosabb.

3.4.2. Gépi érzelemfelismerés

Informatikusok napjainkban olyan programokat próbálnak meg előállítani, amelyek képesek felismerni a hangból a beszélő érzelmi állapotát. Bár az ilyenfajta kutatások még csak nemrég kezdődtek, az igény a fejlesztésükre egyre nagyobb (pl. automatikus adatbányászat telefonos ügyfélszolgálatokon). A gépi érzelemfelismeréses kutatások általában statisztikai módon közelítik meg a problémát. A leggyakoribb módszer, hogy színészek által eljátszott mondatokat elemeznek, vagy spotánbeszéd-részleteket címkéznak fel, és ezekhez rendelnek jellemzővektorokat, amelyek több paraméter és az ezekből származtatott statisztikák (variancia, minimum-maximum stb.) együttesét tartalmazzák.

Láttuk, hogy az emberek a véletlennél nagyobb valószínűséggel képesek felismerni például a szomorúságot, a haragot vagy az örömet. A gépek egyelőre általában kevésbé jól teljesítenek, mint az emberek. Alapvető probléma, hogy minél több érzelmet kell felismernie egy programnak, annál rosszabb a teljesítménye. Csupán két érzelem felismerése esetén azonban a teljesítmény 90 %-os is lehet. (Schuller és mtsai. 2005). Ezen kívül a beszélőfüggő rendszerek teljesítménye sokkal jobb, mint a beszélőfüggetleneké, tehát könnyebb a dolgunk, ha programunknak csupán egy személy érzelmi állapotait kell azonosítania, mint ha sok-sok emberét.

Magyar kutatások is megjelentek ezen a területen. Tüske és munkatársai (2007) beszélőfüggetlen adatbázison 42,9 %, beszélőfüggő adatbázison pedig 60 %-os felismerési pontosságú rendszert fejlesztettek ki. Vicsi és Sztahó (2009) vizsgálatukban egy telefonos ügyfélszolgálat hívásaiban a semleges és az elégedetlenséget kifejező (ideges és panaszkodó) állapot megkülönböztetését tűzte ki célul, és programjukkal átlagosan 77,2 %-os pontossággal el tudták különíteni a két állapotot.

(A szintetizált beszéd előállításáról és a gépi érzelemfelismerésről lásd részletesebben a következő irodalmakat: Cahn, 1990; Ververidis és Kotropoulos, 2006; Vogt és mtsai, 2008; Bosch, 2003; Gobl és Chasaide 2003; Laukka és mtsai, 2011.)

Összegzés

Ebben a fejezetben olyan kutatásokat mutattunk be, amelyek mindegyike a beszédben megfigyelhető érzelmeket vizsgálta. A három általunk bemutatott terület, a társalgáselemzés, a tartalomelemzés és a vokális érzelmerkifejezések vizsgálata más-más oldalról közelíti meg a témát. Míg a társalgáselemzés a szociálpszichológiával, a tartalomelemzés az irodalommal, a vokálisérzelmerkifejezés-kutatás pedig leginkább a fonetikával és az informatikával határos területek.

Egy az érzelmelek és a beszéd kapcsolatát vizsgáló kutatás megtervezésekor több szempont is felmerül. A vizsgálandó jelenség vagy a rendelkezésünkre álló eszközök, illetve a körülöttünk csoportosuló kutatók és az általuk képviselt iskola gyakran meghatározzák, hogy melyik irányzat lesz a domináns vizsgálatunkban. Azonban ez nem mindig van így, legtöbbször nem válnak el ennyire élesen egymástól a megközelítések. Hétköznapi párbeszédet vagy monológokat például éppúgy elemezhetünk a társalgáselemzésben használt GAT jeleinek segítségével, mint tartalomelemzéssel vagy fonetikai szempontból. A legnagyobb különbség talán a használt elemző eszközökben van (pl. Atlas.ti vagy Praat), ugyanakkor az is lehetséges, hogy tartalomelemző programot használunk nem szókategóriák, hanem például beszédbeli rontások, megakadások címkézésére, mert ez teszi lehetővé számunkra legjobban a későbbi statisztikai elemzést, vagy fonetikai elemző program segítségével mérjük le a társalgó felek megszólalásai között eltelt pontos időt társalgáselemzéses kutatásban.

Míg az első vizsgálatok leginkább kvalitatívak, leíró jellegűek voltak, a kutatások a kvantitatív irányba mozdultak el. Ma már szinte minden elemzés valamilyen szoftver segítségével történik, és tartalmaz statisztikai számításokat. A legújabb kutatások az automatizálódás irányába mutatnak, olyan programokat készítenek informatikusok, amelyek képesek például a hangzó beszédet írásos szöveggé átkonvertálni, bizonyos szavakat és szókategóriákat egy szövegkorpuszban megtalálni, valamint az emberi beszéd érzelmi töltetét akusztikus és más jegyek alapján automatikusan felismerni. Elképzelhető, hogy a jövőbeli kutatások célja e területek összekapcsolása lesz.

- Apple, W., Streeter, L. A., Krauss, R. M. (1979) Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, 37, 715-727.
- Aubergé, V. és Lemaître, L. (2000) The prosody of smile. *SpeechEmotion-2000*, 122–126.
- Audibert, N., Aubergé, V. & Rilliard, A. (2005) The prosodic dimensions of emotion in speech: the relative weights of parameters. In: *Interspeech'2005 – Eurospeech*. Lisszabon, Portugália, 525–528.
- Banos, R. M., Botella, C., Liaño, V., Rey, B., Guerrero, B., Alcaniz, M. (2003) Virtual reality as mood induction procedure. Presented at PRESENCE 2003, 6th Annual International Workshop on Presence.
- Banse, R. és Scherer, K. R. (1996) Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology* 70 (3), 614–636.
- Bänzinger, T. és Scherer, K. R. (2005) The role of intonation in emotional expressions. *Speech Communication* 46, 252–267.
- Battacchi, M. W., Suslow, T., Renna, M. (1996) *Emotion und Sprache. Zur Definition der Emotion und ihren Beziehungen zu kognitiven Prozessen, dem Gedächtnis und der Sprache*. Frankfurt am Main: Lang.
- Boersma, P. és Weenick, D. (2009) Praat: doing phonetics by computer (Version 5.1.05.) Computer program, Insitute of Phonetic Sciences, University of Amsterdam, Hollandia. www.praat.org Letöltés ideje: 2009.05.01.
- Bolla, K. (1992) Érzelmek kifejezése a megnyilatkozás szupraszegmentális hangszerkezetében. Szupraszegmentális elemzések. *Egyetemi Fonetikai Füzetek* 7. 90–116.
- Bosch, L. T. (2003) Emotions, speech and the ASR framework. *Speech Communication*, 40, 1, 213-225.
- Böhm, T. (2009) Analysis and modeling of speech produced with irregular phonation, PhD disszertáció. BME Villamosmérnöki és Informatikai Kar, Budapest.

- Breitenstein, C., D. Van Lancker, I. Daum (2001) The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Cognition & Emotion*, 15, 1, 57-79.
- Brewer, D., Doughtie E. B., Lubin, B. (1980) Induction of mood and mood shift. *Journal of Clinical Psychology*, 36, 215-226.
- Bucci, W. és Freedman, N. (1981) The language of depression. *Bulletin of Menninger Clinic*, 45, 334-58.
- Cahn, J. E. (1990) The Generation of Affect in Synthesized Speech. *Journal of the American Voice I/O Society*, 1-19.
- Campbell, R. S. és Pennebaker, J. W. (2002) The secret life of pronouns: flexibility in writing style and physical health. *Psychological Science*, 14, 60-65.
- Cohen, A. S., Hong, S. L., Guevara, A. (2010) Understanding emotional expression using prosodic analysis of natural speech: Refining the methodology. *Journal of Behavior Therapy and Experimental Psychiatry*, 41, 150-157.
- Cowie, R. R. és Cornelius, R. (2003) Describing the emotional states that are expressed in speech. *Speech Communication*, 40, 5-32.
- Deppermann, A. és Lucius-Hoene, G. (2005) Trauma erzählen – kommunikative, sprachliche und stimmliche Verfahren der Darstellung traumatischer Erlebnisse. *Psychotherapie und Spezialwissenschaft. Zeitschrift für Qualitative Forschung und klinische Praxis*. 1. 35-73.
- Douglas-Cowie, E. és Cowie, R. (2002). *The vocal correlates of stress: a summary*. HTML dokumentum, URL:
http://www.image.ece.ntua.gr/oresteia/private/contributed_docs/reports/Voice-and-Stress.pdf Letöltés ideje: 2008. március 13.
- Douglas-Cowie, E., Campbell, N., Cowie, R., Roach, P. (2003) Emotional speech: Towards a new generation of databases. *Speech Communication*, 40, 33-60.
- Drahota, A., Costall, A., Reddy, V. (2008) The vocal communication of different kinds of smile. *Speech Communication*, 50, 278-287.
- Dyck, M. J., Farrugia, Ch., Shochet, I. M., Holmes-Brown, M. (2004) Emotion recognition/understanding ability in hearing or vision-impaired children: do sounds,

- sights, or words make the difference? *Journal of Child Psychology and Psychiatry*, 45, 4, 789–800.
- Ehmann, B. (2002) *A szöveg mélyén*. Budapest, Új Mandátum Könyvkiadó.
- Eibl-Eibesfeldt, I. (1973) The expressive behavior of the deaf-and-blind-born. In Von Cranach M. & Vine I. (eds.) *Social communication and movement* (163–194). New York, Academic Press.
- Ekman, P. (1999) Basic emotions. In T. Dalgleish & M. Power, *Handbook of cognition and emotion*. Pp 45-60. Sussex, UK, John Wiley & Sons.
- Fék M., Olaszy, G., Németh, G., Gordos, G. (2005). Érzelem kifejezése gépi beszéddel. *Beszéd kutatás 2005*. MTA Nyelvtudományi Intézet, 134–144.
- Fiehler, R. (1990) *Kommunikation und Emotion. Theoretische und empirische Untersuchungen zur Rolle von Emotionen in der verbalen Interaktion*. Walter de Gruyter, Berlin.
- Fiehler, R. (2001) Emotionalität im Gespräch. In Klaus, B., Gerd, A., Wolfgang, H., Sven, S. F. (Hrsg) *Text- und Gesprächslingistik. Ein internationales Handbuch zeitgenössischer Forschung*. Halbband 2. Walter de Gruyter, Berlin, New York, 1425–1438.
- Fónagy I. (1981) Emotions, voice and music. In: J. Sundberg (ed.) *Research aspects on singing. Papers given at a seminar organized by the Committee for the Acoustics of Music*, 51-79. Stockholm: The Royal Swedish Academy of Music.
- Fónagy, I. (1998) *A költői nyelvről*. Corvina, Budapest.
- Fónagy I. (2001) *Languages within language: an evolutive approach*. Amsterdam; Philadelphia: John Benjamins.
- Fónagy, I. és Magdics, K. (1967) *A magyar beszéd dallama*. Akadémiai Kiadó, Budapest.
- Fontaine, J., Scherer, K., Roesch, E., Ellsworth, P. (2007) *The world of emotions is not two-dimensional*. *Psychological Science*, 18 (12), 1050-1057.
- Forgács, J. (2003) *Az érzelmek pszichológiája*. Kairosz Könyvkiadó, Budapest.
- Földi, É. (1996) Az érzelemkifejezés szupraszegmentális formái és percepciója. *Egyetemi Fonetikai Füzetek* 20.

- Frolov M. V., Milovanova G. B., Lazarev N. V., Mekhedova A. Y. (1999) Speech as an Indicator of the Mental Status of Operators and Depressed Patients. *Hum. Psychol.* 66, 398–412.
- Frye, N. (1998) *A kritika anatómiája*. Helikon, Budapest.
- Fussell, S. R., Ed. (2002). *The verbal communication of emotion: Interdisciplinary perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Fülöp, É. (2008) Az érzelmek reprezentációja történelmi regényekben és történelemkönyvekben. *Megismerés, reprezentáció, értelmezés, 10 éves a PTE Pszichológiai Doktori Iskolája*, 99-114.
- Gobl, C. és Ni Chasaide, A. (2003) The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, 40 (1-2), 189-212.
- Goldman-Eisler, F. (1968) *Psycholinguistics: experiments in spontaneous speech*. London: New York, Academic Press.
- Gottschalk, L. A., Gleser, G. C. (1969) *The measurement of psychological states through the content analysis of verbal behavior*. Univ. Of California Press, Berkeley. MONITOR, 4. sz., 1972. OIEI, Kézirat.
- Gottschalk, L. A., Winget, C. N., Gleser, G. C. (1969) *Manual of Instructions for Using the Gottschalk-Gleser Content Analysis Scales: Anxiety, Hostility, Social Alienation- Personal Disorganization*. Los Angeles, Berkeley, University of California Press.
- Hart, R. P. (2001) Redeveloping DICTION: theoretical considerations. *See West*, pp. 43–60.
- Iványi, Zs. (2001) A nyelvészeti konverzációelemzés. *Magyar Nyelvőr* 125, 74–93.
- Izard, C. E. (1993) Organizational and motivational functions of discrete emotions. In M. Lewis & J. M. Haviland, *Handbook of emotions*, pp. 631-41. New York, Guilford Press.
- Johannes, B., Salnitski, P., Gunga, H-C., Kirsch, K. (2000) Voice stress monitoring in space – possibilities and limits. *Aviation, Space and Environmental Medicine*, 71, 9 (11), A58-A65,
- Johnson W. F., Emde, R. N., Scherer K. R., Klinnert M. D. (1986) Recognition of Emotion From Vocal Cues. *Archives from General Psychiatry*, 43(3), 280-283.

- Johnstone, T. és Scherer, K. R. (2000). Vocal communication of emotion. In Lewis, M. és Haviland-Jones, J. (Eds.), *Handbook of Emotions, Second Edition*. Guilford Press, New York, 220–235.
- Juslin, P. N., és Laukka, P. (2001) Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion*, 1, 4, 381-412.
- Juslin, P. N., és Scherer, K. R. (2005) Vocal expression of affect. In J. A. Harrigan, R. Rosenthal, K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 65-135). New York: Oxford University Press.
- Karczag, J. (1988) A Gottschalk-Gleser-féle tartalomelemzési módszer. In Mérei, F. és Szakács, F. (szerk.) *Pszichodiagnosztikai vademecum*. Budapest, Tankönyvkiadó, 273–353.
- Kienast, M. és Sendlmeier, W. F. (2000) Acoustical analysis of spectral and temporal changes in emotional speech. *SpeechEmotion-2000*, 92-97.
- Lalouschek, J. (2002) Frage-Antwort-Sequenzen im ärztlichen Gespräch. In: Brünner, G., Fiehler, R. & Kindt, W. (Hrsg) *Angewandte Diskursforschung*, Band 1. Verlag für Gesprächsforschung.
- Lalouschek, J. (2002b) Ärztliche Gesprächsausbildung. Eine diskursanalytische Studie zu Formen des ärztlichen Gesprächs. Radolfzell, Verlag für Gesprächsforschung.
- Lalouschek, J. és Menz, F. (2002) Empirische Datenerhebung und Authentizität von Gesprächen. In Brünner, G., Fiehler, R., Kindt, W. (Hsg.) *Angewandte Diskursforschung*, Band 1., 46 – 68. Radolfzell, Verlag für Gesprächsforschung.
- Lalouschek, J., Menz, F., Wodak, R. (1990) *Alltag in der Ambulanz. Gespräche zwischen Ärzten, Schwestern und Patienten* Gunter Narr Verlag: Tübingen.
- Lang, P. J., Bradley, M. M., Cuthbert, B. N. (2005) International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-6. Gainesville, FL: University of Florida.
- László, J. (2005a) A narratív pszichológiai tartalomelemzés. *Magyar Tudomány*, 11, 1366–1376.
- László, J. (2005b) A történetek tudománya: Bevezetés a narratív pszichológiába. Pszichológiai Horizont; Budapest, Új Mandátum Könyvkiadó.

- Laukka, P. (2004) Vocal expression of emotion: Discrete-emotions and dimensional accounts. In Comprehensive summaries of Uppsala dissertations from the faculty of social sciences, No. 141 (pp. 1-80). Uppsala, Sweden: Acta Universitatis Upsaliensis.
- Laukka, P., Juslin, P. N., & Bresin, R. (2005). A dimensional approach to vocal expression of emotion. *Cognition and Emotion*, 19, 633-653.
- Laukka, P., Neiberg, D., Forsell, M., Karlsson, I., Elenius, K. (2011) Expression of affect in spontaneous speech: Acoustic correlates and automatic detection of irritation and resignation. *Computer Speech and Language*, 25, 84–104.
- Lorenz, M. és Cobb, S. (1952) Language behavior in manic patients. *Arch. Neur. Psych.*, 67, 763–70
- Markó, A. (1999) Intonációról az ezredfordulón, avagy minden így marad? *Egyetemi Fonetikai Füzetek* 25., 29–41.
- Menz, F., Lalouschek, J., Gstettner, A. (2008) *Effiziente ärztliche Gesprächsführung. Optimierung kommunikativer Kompetenz in der ambulanten medizinischen Versorgung. Ein gesprächsanalytisches Trainingskonzept*. LIT-Verlag: Wien/Münster.
- Mergenthaler, E. (1996) Emotion-abstraction patterns in verbatim protocols: a new way of describing psychotherapeutic processes. *Journal of Consulting and Clinical Psychology*, 64, 1306–15
- Mitchell, C. J., Menezes, C., Williams, J. C., Pardo, B., Erickson, D., Fujimura, O. (2000) Changes in syllable and boundary strengths due to irritation. *SpeechEmotion-2000*, 98-103.
- Oatley, K. és Jenkins (2001) *Érzelmek*. Osiris, Budapest.
- Osgood, C.E., Suci, G., & Tannenbaum, P. (1957) *The measurement of meaning*. Urbana, IL: University of Illinois Press.
- Pennebaker, J. W. és Francis, M. E. (1996) Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion*, 10, 601–26.
- Pennebaker, J. W., Francis, M. E., Booth, R. J. (2001) *Linguistic Inquiry and Word Count (LIWC): LIWC 2001*. Mahwah, NJ: Erlbaum
- Pereira, C. (2000) Dimensions of emotional meaning in speech. *Proceedings of the ISCA Workshop on Speech and Emotion*, 25–28.

- Pittam, J. és Scherer, K. R. (1993). Vocal expression and communication of emotion. In Lewis, M. & Haviland, J. M. (Eds.) *Handbook of emotions*. Guilford Press, New York, 185–197.
- Pléh, Cs., és Czigler, I., (1979) Kísérlet politikai kifejezések megítélésére alkalmas szemantikus differenciál skálák kidolgozására és alkalmazására. In: Hunyadi György (ed.), *Pszichológiai Tanulmányok 15*. Budapest: Akadémiai Kiadó, 479-533.
- Pohárnok M. (2004) A térben való mozgás narratív dimenziójának vizsgálata borderline és depressziós betegek élettörténeti epizódjaiban. In: Erős, F. (szerk.), *Az elbeszélés az élmények kulturális és klinikai elemzésében*. Akadémiai, Budapest, 153-166.
- Pólya, T. (2004) A narratív perspektíva hatása az elbeszélő személy észlelésére. In: Erős, F. (szerk.), *Az elbeszélés az élmények kulturális és klinikai elemzésében*. Akadémiai, Budapest, 89-107.
- Pólya T., Kis B., Naszódi M., László J. (2005) Az érzelmi tapasztalat minősége az élettörténeti elbeszélésben. A LAS-verticum perspektíva modulja. *Pszichológia*, 25, 2, 143-155.
- Pólya, T., Kis, B., Naszódi, M., László, J. (2007) Narrative perspective and the emotion regulation of a narrating person. *Empirical Text and Culture Research* 3, 50-61.
- Putnoky, J. (1978) A „motoritás” a szójelentés három főbb működési körében. *Magyar Pszichológiai Szemle*, 35, 203-213.
- Roessler, R. és Lester J. W. (1976) Voice predicts affect during psychotherapy. *Journal of Nervous and Mental Disease*, 163, 3, 166-76.
- Sator, M. (2003) *Zum Umgang mit Relevanzmarkierungen im ÄrztInnen-PatientInnen-Gespräch. Eine konversationsanalytische Fallstudie eines Erstgesprächs auf der onkologischen Ambulanz*. Wien, Univ., Dipl.-Arb.
- Scherer, K. R. (2000) Psychological models of emotion. In Borod, J. (Ed.). *The neuropsychology of emotion* (pp. 137–162). Oxford/New York: Oxford University Press.
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication* 40. 227–256.

- Scherer, K. R. és Zentner, M. R. (2001) Emotional effects of music: production rules. In Juslin, P. N. & Sloboda, J. A. (ed) *Music and emotion: theory and reseatch*. Oxford University Press, Oxford, New York, 361–392.
- Scherer, K. R., Banse, R., Wallbott, H.G., Goldbeck, T. (1991) Vocal Cues in Emotion Encoding and Decoding. *Motivation and Emotion*, 15(2), 123–148.
- Schuller, B., Müller, R., Land, M., Rigoll, G. (2005) Speaker independent emotion recognition by early fusion of acoustic and linguistic features within ensembles. In *Proceedings of INTERSPEECH-2005*. Lisbon, 805–808.
- Selting, M., Auer, P., Barden, B., Bergmann, J., Couper-Kuhlen, E., Günther, S., Meier, C., Quasthoff, U., Schlobinski, P., Uhmann, S. (1998) Gesprächsanalytisches Transkriptionssystem. *Linguistische Berichte* 173. 91–122.
- Sobin, C. és Alpert, M. (1999) Emotion in Speech: The Acoustic Attributes of Fear, Anger, Sadness, and Joy. *Journal of Psycholinguistic Research*, 28, 4, 347–365.
- Stirman, S. W. és Pennebaker, J. W. (2001) Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine*, 63, 517–22.
- Stone, L. D., Pennebaker, J. W. (2002) Trauma in real time: Talking and avoiding online conversations about the death of Princess Diana. *Basic and Applied Social Psychology*, 24, 172–82.
- Szabó, E. (2008) A szomorú és a vidám érzelmi állapot megjelenése a beszédben. *Magyar Pszichológiai Szemle*, 63, 4, 651–668.
- Tüske, Z., Simon, M., Mihajlik, P., Fegyő, T. (2007) Érzelmek automatikus felismerése a beszéd akusztikus jellemzői alapján. In: Gósy M (ed.) *Beszéd kutatás 2007*. Budapest: MTA Kiadó, pp. 151-161.
- Velten, E. (1968). A laboratory task for induction of mood states. *Behavior Research & Therapy*, 6(4), 473-482.
- Ververidis, D. és Kotropoulos, C. (2006) Emotional speech recognition: resources, features, and methods. *Speech Communication*, 48, 1162–1181.
- Vicsi Klára, Sztahó Dávid (2009) Ügyfél érzelmi állapotának detektálása telefonos ügyfélszolgálati dialógusban. In: Tanács Attila, Szauder Dóra, Vincze Veronika (szerk.)

VI. Magyar Számítógépes Nyelvészeti Konferencia. Szeged, Magyarország, 2009.12.03-2009.12.04. Szeged: JATEPress, pp. 217-225.

Vogt, T., André, E., Wagner, J. (2008) Automatic Recognition of Emotions from Speech: a Review of the Literature and Recommendations for Practical Realisation. In Peter, Ch., Beale, R. (szerk.) *Affect and Emotion in Human-Computer Interaction, From Theory to Applications. Lecture Notes in Computer Science*, 4868, 75–91. Springer.

Weintraub, W. (1981) *Verbal Behavior: Adaptation and Psychopathology.* New York: Springer

Westermann, R., Spies, K., Stahl, G., Hesse, F. W. (1996) Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, 26, 4, 557-580.

Wodak, R. (1981) *Das Wort in der Gruppe. Linguistische Studien zur therapeutischen Kommunikation.* Vienna: Akademie der Wissenschaften.

Yanushevskaya, I., Gobl, C., Ní Chasaide, A. (2005). Voice quality and f0 cues for affect expression: implications for synthesis. In *Interspeech'2005 – Eurospeech*. Lisszabon, Portugália, 1849–1852.

Tanulmány [2]

Tisljár-Szabó, E., Pléh, C. (2014) Ascribing emotions depending on pause length in native and foreign language speech. *Speech Communication*, 56, 35–48. DOI 10.1016/j.specom.2013.07.009 (IF: 1,283)

(a megjelent változattal tartalmilag megegyező kézirat változat)

Although the relationship between emotions and speech is well documented, little is known about the role of speech pauses in emotion expression and emotion recognition. The present study investigated how speech pause length influences how listeners ascribe emotional states to the speaker. Emotionally neutral Hungarian speech samples were taken, and speech pauses were systematically manipulated to create five variants of all passages. Hungarian and Austrian participants rated the emotionality of these passages by indicating on a 1-6 point scale how angry, sad, disgusted, happy, surprised, scared, positive, and heated the speaker could have been. The data reveal that the length of silent pauses influences listeners in attributing emotional states to the speaker. Our findings argue that pauses play a relevant role in ascribing emotions and that this phenomenon might be partly independent of language.

Keywords: Emotion ascribing; Cross-linguistic; Speech pauses; Silent pause duration; Foreign language

1. Introduction

The goal of vocal emotion research is to understand the relationship between emotions and certain speech parameters. Most of these studies are encoding or decoding studies (for reviews see [1, 2]). In the case of the encoding studies, the acoustic pattern of expressions uttered in different emotional states is analysed, while in the case of the decoding studies, the focus is on the recognition of emotions. Studies in the area of decoding examine whether

listeners are able to recognise emotions in speakers' speech and if so, which parameters help listeners with this activity.

Although the relationship between emotions and speech is well documented [3-7], few studies have examined the role of speech pauses in emotion expression and emotion recognition. In the present study, the main question concerned how listeners' emotion judgments are influenced by the length of the speech pauses. In addition, a second aim was to compare two different nations (Hungary and Austria) on the same set of stimuli to address questions about sample differences in emotion decoding according to speech pause length. To examine this concept, we conducted two experiments on the same speech material: one in a native context and one in a foreign language context.

1.1. Emotions and speech

The combination of coding and decoding studies is used as another method to examine the effect of change in certain speech parameters on listeners' emotion judgments (for an overview see [2]). For example, with the help of computer programs, acoustic cues in neutral or meaningless sample utterances are systematically manipulated, and researchers examine how this phenomenon influences listeners in terms of the type and strength of the emotions that they ascribe to the speaker. There are many parameters to manipulate, including fundamental frequency (F0), range, intensity, emphasis, and temporal cues [8-10].

For example, Bergman et al. [9] manipulated F0 range, jitter, intensity, intonation contour, and duration in expressions produced by actors. The listeners' task was to rate the emotionality of the speaker with the help of a rating scale with labels for emotion and attitude categories. According to their main results, if F0 range increased, the speakers were rated as more emphatic, angrier, and more aroused, while if F0 range decreased, speakers were rated as sadder. If the intensity of the sentence increased, ratings on reproachful, angry, and empathic scales were higher. The systematic modification of accent-high had the most effect on scales that measured cognitive state or attitude. The higher the accent was, the more contradictory, the more reproachful, and the more empathic the speaker was rated. In the study, the length of the accent was also modified by lengthening its vowel. According to the results, a short vowel length leads listeners to rate speakers as happier, while a long vowel length leads listeners to rate the speakers as sadder.

In other studies, both neutral and emotionally loaded sentences are modified [9, 10]. Breitenstein et al. [10] used neutral and emotionally loaded sentences belonging to four emotion categories from a previous study [11]. Stimulus material was compiled by systematic variation of the F0 variation and speech rate of the sentences in six ways for both parameters. The listeners' task was to rate the emotional tone of voice of each sentence by choosing one of the emotion labels. The results showed that despite the manipulation, subjects most frequently chose the emotion label that corresponded to the emotion category in which the speaker had originally made the utterance. Recognition of happy and sad sentences was most markedly affected by increased speech rate. When the speed of the originally happy sentence was increased, the rate of neutral, frightened, and angry answers increased. When the speed of the originally sad sentences was increased, again, the rate of neutral and frightened answers increased. In most of the sentences, independent of the emotional category, if the utterance became slower, then the frequency of sad ratings increased, while the frequency of frightened, angry, and neutral ratings decreased.

The results of the above mentioned studies show us that modifications in even one speech parameter can change listeners' opinions about the speaker; moreover, listeners are able to perceive changes in temporal parameters of speech, and accordingly, they ascribe different emotional states to the speaker.

1.2. Relationship between speech pauses and emotions

Only a few studies have investigated the role of speech pauses, most of which have studied the relationship between pauses and the cognitive activity needed for speech. According to Goldman-Eisler [12, 13] and Rochester [14], an increase in speech disfluencies and pauses is related to increased cognitive activity. Thus, pauses are more likely to occur before a word that is harder to access in the mental lexicon, or before a syntactically harder expression where the verbal process requires greater cognitive effort.

The relationship between speech pauses and emotions is the research topic of several approaches. Some researchers [15-18] have investigated the relationship between speech errors and emotions, and they count silent pauses as one type of error. Other so-called vocal emotion expression studies [19-21] investigate the presence of different emotions in speech or voice, and speech pauses appear as part of the speech parameters. A third approach is

linguistic conversation analysis, which has also been conducted in studies investigating the role of speech pauses during emotionally loaded situations [22].

The early studies that investigated the relationship between speech disfluencies (or speech errors) and anxiety found that in an anxious state, the number and length of pauses increase [15-18]. Later, Hofmann et al. [23], as well as Laukka et al. [24], analysed the temporal characteristics of the speech of social phobic patients with public speaking anxiety. According to Hofmann's study [23], social phobics produce longer speech pauses, pause more frequently, and spend more time pausing than do controls when giving a speech. Laukka et al. [24] compared speech samples from social phobic patients before and after pharmacological treatment. The patients' task was to give a speech about a vacation or travel experience before an audience of six to eight people. According to the results, the rate of silent pauses decreased in the speech samples of people who responded to the treatment, while in the case of non-responders, the pause ratio increased. In connection with speech tempo, the researcher did not find any significant differences.

Deppermann and Lucius-Hoene [22] analysed trauma narratives from TV shows using conversation analysis. They found that there were more pauses in speakers' speech while they were telling sad stories. Speakers even took longer pauses within syntactical units. In Fairbanks and Hoaglin's [19] study, a 27-word-long passage was spoken by six male actors, simulating five emotional states (contempt, anger, fear, grief, and indifference). The observation of durational features was the main focus of that study. According to the results, fear, indifference, and anger had the fastest tempo, followed by grief and contempt. Fear, indifference, and anger were characterised by a rapid rate, a short duration of phonations and short pauses with much variability, while in the cases of contempt and grief, a slower speech rate was measured. Contempt and grief, however, differed in terms of the cause of this slower rate. In the case of contempt, the slower speech rate was caused by prolongation of both phonations and pauses, and the rate of pauses was approximately the same as the three other emotions. However, the slow rate of grief was caused almost entirely by prolongation of pauses; thus, the pause ratio significantly differed from that of the other emotions. In the case of grief, the rate of pauses was greater than 50 percentages in the speech of three of the actors.

Jovicic et al. [20] established a Serbian emotional speech database. They asked actors to simulate emotional states while uttering words, sentences, and a passage. Speech duration, pause duration, and speech/pause ratio were studied and measured. Researchers found that

pause duration was more discriminative between emotions than speech duration and that the speech/pause ratio was a good indicator for speech rhythm and differentiated well between angry, happy, frightened, sad, and neutral emotional states. In Szabó's research [21], sad and happy emotional states were induced by using music and autobiographical recall tasks, and half-minute long speech samples were analysed. The results showed that in a sad emotional state, the average length of silent pauses and the rate of pauses to the whole speech increased compared to a happy emotional state.

The results of the above mentioned studies indicate that length and rate of speech pauses can change in different emotional states. However, most of these studies are connected by the emotional states of anxiety or fear, and little is known about other states such as happiness, sadness or anger.

1.3. Previous studies on vocal emotion recognition in foreign languages

Many studies have investigated emotion recognition in different cultures, most of which have focused on whether subjects belonging to different cultures can recognise expressions of emotion on the faces of others [25-30]. These studies demonstrated that although emotion expression is regulated by display rules, cultural norms, and individual habits, individuals are able to accurately recognise emotional displays of the face independent of different cultural backgrounds. Thus, these studies argue that emotional display involves universal principles.

There are fewer studies that have investigated how individuals from different linguistic and cultural backgrounds recognise one another's emotions from vocal expressions [31-36]. In these studies, generally the emotionality of sample utterances by lay people or professional actors is judged, and researchers compare how the accuracy in recognising emotions differs when the speaker and listener share cultural and linguistic background and when they do not.

In Beier and Zautra's study [31], American, Japanese and Polish subjects rated the emotionality of English sentences that were neutral based on content. Their results showed that when the judged expressions were short, Americans were better at emotion recognition than Japanese and Polish subjects. At the same time, by increasing the length of the expressions, the differences disappeared. Van Bezooijen et al. [32] asked Dutch, Japanese, and Taiwanese subjects to rate the emotionality of sentences simulated by Dutch actors. The

results were similar to those of the previous study: native speaking Dutch subjects performed significantly better than the other two groups, but Japanese and Taiwanese participants were also able to recognise the majority of emotions with above-chance accuracy, and the pattern of confusion was similar across judgement groups.

Later studies used so-called pseudo-utterances to filter out the effect of sentence meaning. The pseudo-utterances were made by connecting meaningless syllables from different languages in such a way that utterances that sounded like speech, but were semantically meaningless, were created. In a study by Scherer et al. [36], professional German actors simulated five emotions by uttering pseudo-sentences. Subjects from nine countries participated in the experiment, and their task was to judge the emotional content of the utterances on five emotion scales ranging from 0 to 6. Their results showed that German participants performed significantly better than the other language groups; thus, accuracy was highest when the speaker's and listener's language was the same. Further results showed that participants, whose language was more linguistically different from German, performed less accurately than those from a highly similar language. Thus, German subjects performed the best, followed by the Swiss-French, English, Dutch, American, Italian, French, and Spanish groups. The worst accuracy level was produced by the Indonesian subjects, but even this group performed with better than chance accuracy. Based on these findings, the authors formulated the "language distance hypothesis", which says that language similarity plays an important role in emotion decoding in a foreign language.

In another study conducted by Pell et al. [37], pseudo-utterances from native speakers of English, German, Hindi, and Arabic were rated by subjects from the same languages. Emotion recognition and acoustic patterns were analysed. The results showed that the mean, variance in fundamental frequency, and speech rate were responsible for 70–80 percent of the variance in the acoustic data across emotion types. The authors argue that these parameters determine the vocal expressions of basic emotions on a large scale, which might be a universal cue.

Based on previous studies [32, 35, 36, 38-40], we argue that listeners can accurately detect and categorise emotional states from speakers with a different language and cultural background. Processes of vocal emotion recognition involve universal principles; thus, the phenomenon is at least partly universal. However, some studies [31, 32, 36] have reported an in-group advantage, as vocal emotions that are simulated by speakers of the same culture and language are more accurately identified when compared to speakers of a different culture and

language. Therefore, beyond universality, social aspects and language-specific prosodic features are also important in communicating emotions [35, 41-44]. This phenomenon is described by the “language distance hypothesis” of Scherer and collaborators [36] and by the “cultural proximity” hypothesis of Elfenbein and Ambady [45] as well. According to this last hypothesis, people who are more similar in their cultural background can more accurately decode one another’s emotions than people whose cultures are different.

1.4. Cross-section of intercultural and inference studies

The research of Breitenstein et al. [10] is a good example of the cross-section of intercultural and inference studies. As mentioned previously, in Breitenstein and collaborators’ experiment, the speech rate and variance of the fundamental frequency of emotional sentences produced by a German actress were systematically manipulated. The emotionality of utterances was judged not only by German raters, but also by American raters. According to the results, the judgments of the two groups did not greatly differ, as despite the manipulation, subjects most frequently chose the emotion label that corresponded to the emotion category in which the speaker had originally spoken the utterance. The only exception was for the happy sentence, which American participants most frequently placed in the neutral category. Further analysis of the data showed that American raters were more influenced by tempo manipulations than Germans. For example, in the case of the happy sentence, German participants more frequently chose the happy category at faster rates, while American raters chose the angry or frightened answers more frequently. If the rate of angry sentences decreased, Americans were less accurate than the Germans and often responded with the neutral category. For frightened and neutral sentences at a slower rate, Americans more often answered with the sad category than did Germans. Manipulation in F0 variation influenced German and American subjects in the same way, and there were only small differences between the two groups.

In a study by Burkhardt et al. [46], French, German, Greek, and Turkish sentences were systematically manipulated with respect to pitch range, duration, and jitter. In a perception experiment, French, German, Greek, and Turkish subjects listened to the original and the manipulated utterances, each in the subjects’ native language, and rated how appropriate the manner of speaking was for the meaning of the sentence (e.g., happy, disgusted, etc.). The

results showed that regardless of the language, listeners interpreted the emotional load of the phrases as intended, with the exception of the bored phrases. As expected, phrases with the original prosody were generally judged as more appropriate than those that were manipulated. Furthermore, negative emotions were found to be better displayed by the speech synthesizer than the positive ones. As there were language differences, regarding what kind of pitch range, duration, or jitter manipulation had an effect on the emotion judgements, the authors argue that emotion simulation cannot work the same way in all languages.

1.5. The current study

The present study differs from previous studies in many respects. The so-called inference studies usually manipulate sentences that are uttered by actors [8-10, 47-50]. The advantage of using sentences uttered by actors is that it produces well controlled and high quality speech samples with intensive emotions, and changes can be well observed and measured. However, the disadvantage of this method is that these speech samples can differ in numerous ways (e. g. pausing, accentuation, speech errors) from those that occur in real life (see [1, 2]). Therefore, in our experiment we used natural speech samples, namely, samples from interviews from a Hungarian speech database [51] and from an experiment [21] in which speakers were volunteers and talked about themselves for a few minutes. The speech material sounded natural and lifelike compared to emotional expressions by actors. We selected one-minute long monologic sections to study speech pauses.

Several studies [9, 10, 50] have investigated the relationship between tempo manipulation and emotion ascribing, as well as modified articulation rate and speech rate⁸ together. In these studies, the articulation rate and possible pauses between sentence units, and thus speech rate, were both modified by lengthening or shortening the sentence. However, there are only few studies which examined the role of speech pauses in emotion expression (for an example see [3]), thus little is known about this topic. For investigating this phenomenon, utterances with

⁸ Speech tempo is defined as the number of production units (usually phonemes or syllables) per unit of time (usually seconds or minutes), where total speech time includes pauses; while articulation rate is defined as the number of production units per unit of time after subtracting pauses. (see 52.

Crystal, T.H. and A.S. House, *Articulation Rate and the Duration of Syllables and Stress Groups in Connected Speech*. Journal of the Acoustical Society of America, 1990. **88**(1): p. 101-112.)

naturally occurring pauses are needed. The present study used this kind of material and manipulated the pauses of them. In contrast to previous studies (e. g. [10]), speech tempo was modified only by manipulating silent pauses and by leaving the articulation rate fixed.

Thus, in this study, we investigated how listeners' attributions of the emotional state of speakers are affected by the systematic manipulation of the length of silent pauses in originally natural speech samples. In Experiment I, pauses in Hungarian speech samples were modified in four different ways, and variants were rated by listeners. Because, to our knowledge, there is no work that has investigated the relationship between speech pauses and emotions by using manipulated speech samples, we mostly based our hypotheses on the results of previous encoding and decoding studies. [1, 10, 19, 21-24]. According to our hypothesis, in the case of longer speech pauses, speakers would be rated as sadder, less happy, less angry, more scared, more disgusted, more surprised, less positive, and less heated.

In Experiment II, the same Hungarian experimental speech material was listened to and judged by Austrian subjects with a German mother tongue. Although Hungary and Austria are neighbouring countries, the German and Hungarian languages belong to different language families, and are very different concerning the vocables or the grammar. Thus for German speaking Austrians, Hungarian sentences were meaningless. By using this kind of stimuli, we had the possibility to investigate whether the role of the pauses in emotion ascribing is a language independent phenomenon or not. Based on previous intercultural vocal emotion recognition studies [32, 35, 36, 39] and the cultural proximity hypothesis [45], we assumed similarities in the relationship between speech pauses and emotions in Hungarian and German. Thus, we hypothesised that if native German speaking subjects listen to the Hungarian speech samples, pause manipulation would have a similar effect on trends in emotion ratings, as in the case of the Hungarians. Thus, in the case of pause elongation, speakers would be rated as sadder, less happy, less angry, more scared, more disgusted, more surprised, less positive, and less heated.

2. Material and methods

2.1. Experimental material

In the experiment, we used five emotionally neutral speech samples that originated from the Hungarian Speech Database [51] and a previous study we had conducted [21]. On the recordings, five female volunteers talked about their schools, jobs or an ordinary day in their lives. The speech samples were approximately one minute long (49.1-64.4 seconds, average: 55.8 seconds). The content was emotionally neutral and did not contain emotional words (such as “it was a very happy part of my life”) or laughter. The speech samples were all monologues and were extracted from longer interviews. Here we give an example where speaker A is talking about an ordinary day of her. Already present speech pauses are marked by **<pause>**.

“Reggel általában olyan 7 óra felé kelek fel **<pause>** utána rögtön az az első, hogy elkészítem a reggelit **<pause>** tehát az a legfontosabb, a reggel fénypontja, **<pause>** és arra szánok a legtöbb időt is, tehát legalább fél óra kell, hogy én reggelizzek, és utána jöhet a többi készülődés, rendet rakás satöbbi. **<pause>** De ugye a nap nagy részét a suliban töltöm, **<pause>** az órák mellett gyakran el szoktam menni könyvtárba **<pause>** kiegészíteni a **<pause>** tanulmányaimat... **<pause>** Ööö, **<pause>** illetve **<pause>** Ööö **<pause>** Igen **<pause>** És az iskola után pedig **<pause>** bevásárolni igen gyakran szoktam menni.. **<pause>**. Utána, **<pause>** Ööö **<pause>** miután hazaérkeztem, **<pause>** megfőzök másnapra, **<pause>** de ez nem mindig így van, tehát hetente kétszer **<pause>** szokott ez lenni. Aztán hogy ha kell valamit készülnöm másnapra, azt is megteszem, **<pause>** és utána pedig vagy a párommal elmegyünk valahová **<pause>**, vagy éppen táncolni, amit említettem, vagy moziba **<pause>**, vagy egyszerűen csak tévézünk és pihenünk otthon.”

The English translation of this section (literal translation, retaining much of the Hungarian word order, grammar, and disfluencies):

In the mornings I usually get up about 7 o'clock. Then first I make my breakfast, this is the most important part, the highlight of the morning, so I spend the most time on it, therefore I need at least half an hour to have my breakfast, and only then come the other preparations like straighten up and so on... But I spend most of the day at school, after class I often go to the library, to complete my studies, um, well, um, yes. And, after school I often go for shopping, and then, um, after I get home, I cook for the next day, but it is not always so, but twice a week. Then if I have to prepare something for the next day, I do it and then we go out with my boyfriend, or go dance what I have already mentioned, or we go to the cinema or we just watch TV and have a rest at home.

These approximately one-minute long speech samples were taken as “original”, and each was modified in four different ways. Using the program Audacity, we manually measured pauses that were longer than 100 milliseconds (based on the Hungarian tradition [52] [53] and international studies [54] [55]) and performed the following four manipulations to each: lengthening by 18 %, lengthening by 50 %, shortening by 21 %, and shortening by 50 %. Only one manipulation was performed on each sample. For example, variations of speech samples with 21 % shortened pauses were created so that all of the pauses that were longer than 100 milliseconds in the original speech samples were shortened by 21 %.

Modifications with 21 and 18 % were based on our previous results [21] where music and autobiographical recall were used to induce sad and happy emotional states. According to the results, the durations of silent pauses were 18 % longer in a sad emotional state and 21 % shorter in a happy emotional state compared to the neutral state. Based on observations in a pre-test, we also increased modifications by 50 %.

Manipulations resulted in five variants for each of the five speech samples: variants with pauses elongated by 18 or 50 %, variants with pauses shortened by 21 or 50 %, and the original version. Thus, pauses became 50, 79, 100 (original), 118, and 150 % of the original versions (resulting in the following variants: 0.5, 0.7, 1 (original), 1.18, and 1.5.). (The description of stimuli is presented in Tables 1 and 2.)

Speaker	Mean fundamental frequency (SD) in Hz	Mean intensity (SD) in dB	Number of pauses
A	204.5 (8.3)	64.2 (0.7)	22
B	188.0 (20.4)	59.1 (2.1)	20
C	144.1 (19.7)	63.5 (1.7)	22
D	135.8 (11.3)	62.8 (1.4)	22
E	140.3 (9.4)	64.1 (1.3)	16

Table 1. *Mean fundamental frequency, mean intensity, and the number of pauses of the original material.*

Speaker	Length of sound file (sec)					Average length of pauses (msec)					Total length of pauses (sec)					Pause ratio (%)				
	Pause modification					Pause modification					Pause modification					Pause modification				
	*0.5	*0.79	*1	*1.18	*1.55	*0.5	*0.79	*1	*1.18	*1.55	*0.5	*0.79	*1	*1.18	*1.55	*0.5	*0.79	*1	*1.18	*1.55
A	52.4	55.7	58.0	60.0	63.5	253	402	506	597	759	5.6	8.8	11.1	13.1	16.7	10.6	15.9	19.2	21.9	26.3
B	45.0	48.6	51.1	53.3	57.2	306	486	612	722	918	6.1	9.7	12.2	14.4	18.4	13.6	20.0	24.0	27.1	32.1
C	59.1	62.4	64.4	66.6	70.1	250	396	499	589	749	5.5	8.7	11.0	13.0	16.5	9.3	14.0	17.0	19.5	23.5
D	50.7	54.0	56.3	58.2	61.6	250	397	500	590	750	5.5	8.7	11.0	13.0	16.5	10.8	16.2	19.6	22.3	26.8
E	44.3	47.2	49.1	50.9	53.8	302	480	605	714	907	4.8	7.7	9.7	11.5	14.5	10.9	16.3	19.7	22.4	26.9

Table 2. Length of the sound files, average length of pauses, total length of pauses, and pause ratio of the sound files across all 5 samples for each speaker

To assess the naturalness of the materials, we conducted an additional rating study, where 75 university students (21 male and 54 female) with a mean age of 22.86 ($SD = 4.50$) differed from those who participated in the experiment, rated the naturalness of the sound files. In the same arrangement as in the experiment (see Table 3), five groups of participants listened to the speech samples and subjects were asked to rate how natural the sound files sounded on a scale of 1-6 (1: very artificial sounding, 6: absolutely natural sounding). Participants rated the naturalness of the sound files for an average of 4.66 ($SD = 1.17$), indicating that the material sounded quite natural and pause manipulation did not make the sounding artificial.

2.2. Participants

2.2.1. Hungarian sample

Fifty Hungarian raters (16 male and 34 female) participated in the first experiment, all of whom were students of the Budapest University of Technology and Economics. Participants were studying engineering or communications. They had a mean age of 20.12 years ($SD = 1.41$). All of these participants were native Hungarian speakers. Participants received university credits for their participation.

2.2.2. Austrian sample

Thirty-eight (8 male and 30 female) native German speaking Austrian subjects participated in the second experiment. Participants were all students of the Linguistic Institute at the University of Vienna, but twelve of them also studied other majors, such as history, pedagogy or languages. They had a mean age of 27.58 years ($SD = 12.00$). German was their native language, but three of them also spoke additional native languages. One subject spoke German and English, one spoke German and Dutch, and one spoke German, Polish, and Russian. With the exception of one participant, every subject spoke at least one foreign language. On average, participants spoke 2.74 foreign languages ($SD = 1.14$), but none of them understood or spoke Hungarian. Participation was voluntary, and subjects did not receive any compensation for their participation.

2.3. Procedure

Before the experiment began, participants were informed that the relationship between certain emotions and certain speech parameters would be studied. Participants were asked to judge five speech samples by filling out a questionnaire. After listening to each speech sample, participants were asked to rate how angry, sad, disgusted, happy, surprised, scared, positive, and heated the speaker seemed on a scale of 1-6 (not at all – very much). In addition to the basic emotions, *positive* and *heated* labels were believed to signal the valence and arousal dimensions. Austrian raters received questionnaires in German (the translation of the expressions was: *der Sprecher war: wütend, traurig, angeekelt, froh, überrascht, hatte Furcht, war positiv, war erregt*). Hungarian participants also rated the neutrality of the *content* of the speech samples on a scale of 1-6 (1: not at all emotionally-neutral content, 6: completely emotionally-neutral content). Thus, we tried to emphasise the fact that content and emotions can be independent of one another. In all cases, we emphasised that participants should rate the entire speech sample and should not decide too early in the utterance, and questionnaires must only be completed after listening to the full sample during the minute long pause between speech samples. Participants were instructed not to think too much about ratings, but to decide spontaneously.

One-minute long speech pauses were provided for answering the questions, and this time was intended to delete the “effect” of the speech that had been previously listened to so that participants had less opportunity to compare speech samples. Accordingly, listeners had to wait until the end of the one minute pauses. They were not allowed to fast-forward, even when they had answered all of the questions. All listeners participated in the experiment individually, and the speech samples were played on a computer via a headset. They answered the questions on a paper form questionnaire. The volume of the speech samples was the same for all participants, and the subjects were not allowed to adjust it. We previously tested the appropriateness of the loudness of the sound files, based on the opinion of several persons. During the experiments all participants reported that they could hear and understand the utterances well.

2.4. Experimental arrangement

In both of the experiments, we had five experimental conditions. Each participant judged each speaker only once, and all of the speech samples he or she listened to were modified with a different pause ratio (experimental arrangement is seen in Table 3). Participants belonging to the same group listened to speech samples according to a “Balanced Latin Square” arrangement [56], which limits the risk of carryover effects. With this method we aimed to avoid having one person listen to the same speech more than once.

Group A	Group B	Group C	Group D	Group E
a 0.5	a 1.5	a 1.18	a 1	a 0.79
b 0.79	b 0.5	b 1.5	b 1.18	b 1
c 1	c 0.79	c 0.5	c 1.5	c 1.18
d 1.18	d 1	d 0.79	d 0.5	d 1.5
e 1.5	e 1.18	e 1	e 0.79	e 0.5

Table 3. *Experimental arrangement. Letters (a-e) indicate the speech samples, while numbers indicate the rate of pause modification. Participants in group A listened to speech a with 50 % abbreviated pauses (a0.5), speech b with 21 % abbreviated pauses (b0.79), speech c in the original version (c1) etc.*

3. Results of the first experiment: ratings of Hungarian participants

3.1. Differences between speakers

To evaluate whether speakers were rated as different regarding the emotionality of their speech’s content, a univariate analysis of variance (ANOVA) using the GLM procedure was conducted with the original, non-manipulated speech samples of the speakers (5: A, B, C, D, E) as independent variable and the ratings on the *content neutrality* scale (1-6, where 1: not at all emotionally neutral, 6: completely emotionally neutral) as dependent variable. (We selected ratings on original speech samples from all of the answers.) The content of the original speech samples was rated by the Hungarian participants between 2.8 and 4.8, and

ANOVA revealed a trend level result (main effect of speaker: $F(4, 45) = 2.355$, $p = .068$, $\eta_p^2 = .173$), indicating that the five original speech samples were slightly different regarding the neutrality of their content.

In an analysis with the pause length manipulation (5 levels: 0.5, 0.79, 1, 1.18 vs. 1.5) as the independent variable and *content neutrality* ratings (1-6, where 1: not at all emotionally neutral, 6: completely emotionally neutral) as the dependent variable, we found that Hungarian participants rated the content of the speech samples between 3.64 and 4.24, and there was no significant difference between the five speech samples (main effect of pause length: ANOVA; $F(4, 245) = 1.07$, $p = .37$, $\eta_p^2 = .017$). In conclusion, pause manipulation did not have an effect on content neutrality ratings.

We also analysed whether original samples were different regarding emotion ratings in an ANOVA with the non-manipulated speech samples of the speakers (5: A, B, C, D, E) as independent variable and ratings (1-6) on emotion scales (angry, sad, disgusted, happy, surprised, scared, positive, and heated) as the dependent variables. (Again, we selected ratings on original speech samples from all of the answers.) Separate ANOVAs for each emotion scale yielded significant differences on disgusted ($F(4, 45) = 3.59$, $p = .013$, $\eta_p^2 = .242$), happy ($F(4, 45) = 10.27$, $p < .001$, $\eta_p^2 = .477$), positive ($F(4, 45) = 7.44$, $p < .001$, $\eta_p^2 = .398$), and heated ($F(4, 45) = 5.28$, $p = .001$, $\eta_p^2 = .32$) ratings. Post-hoc LSD pairwise comparisons indicated that speaker A was rated as happier and more positive than other speakers and was less disgusted than speaker C. Speaker E was rated as the most heated and was significantly different from speakers A, B, C, and D. Again, speaker E was rated as more disgusted than speakers A, B and D (the exact p values are presented in Table 4). All in all, the original speech samples were not totally neutral from the perspective of content and emotional load, and listeners credited different emotional characteristics to them.

Emotion scale	Speaker	Direction of the relationship	Speaker	p value
Disgusted	C	>	A	.026
		>	A	.003
		>	B	.017
		>	D	.011
Happy	A	>	B	<.001
		>	C	<.001
		>	D	<.001
		>	E	<.001
Positive	A	>	B	<.001
		>	C	.001
		>	D	.001
		>	E	<.001
Heated	E	>	A	.001
		>	B	.008
		>	C	.002
		>	D	<.001

Table 4. *Differences between speakers: significant effects on certain emotion scales, the direction of the relationship, and exact p values from significant LSD multiple comparisons (Hungarians' ratings)*

3.2. Relationship between pause modification and emotion judgments

To examine our research question, we investigated how pause conditions are related to judgments on the emotion scales. We performed univariate ANOVAs using the GLM procedure, and tested differences in mean ratings between the five pause modifications separately for each emotion scale (angry, sad, disgusted, happy, surprised, scared, positive, and heated). For all ANOVAs, the independent variable was pause length modification (5 levels: 0.5, 0.79, 1, 1.18 vs. 1.5), while the dependent variables were the ratings (1-6) given on the emotion scales. We used an alpha level of .05 for all ANOVAs and LSD post-hoc tests.

The ANOVA yielded statistically significant differences on the sad scale and trend level differences on the happy scale (the results of the ANOVAs and mean ratings are presented in Table 5).

Emotion scale	Mean ratings on emotion scales (SD)					<i>Df</i> <i>effect</i>	<i>Df</i> <i>error</i>	<i>F</i>	<i>p</i>	η_p^2
	Pause manipulation									
	0.5	0.79	1	1.18	1.5					
Angry	1.46 (.9)	1.46 (.9)	1.68 (1.2)	1.62 (1.0)	1.50 (0.9)	4	245	.49	.743	.008
Sad	2.00 (1.2)	2.2 (1.3)	2.32 (1.4)	2.58 (1.6)	2.96 (1.6)	4	244	3.33	.011*	.052
Disgusted	1.62 (1.1)	1.82 (1.3)	2.16 (1.4)	2.14 (1.2)	2.12 (1.3)	4	245	1.79	.131	.028
Happy	2.86 (1.4)	2.6 (1.6)	2.46 (1.4)	2.3 (1.3)	2.1 (1.3)	4	245	2.01	.094	.032
Surprised	1.66 (1.1)	1.70 (1.1)	1.44 (1.0)	1.78 (1.3)	1.90 (1.3)	4	245	1.07	.373	.017
Scared	1.52 (0.9)	1.58 (1.2)	1.56 (1.0)	1.82 (1.3)	2.08 (1.6)	4	245	1.87	.117	.030
Positive	3.28 (1.5)	3.14 (1.6)	3.06 (1.5)	2.84 (1.5)	2.62 (1.4)	4	245	1.52	.198	.024
Heated	2.24 (1.3)	2.14 (1.3)	2.10 (1.3)	2.02 (1.2)	1.68 (.9)	4	245	1.59	.178	.025

Table 5. *Mean ratings and standard deviations of Hungarian participants on emotion scales, regarding pause manipulation and ANOVA Summary*

* ANOVA is significant at the 0.05 level

Post- hoc tests were used to investigate these differences in greater detail. LSD pairwise comparisons indicated that variants 1.5 were rated as significantly sadder than variants 0.5, 0.79, and 1; and variants 1.18 were rated as significantly sadder than variants 0.5. For the happy scale, there was a significant difference between variants 0.5 and variants 1.5, indicating that utterances with shorter pauses were rated as happier (the exact p values are presented in Table 6).

	Variant	Direction of the relationship	Variant t	p value
Sad	1.18	>	0.5	.044
	1.5	>	0.5	.001
		>	0.79	.009
		>	1	.027
Happy	0.5	>	1.5	.008

Table 6. *Differences in sad and happy emotion ratings regarding pause manipulation: the direction of the relationship and exact p values from significant LSD multiple comparisons (Hungarians' ratings)*

4. Results of the second experiment: ratings of Austrian participants

4.1. Differences between speakers

As with the Hungarian participants, we analysed whether Austrian raters judged Hungarian speakers' emotions differently. Univariate ANOVAs with original speech samples of the speakers (5: A, B, C, D, E) as independent variable and ratings (1-6) on each emotion scales (angry, sad, disgusted, happy, surprised, scared, positive, and heated) as dependent variable were conducted. (We selected ratings on original speech samples from all of the answers.) ANOVA yielded statistically significant differences for the sad ($F(4, 33) = 5.04, p = .003, \eta_p^2 = .379$), disgusted ($F(4, 32) = 5.23, p = .002, \eta_p^2 = .395$), happy ($F(4, 32) = 14.73, p < .001, \eta_p^2 = .648$), scared ($F(4, 32) = 2.86, p = .039, \eta_p^2 = .263$), and positive ($F(4, 32) = 15.32, p < .001, \eta_p^2 = .657$) scales.

Post-hoc LSD pairwise comparisons indicated that the Austrians considered speakers A and E as happier and more positive than speakers B, C, and D, and speaker A was rated as even more positive and happier than speaker E. In line with this finding, A and E were rated as less sad than speakers B, C, and D. Speaker B was rated as more disgusted than other speakers and more scared than speakers A and E (the exact p values are presented in Table 7). In sum, Austrian participants ascribed different emotional characteristics to the original speech samples.

Emotion scale	Speaker	Direction of the relationship	Speaker	p value
Sad	A	<	B	.007
		<	C	.003
		<	D	.017
	E	<	B	.007
		<	C	.003
Disgusted	B	<	D	.017
		>	A	<.001
		>	C	.022
	A	>	D	.001
		>	E	.002
Happy	A	>	B	<.001
		>	C	<.001
		>	D	<.001
	E	>	E	.016
		>	B	.002
Positive	A	>	C	.001
		>	D	.001
		>	B	<.001
	E	>	C	<.001
		>	D	<.001
Scared	B	>	E	.008
		>	B	.001
		>	C	.001
	A	>	D	.007
		>	A	.010
		>	E	.007

Table 7. *Differences between speakers: significant effects on certain emotion scales, the direction of the relationship, and exact p values from significant LSD multiple comparisons (Austrians ratings)*

4.2. Relationship between pause modification and emotion judgments

A univariate ANOVAs with pause length modification (5 levels: 0.5, 0.79, 1, 1.18 vs. 1.5) as the independent variable and the Austrians' ratings (1-6) on the emotion scales (angry, sad, disgusted, happy, surprised, scared, positive, and heated) as the dependent variables yielded no statistically significant effects. Trend level differences for the happy and positive scales were observed (the results of the ANOVAs and mean ratings are presented in Table 8).

Emotion scale	Mean ratings on emotion scales (SD)					<i>Df</i> <i>effect</i>	<i>Df</i> <i>error</i>	<i>F</i>	<i>p</i>	η_p^2
	Pause manipulation									
	0.5	0.79	1	1.18	1.5					
Angry	1.95 (1.2)	2.00 (0.9)	1.89 (1.2)	1.76 (0.9)	2.08 (1.2)	4	184	0.448	.773	.010
Sad	2.74 (1.7)	2.63 (1.6)	2.71 (1.7)	3.18 (1.8)	3.47 (1.8)	4	185	1.738	.143	.036
Disgusted	1.82 (1.3)	2.16 (1.4)	1.97 (1.3)	1.74 (1.1)	1.87 (0.9)	4	184	0.713	.584	.015
Happy	2.53 (1.5)	2.37 (1.3)	2.65 (1.7)	2.47 (1.6)	1.74 (1.4)	4	184	2.14	.077	.045
Surprised	2.13 (1.4)	2.00 (1.1)	2.27 (1.4)	2.29 (1.4)	2.18 (1.3)	4	184	0.291	.883	.006
Scared	2.11 (1.4)	2.00 (1.3)	2.11 (1.4)	2.32 (1.4)	2.79 (1.7)	4	184	1.859	.120	.039
Positive	2.87 (1.6)	2.82 (1.6)	2.97 (1.7)	2.97 (1.7)	2.05 (1.6)	4	184	2.16	.076	.045
Heated	3.21 (1.4)	2.97 (1.3)	3.14 (1.5)	2.63 (1.2)	3.11 (1.4)	4	183	1.052	.382	.022

Table 8. Mean ratings and standard deviations of Austrian participants on emotion scales, regarding pause manipulation, and ANOVA Summary

To explore the differences in more detail, post-hoc tests were made. LSD pairwise comparisons indicated that variants 1.5 were rated as significantly less happy than variants 0.5, 1, and 1.18, and variants 1.5 were rated as significantly less positive than all the other variants (the exact *p* values are presented in Table 9). The results indicate that utterances with shorter pauses were rated as happier and more positive.

	Variant	Direction of the relationship	Variant t	p value
Happy	1.5	<	0.5	.023
		<	1	.009
		<	1.18	.034
Positive	1.5	<	0.5	.030
		<	0.79	.042
		<	1	.015
		<	1.18	.015

Table 9. Differences in happy and positive emotion ratings regarding pause manipulation: the direction of the relationship and exact p values from significant LSD multiple comparisons (Austrians' ratings)

5. Comparing the ratings of the two groups of participants

To test whether there is a difference between the two groups regarding emotion judgments or there is an interaction between the independent variables, we performed a MANOVA with language (2 levels: Hungarian, German), pause modification (5 levels: 0.5, 0.79, 1, 1.18 vs. 1.5), and speaker (5 levels: A, B, C, D, E) as independent variables, and ratings (1-6) on emotion scales (angry, sad, disgusted, happy, surprised, scared, positive, and heated) as dependent variables. MANOVA revealed significant main effect of language ($F(8, 378) = 14.13, p < .001, \eta_p^2 = .230$), pause modification $F(32, 1524) = 1.47, p < .05, \eta_p^2 = .030$, and speaker ($F(32, 1524) = 8.38, p < .001, \eta_p^2 = .150$). In addition, there was an interaction of language and speaker ($F(32, 1524) = 3.01, p < .001, \eta_p^2 = .059$). No other interactions were significant.

The effect of language was significant for the angry ($F(1, 385) = 17.11, p < .001, \eta_p^2 = .043$), the sad ($F(1, 385) = 15.61, p < .001, \eta_p^2 = .039$), the surprised ($F(1, 385) = 17.13, p < .001, \eta_p^2 = .043$), the scared ($F(1, 385) = 22.28, p < .001, \eta_p^2 = .055$), the positive ($F(1, 385) = 4.07, p < .05, \eta_p^2 = .010$) and the heated ($F(1, 385) = 73.45, p < .001, \eta_p^2 = .16$) scales. Austrian participants rated Hungarian speakers as angrier, sadder, more surprised, more scared, less positive, and more heated than did Hungarian raters (see Figure 1).

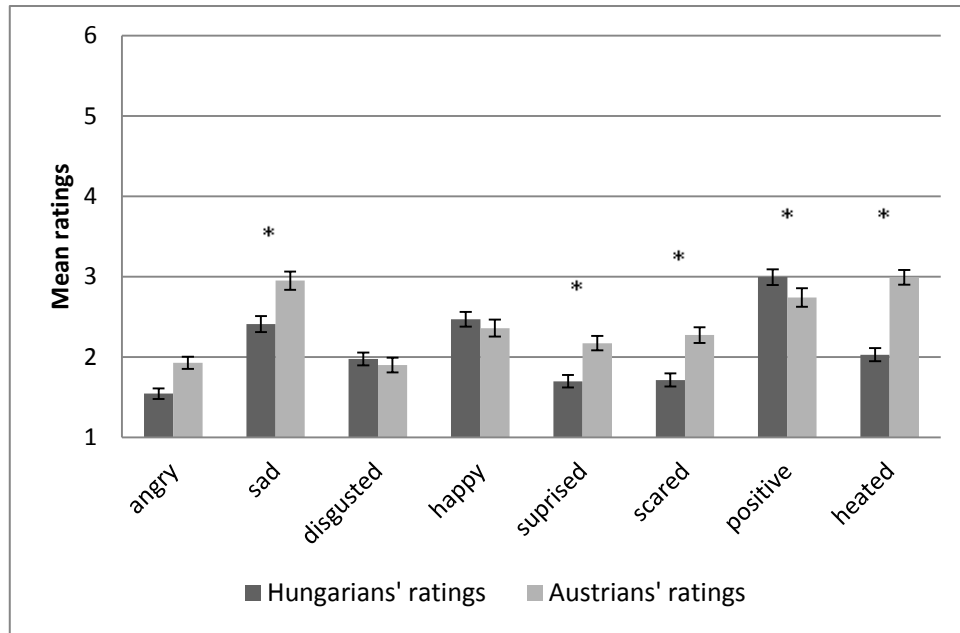


Figure 1. Hungarians' and Austrians' mean ratings on the emotion scales across all 5 samples for each speaker (significant differences are marked by *) Error bars represent the standard error.

The main effect of pause modification was significant for the sad ($F(4, 385) = 4.71, p < .05, \eta_p^2 = .047$), the happy ($F(4, 385) = 4.03, p < .05, \eta_p^2 = .040$), the scared ($F(4, 385) = 4.28, p < .05, \eta_p^2 = .043$), and the positive ($F(4, 385) = 3.16, p < .05, \eta_p^2 = .032$) scales. Post hoc LSD pairwise comparisons indicated that variants 1.5 were rated as significantly sadder than variants 0.5, 0.79, and 1; and variants 1.18 were rated as significantly sadder than variants 0.5, 0.79. Variants 1.5 were rated as significantly less happy and less positive than the other variants. Variants 1.5 were rated as more scared than the other variants of pause modification. In sum, utterances with longer pauses were rated as sadder, less happy, more scared and less positive (the exact p values are presented in Table 10).

Emotion scale	Variant	Direction of the relationship	Variant t	p value
Sad	1.18	>	0.5	.019
		>	0.79	.036
	1.5	>	0.5	<.001
		>	0.79	<.001
Happy	1.5	>	1	.001
		<	0.5	<.001
		<	0.79	.004
		<	1	.002
Scared	1.5	<	1.18	.020
		>	0.5	.001
		>	0.79	.001
		>	1	.001
Positive	1.5	>	1.18	.042
		<	0.5	.001
		<	0.79	.004
		<	1	.002
		<	1.18	.018

Table 10. *Differences in emotion ratings regarding pause manipulation: the direction of the relationship and exact p values from significant LSD multiple comparisons (results from the Hungarian and the Austrian raters together)*

Given that there was no interaction between language and pause modification, the results indicate that pause length plays a similar role for the Hungarian and the German speaking raters in ascribing sad, happy, scared and positive emotional states, and longer pauses indicate sadder, less happy, more scared and less positive emotional states. (Mean ratings of the Hungarians and the Austrians are seen in Figure 2.)

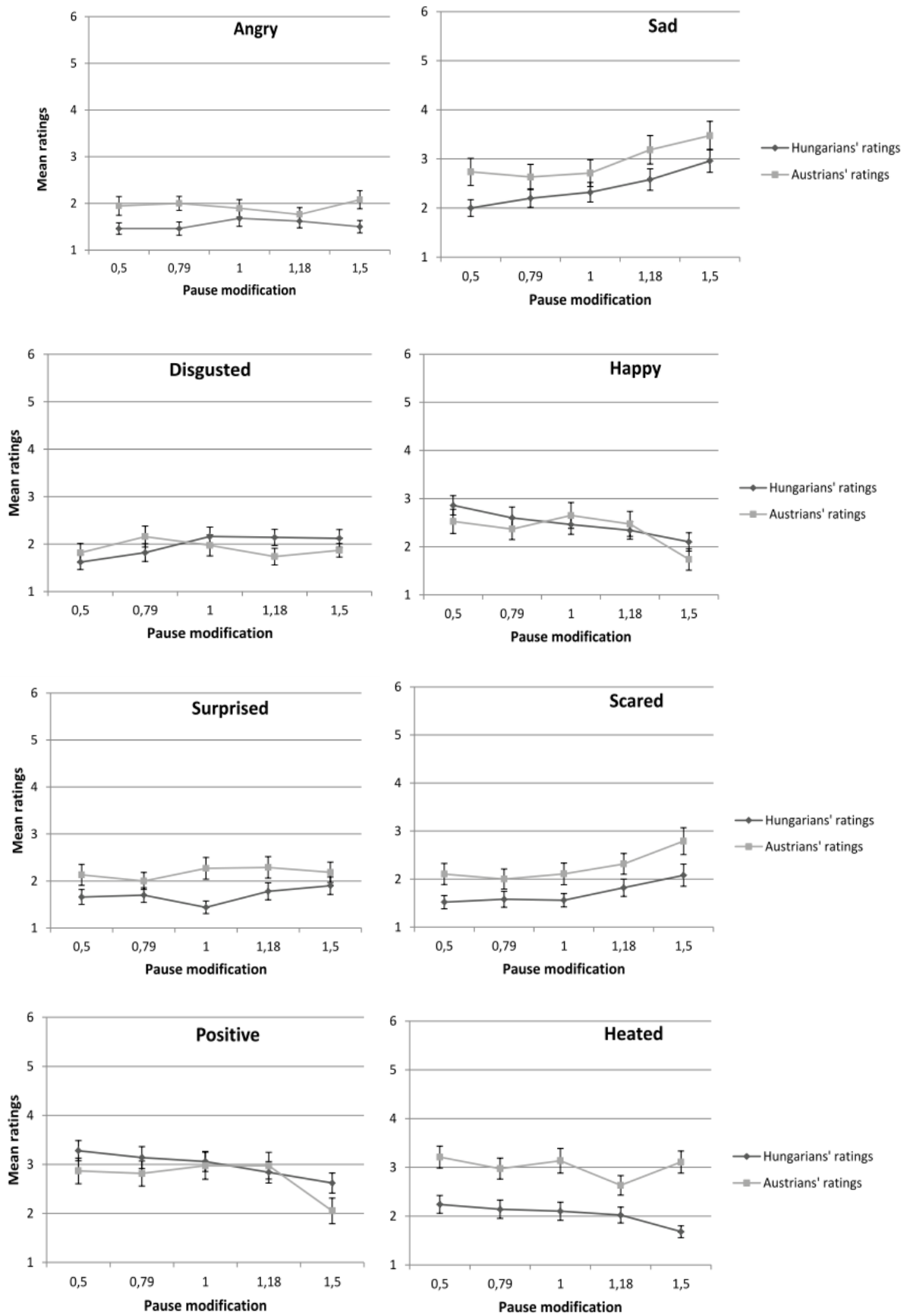


Figure 2. Mean ratings of Hungarian and Austrian participants on the emotion scales, divided by pause modification. Error bars represent the standard error.

The effect of speaker was analysed more detailed separately for the two language groups (see section 3.1 and 4.1). The speaker and language interaction was significant on the angry ($F(4, 385) = 4.21, p < .05, \eta_p^2 = .042$), sad ($F(4, 385) = 5.32, p < .001, \eta_p^2 = .052$), disgusted ($F(4, 385) = 5.6, p < .001, \eta_p^2 = .055$), scared ($F(4, 385) = 5.67, p < .001, \eta_p^2 = .056$), positive ($F(4, 385) = 5.33, p < .001, \eta_p^2 = .052$), and heated scales ($F(4, 385) = 4.61, p < .05, \eta_p^2 = .046$). Post hoc LSD pairwise comparisons indicated that Austrians rated the speakers as angrier, sadder, more surprised, more scared, less positive, and more heated, in the case of speaker E and emotions angry, sad, and positive, ratings were in the opposite direction. Moreover, speaker A and E were rated as similarly scared, and E was rated as similarly heated by the two groups ($ps < .05$).

6. Discussion

The aim of the study was to investigate how changes in only one parameter influence the emotion judgments of listeners. Trends in the data showed that both in the Hungarian and Austrian groups, the length of silent pauses influenced listeners in attributing emotional states to the speaker. In the case of the Hungarians, the effect of pause length was most evident on the sad scale, and pause length manipulation had a significant effect on emotion ratings. The same speech samples were rated as sadder when pauses were longer. There was a smaller, trend level effect on the happy scale, indicating that listeners rate slightly differently the same speech samples with different pause length in generally, variants that contained shorter pauses were rated as significantly happier than variants that contained longer pauses. In the case of the German speaking raters, the pause length manipulation had a general trend level effect on the happy and positive scales. Those variants which pauses were elongated by 50 percentages were rated as significantly less happy than those variants where the pauses were shortened by 50 percentages or were elongated by 18 percentages or the original versions. The variants, containing the 50 percentages elongated pauses were rated as significantly less positive than all the other variants.

Combining the results from the Hungarian and the Austrian raters, we found that pause length modification has a significant main effect on the sad, happy, scared, and positive scales. Given that there was no interaction between language and pause modification, the

results indicate that pause length plays a similar role for the Hungarian and the German speaking raters in ascribing sad, happy, scared and positive emotional states; longer pauses indicate sadder, less happy, more scared and less positive emotional state. The strongest effect was observed in the case of the sad emotion by the Hungarians. The results are consistent with Breitenstein et al.'s [10], Fairbanks and Hoaglin's [19] and Szabó's [21] studies which found that speech rate and speech pauses are strongly related to happy and sad emotional states.

With respect to the fear ratings, we found that with the increase of pause length, speakers were rated as being more scared. As fear and anxiety are interconnected emotions [57], our result can be related to previous studies on anxiety [15-18, 23, 24], which found that in speech during anxious situations or in the speech of social phobics, the number of pauses increased. Our results seem to contradict those of Breitenstein et al. [10], who found that with the increase in speech rate, speakers were rated as more scared. However, in Breitenstein's experiment, the tempo was modified by slowing down and speeding up sentences in a supposedly continuous way while also modifying the articulation rate and speech rate. In our case, tempo modification was the result of just manipulating pause length and leaving articulation fixed. Thus, the results of the two studies do not necessarily contradict one another, but rather they are compatible: fear/anxiety causes a faster articulation rate [19] *and*, at the same time, longer pauses in speech. When these speech features are perceived, fear is ascribed to the speaker. Therefore, according to the results speech rate and the speech/pause ratio seem to be related to happy-sad (valence dimension, see [57]) and scared emotional states. When we are happy, sad, or scared, we show our emotional state to listeners by modifying the pause length in our speech. Furthermore, when with a person we are speaking to perceives changes in the speech/pause ratio, he or she will ascribe these emotional states to us. In the case of the other emotions, there was no observable effect of the pause manipulation. Probably, speakers tend to express anger, surprise, disgust, and heatedness by using shorter utterances and pauses play no role.

Speech samples were produced by Hungarian volunteers and not by actors, as the role of the pauses in natural speech was the focus of the study. Though the original speech samples were not totally neutral regarding their content or emotional load, the experimental arrangement (Latin square and balanced Latin square, eliminating order effect) allowed us to use pause length modification as the only independent variable. Thus, the results of the statistical analysis indicated that differences in emotion judgments were only affected by pause length modification. However, we have to stress it here that when interpreting the

results, it has to be taken into account that effect sizes were small and the LSD post-hoc tests used in this experiment are considered as quite liberal.

The fact that there was no big difference between the emotion ratings that were given on certain variants is not surprising, as the topic of all speech samples was neutral, and intonation, loudness, articulation rate and other speech parameters corresponded with this emotionally neutral content. In addition, the rate of pause modification was not high. When listening to the variants of the same sound file one after another, differences are hard to detect. Furthermore, in the additional rating study, participants rated the sound files for quite natural sounding and none of the participants mentioned that the speech samples were artificial or strange.

In the experimental instructions, we emphasised to both groups that the task was to judge the entire speech sample because we assumed that modifications in speech pauses only have an effect during longer passages. Manipulating the length of speech pauses is not only interesting in itself, but also determines the speech-pause ratio. Thus, not only the length of pauses followed by one another, but also the ratio of pauses to the whole speech is what listeners perceive and what influences emotion ascribing.

Hungarian and Austrian subjects participated in our research. Previous intercultural/interlinguistic emotion recognition studies [32, 35, 36, 38-40] have found that people can identify emotions from foreign languages with better than chance accuracy. At the same time, individuals are more accurate with regard to sentences uttered in their native language than those in a foreign language [31, 32, 36]. Based on our results, emotion ascribing works in a similar manner for Hungarians and Austrians with a German native language when the length and rate of speech pauses change. In our opinion, the relationship between speech pauses and emotions might be universal, but more studies are needed to confirm this hypothesis. A good theory to explain the similar results is the Elfenbein-Ambady cultural proximity hypothesis [45], which argues that people from similar cultures are similar in coding emotions, and thus decode one another's emotions more easily than people from more distant cultures. Hungarian and Austrian cultures were often interconnected throughout history, and because they are neighbouring countries, they are also geographically close to one another. Thus, although the German and Hungarian languages are very different (they belong not even into the same language family), the function of speech pauses in emotion expression might be the same, and longer pauses shift answers to the *sadder* and *less happy*, *less positive* and *more scared* direction.

At the same time, we found differences between the two groups. First, the two groups differed in the emotion judgments of speech samples. Based on the statistical analysis, Austrian participants rated the Hungarian speakers as angrier, sadder, more surprised, more scared, less positive, and more heated than did Hungarian raters. The results cannot be connected with the valence or the arousal dimensions, as although speakers generally were rated by the Austrians as less positive, but one exception is the surprised scale, and more heated yet again, an exception is the sad scale. (see [57]). It is conceivable that the results were affected by language differences. Thus, speech parameters such as loudness, pitch, and melody and the common appearance of these factors made the Austrians raters, compared to the Hungarian ones, feel more like that the Hungarian speakers were angrier, sadder, more surprised, scared, negative, and heated. Of course, results might also have been influenced by the fact that the sentences were meaningful for the Hungarian raters, while Austrians could only base their rating on the vocal and temporal features.

Speakers' voices were judged differently by Hungarian and Austrian participants as well, and for several speakers, higher ratings were given on certain emotion scales. Judgments of the two groups were only partly similar, as speaker A was rated as happier and as more positive than the other speakers. As speaker A had the highest pitch and high pitch reveals positive emotional states no matter of culture, is likely that this effect is responsible for this similarity. In sum, individual characteristics of speech and the linguistically and culturally determined way of emotion expression fundamentally influence the attributes and emotional states we ascribe to speakers.

Our work highlights the role of speech pauses in emotion expression and recognition and stresses the importance of the topic. Our results show that speech pauses play an important role in emotion ascribing, which might be a regional/cultural, if not universal, phenomenon. To confirm this assertion, an "inversion" of the experiment would be needed, or in other words, manipulating pauses in German speech samples and getting them rated by subjects with a German and Hungarian mother tongue.

Acknowledgements

We thank Florian Menz, professor at the Linguistic Institute, the University of Vienna for helping in providing conditions for the experiment. This work was supported by a scholarship from the Austria-Hungary Action Foundation to the first author. We thank Csaba Szabó, Ágnes Lukács, Dezső Németh, Karolina Janacsek, and István Winkler for valuable feedback and suggestions.

References

1. Juslin, P.N. and K.R. Scherer, *Vocal expression of affect*, in *The new handbook of methods in nonverbal behavior research*, J.A. Harrigan, R. Rosenthal, and K.R. Scherer, Editors. 2005, Oxford University Press: New York. p. 65-135.
2. Scherer, K.R., *Vocal communication of emotion: A review of research paradigms*. Speech Communication, 2003. **40**(1-2): p. 227-256.
3. Juslin, P.N. and P. Laukka, *Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion*. Emotion, 2001. **1**(4): p. 381-412.
4. Johnson, W.F., et al., *Recognition of Emotion from Vocal Cues*. Archives of General Psychiatry, 1986. **43**(3): p. 280-283.
5. Bänziger, T. and K.R. Scherer, *The role of intonation in emotional expressions*. Speech Communication, 2005. **46**(3-4): p. 252-267.
6. Banse, R. and K.R. Scherer, *Acoustic profiles in vocal emotion expression*. J Pers Soc Psychol, 1996. **70**(3): p. 614-36.
7. Schröder, M., *Experimental study of affect bursts*. Speech Communication, 2003. **40**(1-2): p. 99-116.
8. Ladd, D.R., et al., *Evidence for the independent function of intonation contour type, voice quality, and f0 range in signalling speaker affect*. Journal of the Acoustic Society of America, 1985. **78**(2): p. 435-444.

9. Bergmann, G., T. Goldbeck, and K.R. Scherer, *[Emotional impression of prosodic speech markers]*. Z Exp Angew Psychol, 1988. **35**(2): p. 167-200.
10. Breitenstein, C., D. Van Lancker, and I. Daum, *The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample*. Cognition & Emotion, 2001. **15**(1): p. 57-79.
11. Breitenstein, C., et al., *Erfassung der Emotionswahrnehmung bei zentralnervösen Läsionen und Erkrankungen: Psychometrische Gütekriterien der "Tübinger Affekt Batterie"*. Neurologie & Rehabilitation, 1996. **2**(93-101).
12. Goldman-Eisler, F., *Speech production and the predictability of words in context*. Quarterly Journal of Experimental Psychology, 1958. **10**: p. 96-106.
13. Goldman-Eisler, F., *Psycholinguistics: Experiments in spontaneous speech*. 1968, New York: Academic Press.
14. Rochester, S.R., *The significance of pauses in spontaneous speech*. Journal of Psycholinguistic Research, 1973. **2**(1): p. 51-81.
15. Eldred, S.H. and D.B. Price, *A linguistic evaluation of feeling states in psychotherapy*. Psychiatry, 1958. **21**(2): p. 115-21.
16. Kasl, S.V. and G.F. Mahl, *The Relationship of Disturbances and Hesitations in Spontaneous Speech to Anxiety*. J Pers Soc Psychol, 1965. **95**: p. 425-33.
17. Mahl, G.F., *Disturbances and silences in the patient's speech in psychotherapy*. J Abnorm Psychol, 1956. **53**(1): p. 1-15.
18. Pope, B., et al., *Anxiety and depression in speech*. J Consult Clin Psychol, 1970. **35**(1): p. 128-33.
19. Fairbanks, G. and L.W. Hoaglin, *An experimental study of the durational characteristics of the voice during the expression of emotion*. Speech Monographs, 1941. **8**(1): p. 85-90.
20. Jovicic, S.T., et al., *Serbian Emotional Speech Database: Design, Processing and Evaluation*, in *SPECOM-2004* 2004. p. 77-81.

21. Szabó, E., *A szomorú és a vidám érzelmi állapot megjelenése a beszédben*. Magyar Pszichológiai Szemle, 2008. **63**(4): p. 651–668.
22. Deppermann, A. and G. Lucius-Hoene, *Trauma erzählen – kommunikative, sprachliche und stimmliche Verfahren der Darstellung traumatischer Erlebnisse*. Psychotherapie und Spezialwissenschaft. Zeitschrift für Qualitative Forschung und klinische Praxis, 2005. **1**: p. 35–73.
23. Hofmann, S.G., et al., *Speech disturbances and gaze behavior during public speaking in subtypes of social phobia*. J Anxiety Disord, 1997. **11**(6): p. 573-85.
24. Laukka, P., et al., *In a Nervous Voice: Acoustic Analysis and Perception of Anxiety in Social Phobics' Speech*. Journal of Nonverbal Behavior, 2008. **32**(4): p. 195-214.
25. Ekman, P., E.R. Sorenson, and W.V. Friesen, *Pan-cultural elements in facial displays of emotion*. Science, 1969. **164**(3875): p. 86-8.
26. Izard, C.E., *The Face of Emotion*. 1971, New York: Appleton-Century-Crofts.
27. Izard, C.E., *Innate and universal facial expressions: evidence from developmental and cross-cultural research*. Psychol Bull, 1994. **115**(2): p. 288-99.
28. Ekman, A. and J. Strom, *[Habit forming drugs (7): somatic complications]*. Lakartidningen, 1969. **66**(48): p. 5021-4.
29. Ekman, P. and W.V. Friesen, *Constants across cultures in the face and emotion*. J Pers Soc Psychol, 1971. **17**(2): p. 124-9.
30. Biehl, M., et al., *Matsumoto and Ekman's Japanese and Caucasian Facial Expressions of Emotion (JACFEE): Reliability data and cross-national differences*. Journal of Nonverbal Behavior, 1997. **21**(1): p. 3-21.
31. Beier, E.G. and A.J. Zautra, *Identification of vocal communication of emotions across cultures*. J Consult Clin Psychol, 1972. **39**(1): p. 166.
32. Vanbezooijen, R., S.A. Otto, and T.A. Heenan, *Recognition of Vocal Expressions of Emotion - a 3-Nation Study to Identify Universal Characteristics*. Journal of Cross-Cultural Psychology, 1983. **14**(4): p. 387-406.

33. Mccluskey, K.W. and D.C. Albas, *Perception of the Emotional Content of Speech by Canadian and Mexican Children, Adolescents, and Adults*. International Journal of Psychology, 1981. **16**(2): p. 119-132.
34. Mccluskey, K.W., et al., *Cross-Cultural Differences in Perception of Emotional Content of Speech - Study of Development of Sensitivity in Canadian and Mexican Children*. Developmental Psychology, 1975. **11**(5): p. 551-555.
35. Thompson, W.F. and L.L. Balkwill, *Decoding speech prosody in five languages*. Semiotica, 2006. **158**(1-4): p. 407-424.
36. Scherer, K.R., R. Banse, and H.G. Wallbott, *Emotion inferences from vocal expression correlate across languages and cultures*. Journal of Cross-Cultural Psychology, 2001. **32**(1): p. 76-92.
37. Pell, M.D., et al., *Factors in the recognition of vocally expressed emotions: A comparison of four languages*. Journal of Phonetics, 2009. **37**(4): p. 417-435.
38. Albas, D.C., K.W. Mccluskey, and C.A. Albas, *Perception of Emotional Content of Speech - Comparison of 2 Canadian Groups*. Journal of Cross-Cultural Psychology, 1976. **7**(4): p. 481-490.
39. Pell, M.D. and V. Skorup, *Implicit processing of emotional prosody in a foreign versus native language*. Speech Communication, 2008. **50**(6): p. 519-530.
40. Fónagy, I. and K. Magdics, *Emotional patterns in intonation and music*. Zeitschrift fur Phonetik, 1963. **16**: p. 293-326.
41. Beaupre, M.G. and U. Hess, *Cross-cultural emotion recognition among Canadian ethnic groups*. Journal of Cross-Cultural Psychology, 2005. **36**(3): p. 355-370.
42. Ekman, P., et al., *Universals and cultural differences in the judgments of facial expressions of emotion*. J Pers Soc Psychol, 1987. **53**(4): p. 712-7.
43. Matsumoto, D., *Ethnic-Differences in Affect Intensity, Emotion Judgments, Display Rule Attitudes, and Self-Reported Emotional Expression in an American Sample*. Motivation and Emotion, 1993. **17**(2): p. 107-123.

44. Elfenbein, H.A., et al., *Toward a dialect theory: cultural differences in the expression and recognition of posed facial expressions*. *Emotion*, 2007. 7(1): p. 131-46.
45. Elfenbein, H.A. and N. Ambady, *Cultural similarity's consequences - A distance perspective on cross-cultural differences in emotion recognition*. *Journal of Cross-Cultural Psychology*, 2003. 34(1): p. 92-110.
46. Burkhardt, F., et al., *Emotional prosody - does culture make a difference?*, in *Proceedings of Speech Prosody2006*: Dresden, Germany.
47. Scherer, K.R., D.R. Ladd, and K.E.A. Silverman, *Vocal Cues to Speaker Affect - Testing 2 Models*. *Journal of the Acoustical Society of America*, 1984. 76(5): p. 1346-1356.
48. Cahn, J.E., *The Generation of Affect in Synthesized Speech*. *Journal of the American Voice I/O Society*, 1990. 8: p. 1-19.
49. Carlson, R., *Synthesis - Modeling Variability and Constraints*. *Speech Communication*, 1992. 11(2-3): p. 159-166.
50. Burkhardt, F. and W.F. Sendlmeier, *Verification of Acoustical Correlates of Emotional Speech using Formant-Synthesis*, in *ISCA Workshop (ITRW) on Speech and Emotion2000*: Belfast.
51. Gósy, M., *Magyar spontán beszéd adatbázis - BEA*. *Beszédkutatás 2008*, 2008: p. 116–128.
52. Gósy, M., *A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései*. *Magyar Nyelvőr*, 2003. 127: p. 257–277.
53. Gósy, M. and M. Kovács, *Virtual sentences of spontaneous speech*, in *Human Factors and Voice Interactive Systems*. 2008, Springer: New York ; London. p. 193-207.
54. Henderson, A.I., F. Goldman-Eisler, and A. Skarbek, *Sequential temporal patterns in spontaneous speech*. *Language and Speech*, 1966. 9: p. 207-216.
55. Schönplflug, U., *Pauses in elementary school children's verbatim and gist free recall of a story*. *Cognitive Development*, 2008. 23 p. 385-394.

56. Williams, E.J., *Experimental designs balanced for the estimation of residual effects of treatments*. Australian Journal of Scientific Research, 1949. **Ser. A 2**: p. 149-168.
57. Fontaine, J.R., et al., *The world of emotions is not two-dimensional*. Psychological Science, 2007. **18**(12): p. 1050-7.

Tanulmány [3]

Tisljár-Szabó, E., Rossu, R., Varga, V., Pléh, Cs. (2014) The effect of alcohol on speech production. *Journal of Psycholinguistic Research*. (Online first, DOI 10.1007/s10936-013-9278-y) (IF: 0,642)

(a megjelent változattal tartalmilag megegyező kézirat változat)

Abstract

The present study investigates the effect of acute alcohol consumption on speech in Hungarian subjects. The measures used to reveal these effects were tongue-twisters, which were grouped according to their linguistic features. The number and type of speech errors while uttering the tongue-twisters were compared between intoxicated and sober conditions. The results showed that subjects made more speech errors in alcohol influenced than in sober states in all types of the tongue-twisters except for those using foreign words. Changes in the articulation rate, number of pauses and fundamental frequency were investigated as well. In the intoxicated state, no changes were observed in fundamental frequency and articulation rate, while the number of pauses increased.

Keywords: alcohol, speech error, pitch, tempo, tongue-twisters.

INTRODUCTION

It is a commonplace observation that certain cognitive and motor processes change in states under the influence of alcohol. People are unable to walk straight, reaction time is slower, the number of speech errors increase.

Several biological studies investigated the molecular action of alcohol in the central nervous system (Givens & Breese, 1990; Hunt, 1985; Suzdak & Paul, 1987). The main findings showed that alcohol has a general depressant action on the central nervous system through the impairment of neuronal transmission. Alcohol is considered to be unequally distributed, because it has an effect only on certain functions of the nervous system. The effects show up

first at the cerebral level and then descend to cerebellar and spinal cord levels (Chin & Pisoni, 1997).

As a corollary to the cortical effects, alcohol has an effect on perceptual and motor processes (Chin & Pisoni, 1997; Starmer, 1989). In alcohol influenced states dynamic visual acuity, depth perception, and discrimination between auditory stimuli are impaired, and glare resistance is reduced. The effect of alcohol on reaction time depends on the alcohol dose, the alcohol concentration attained, and the complexity of the task. Several studies investigated the effect of alcohol on cognitive functions (see Starmer, 1989; Wallgren & Barry, 1970) and found that alcohol has the most robust effect on selective attention, divided attention, and short-term memory.

Santamaria (1989, see Chin & Pisoni, 1997) summarized the main behavioral effects of alcohol. At a low level of alcohol, judgment and visual acuity are impaired. On the second level, between 0.05 and 0.08% blood-alcohol concentration (BAC), muscle control is affected, which leads to increases in uncoordinated movements and in reaction time. At around 0.10% BAC nystagmus develops and movements become increasingly uncoordinated. As performance in fine movements and skilled operations are affected, articulation is challenged, and ataxia emerges. At an even higher concentration, people become sleepy and dazed. At about 0.30% coma, and above 0.40% death, may occur.

Studies on the effect of alcohol on speech

Alcohol has an effect on cognitive functions and motor coordination as well (Hollien, DeJong, Martin, Schwartz, & Liljegren, 2001). As the ability to speak needs both of these functions we can assume that alcohol has an effect on speech production. There are two main types of studies investigating the effect of alcohol on speech. Decoding studies examine to what extent listeners are able to determine whether speakers are in an alcohol influenced state. According to previous studies (Hollien et al., 2001; Hollien, Harnsberger, Martin, Hill, & Alderman, 2009) listeners are able to judge talkers as being intoxicated from 62% to 74% of the time. While decoding studies investigate speech perception, encoding studies analyze words, sentences or paragraphs in speech samples which were recorded while talkers were intoxicated. One type of the studies focuses on speech errors (Gyarmathy, 2008; Lester & Skousen, 1974; Sobell & Sobell, 1972; Trojan & Kryspin-Exner, 1968), while others analyze

supra-segmental changes in speech under the influence of alcohol (see, for example Behne et al., 1991; Cooney, McGuigan, Murphy, & Conroy, 1998; Hollien, Liljegren, Martin, & DeJong, 1999; Hollien & Martin, 1996; Klingholz, Penning, & Liebhadt, 1988; Künzel, Braun, & Eysholdt, 1992; Pisoni & Martin, 1989; Sobell, Sobell, & Coleman, 1982; Watanabe et al., 1994).

The studies also differ regarding the speech material they analyze (Chin & Pisoni, 1997). Some studies use spontaneous or semi-spontaneous speech, in which subjects tell stories they have experienced (Trojan & Kryspin-Exner, 1968), have a conversation with each other (Gyarmathy, 2008; Smith, Parker, & Noble, 1975) or describe pictures (Collins, 1980) in an alcohol influenced state. Other studies use the so called naming task, where subjects have to name objects or colors they see in pictures, or give antonyms of adjectives. (e.g. (Hollingworth, 1923; Moskowitz & Roth, 1971).

A further method used to obtain speech samples in an alcohol influenced state is the shadowing task, where “stimulus materials are presented auditorily to subjects, who then repeat what they have heard as quickly as possible” (Chin and Pisoni, 1997, 82.). A final type of task is when subjects are asked to read a list of isolated words, phrases, sentences, or passages in an alcohol influenced state (D. M. Behne & Rivera, 1990; Dawn M. Behne, Rivera, & Pisoni, 1991; Johnson et al., 1993; Pisoni, Hathaway, & Yuchtman, 1985; Sobell & Sobell, 1972; Sobell, Sobell, & Coleman, 1982).

In the international literature since the 1990s researchers have been using only experimental designs and standardized texts which gives the researchers the opportunity to perform detailed cross-subject and cross-condition analyses (Chin & Pisoni, 1997).

Behne et al. (1991) asked nine male university students to shadow words and sentences, as well as to read aloud spondees, sentences and passages. (The same speech material was also used by D. M. Behne and Rivera (1990); Pisoni et al. (1985); Pisoni and Martin (1989) as part of a research program conducted under a contract between General Motors Research Laboratories and Indiana University). Results showed that duration of sentences and monosyllabic words was longer in an intoxicated state, but there was no difference in speech rate in the case of the spondees (words of two long syllables).

In a study by Künzel, Braun and Eysholdt (1992), 33 male police officers read a passage aloud in sober and intoxicated states. Analysis of speech errors such as repetitions, insertions, omissions, and substitutions showed that under alcoholic intoxication the number of errors

increased and speech rate decreased, while the mean length of pauses increased.

The first Hungarian study in this field was made by Gyarmathy (2008), who used spontaneous conversations and a test sentence to analyze the effect of alcohol on speech disfluencies. She used Gósy's (2003) classification, and found more speech disfluencies in participants' speech in an alcohol influenced than in a sober state. The rate of 'errors' increased compared to 'uncertainties'. Gyarmathy did not use any statistical analysis, but based on her raw data, in an alcohol influenced state the rate of expletives, repetitions, restarts, and slips of the tongue seemed to increase, while the number of pauses decreased.

Independently of the type of the speech sample used, the most frequently analyzed parameters are the number of speech errors, the fundamental frequency, and the speech rate, as well as the number and length of silent pauses.

Based on results from several previous studies, one can summarize that in most of the cases in an intoxicated state the speech rate decreases (Dawn M. Behne et al., 1991; Hollien et al., 2001; Künzel et al., 1992; Pisoni et al., 1985; Pisoni & Martin, 1989; Sobell et al., 1982), more speech errors occur (Hollien et al., 2001; Künzel et al., 1992; Trojan & Kryspin-Exner, 1968; Zaimov, 1969), and the length of pauses increases (Künzel et al., 1992). Regarding fundamental frequency, results from previous studies are inconsistent. Many of the studies (D. M. Behne & Rivera, 1990; Hollien et al., 2001; Klingholz, Penning, & Liebhardt, 1988; Trojan & Kryspin-Exner, 1968) found increased pitch in an intoxicated state. However, according to Sobell et al. (1982), Johnson et al. (1993) or Gyarmathy (2008), alcohol has no effect or has an ambiguous effect on fundamental frequency: in most of the cases fundamental frequency increases, but in some cases it decreases or does not change. Schiel and Heinrich (2009) found that fundamental frequency typically increased for women and decreased for men. Watanabe (1994) found decreased pitch even for women.

In this study we investigated the effect of alcohol on repeating Hungarian tongue-twisters. We analyzed speech errors, tempo, pauses and fundamental frequency, and compared these in sober and in alcohol influenced states. We assumed that in an alcohol influenced state, subjects commit more speech errors, speech tempo decreases, and the number of pauses increases, compared to the sober state. We assumed an increase in fundamental frequency in an alcohol influenced state, although because of the previous conflicting results this hypothesis was considered open.

METHOD

Subjects

15 undergraduate students, 7 female and 8 male, participated in the experiment, with a mean age of 20.73 (SD=1.79). They were all native speakers of Hungarian. Participation was voluntary; subjects did not receive any compensation for their participation.

Material

We used 3 questions from the Alcohol Use Disorders Identification Test (Babor, Biddle-Higgins, Saunders, & Monteiro, 2001) to chart the drinking habits of the subjects. We also collected some personal data.

The experimental material consisted of a list of tongue-twisters and Jacob's digit span test (methods and results related to short memory performance are presented in another paper; see Szabó et al. , manuscript).

Tongue-twisters were collected and grouped into 5 categories.

The 5 phonologic tongue-twisters were Hungarian minimal pairs, where one syllable words differed only in one consonant (for example *két kék* [two blue]).

In the morphologic tongue-twisters (4 expressions) subjects had to conjugate a verb with its adjunct (for example *lassan látni* [seeing slowly]) in the imperative, from the first person singular to the third person plural. In Hungarian, this is quite difficult because of the agglutinative nature of the language, and because of the fusional nature of certain allomorphs (for example the first person singular imperative form of *lassan látni* is *lassan lássak*.)

The 10 phonetic and the 6 syntactic tongue-twisters were conventional Hungarian tongue-twisters. In the first case difficulties in utterance are caused by the phonemes themselves and in the second case they are influenced by the syntactical structure of the sentence.

We also used 10 foreign words difficult to pronounce for Hungarians (like *vucsicsevázok*) as tongue-twisters. (For the list of tongue-twisters see the Appendix). We eliminated tongue-twisters which were very popular.

This grouping was ad hoc, made by the authors. In selecting the speech material we preferred to use expressions from real life. To do so, we did not create any artificial linguistic material

which would have matched any speech production model, but collected already existing, difficult to utter Hungarian expressions. Thus, the names of the categories do not exclusively refer to the features of the members. For example in the phonetic tongue-twisters group, expressions with “phonologic” difficulties can occur.

Stimulus materials were recorded by a female university student (age: 21), with Hungarian as her mother tongue and with no speech, hearing, or voice disorders. Recordings were made by a dynamic microphone (Pioneer DM-DV5) using the computer program PRAAT (Boersma & Weenink, 2009), Subjects’ answers were recorded by the same microphone and by the same computer program.

Procedure

Two pretests were done before the experiment where we tested the effect of alcohol on independent subjects and decided how much time should elapse between alcohol consumption and the repetition of the test.

Subjects were requested not to eat for 4 hours before the experiment. In the experiment 3 times 50 ml of a 40% alcohol concentration shot drink (47.25 g ethanol, see AUDIT, Babor et al., 2001) was consumed within a maximum of 5 minutes. After an hour, subjects were asked to repeat the tongue-twisters. Stimuli were presented auditorily through head-phones in a random order, and subjects were instructed to listen carefully to each stimulus sentence or expression and then to repeat it aloud or, in the case of the morphologic tongue-twisters, to make the conjugation. Syntactic tongue-twisters were presented twice. Subjects also rated their intoxication on a 1-to-7 point scale.

Two weeks later the same subjects participated in a sober condition, and did the same task without alcohol consumption (sober/control condition).

For the statistical analysis of data we used the computer program SPSS 15.0. We used a paired sample t-test to investigate whether there are any differences in changes regarding the measured parameters in an alcohol influenced state compared to a sober state.

RESULTS

Alcohol consumption data

Subjects ate 5.67 hours (SD=1.68) before the experiment. 67% of the subjects usually consume alcohol two or three times a week, and drink 3-4 units⁹ of it each time. 80% of the subjects answered that once in a month they drink more than 5 units of alcohol.

Subjects rated themselves as 4.19 (SD=1.38) intoxicated on the 1-to-7 scale. The average weight of the subjects was 68.51 kg and their blood alcohol level (consumed poor alcohol (ethanol)/body weight) was 0.69 g/kg (SD=0.1).

Speech errors

Two types of errors were differentiated. Errors which were related to memory dysfunction belonged to the first group, and errors where subjects could not end the sentence, or eliminated or modified a word belonged to this category. Speech errors such as insertions, repetitions, fillers, and slips of the tongue belonged to the second group. The present study only analyzes speech errors; alcohol's effect on memory dysfunction is investigated in another paper (see Szabó et al., manuscript).

We used a paired sample t-test to test whether subjects make more errors in an alcohol influenced than in a sober state. Results showed that there was a significant difference between the intoxicated (Mean=20.53, SD=3.58) and the sober (Mean=13.07, SD=4.93) condition. All in all, in an alcoholic intoxication state subjects made more errors ($t=8.456$, $df=14$, $p<0.05$). In detail, except for one task (foreign words), in every kind of task subjects made more errors in an alcohol influenced than in a sober state (Table I. shows the significant results).

⁹ One unit is the equivalent of half a bottle (2.5 dl) of beer, 100 ml wine or a 40 ml shot of spirit.

Fundamental frequency

The paired sample t-test showed no difference between the alcohol influenced and the sober states ($t=1.267$, $df=14$, $p=0.226$) regarding the fundamental frequency.

We also investigated possible gender differences. We calculated the degree of the pitch change between the alcohol influenced and the sober state in the case of men and women. In an alcohol influenced state, in the case of men, fundamental frequency increased on average by 3.49 Hz, while in the case of the women it increased by 5.32 Hz, compared to the sober state. The result of the two-sample t-test showed that there is no difference between the two groups regarding the changes ($t=-3.12$, $df=13$, $p=0.76$). Figure 1 shows the value of fundamental frequency by individual speakers in sober and in alcohol influenced states, while Figure 2 shows the change in fundamental frequency in the alcohol condition compared to the sober condition in percentages. In the case of three women pitch decreased (Mean: 7.2 Hz), in the case of four women pitch increased (Mean: 14.7 Hz). In four of the men pitch decreased (Mean: 1.6 Hz) while in another four pitch increased (Mean: 8.6 Hz).

Articulation rate

Articulation rate was defined as the number of production units (here phonemes) per unit time (here seconds) after subtracting pauses (Crystal & House, 1990; Gósy, 2004).

We conducted paired sample t-tests, and found that there was no significant difference between the alcohol influenced and the sober states ($t=-1.165$, $df=14$, $p=0.264$). Results were significant only in one case: in alcoholic intoxication the articulation rate decreased during the syntactic tongue-twisters ($t=-2.264$, $df=14$, $p<0.05$).

Number of pauses

We measured the number of pauses irrespective of whether the subject uttered the tongue-twister correctly or not. All in all, the paired-sample t-test showed a significant difference between the alcohol influenced and the control state. In the alcohol influenced state the number of pauses increased ($t=-3.081$, $df=14$, $p<0.05$).

In detail there was a significant difference in the phonetic ($t=2.717$, $df=14$, $p<0.05$), in the morphologic ($t=3.123$, $df=14$, $p<0.05$), and the syntactic ($t=2.172$, $df=14$, $p<0.05$) tongue-twisters (see Figure 3.). In all of the cases, the number of pauses increased in the alcohol

influenced state. However, there was no difference in the case of the phonologic tongue-twisters and the foreign words.

Relationship between the rate of feeling intoxicated and speech

We used the Pearson correlation to test whether there is any relationship between ratings of self perceived intoxication and speech errors, pauses or speech tempo. The results showed that rate of self perceived intoxication did not correlate with the number of speech errors ($r=-0.023$, $p=0.935$), with the articulation rate ($r=-0.025$, $p=0.93$), or with the number of pauses ($r=0.414$, $p=0.125$).

DISCUSSION

In this study we investigated the effect of alcohol on speech. We used tongue-twisters, and grouped them into five categories according to linguistic features, resulting in the following categories: phonetic, phonologic, morphologic and syntactic tongue-twisters, and tongue-twister-like foreign words. The subjects' task was to repeat the words and the sentences which were presented through head-phones, as well as to conjugate some verbs in alcohol influenced and sober conditions.

We analyzed speech errors, pauses, articulation rate and pitch in Hungarian speech. Our results showed that in an alcohol influenced state, the number of these speech errors increased. This happened in all types of tongue-twisters except those with foreign words. One part of the models of speech production (Garrett, 1988; Levelt, 1989) is based on the analyses of errors occurring in speech. According to these models, errors are related to actual functional disorders of certain levels of speech production (like syntactic planning or lexical retrieval). There are some research studies which investigated alcohol's effect on speech disfluencies, however there are less interpretations in these. Christenfeld and Creager (1996) found that number of one certain type of speech disfluency, of the so called *ums*, decreased when subjects were intoxicated. According to their interpretation, under alcoholic intoxication rate of self-consciousness and self-monitoring decreases which leads to a more automatic speech and furthermore to less *ums*. "This approach also suggests that people who make an

effort to speak well may use *ums* more often than people who do not care how well they are talking” (Christenfeld and Creager, 1996, p. 459.).

At the same time, other studies (Forney & Hughes, 1961; Gyarmathy, 2008; Hollien et al., 2001; Künzel et al., 1992; Zaimov, 1969) found that in an alcohol influenced state, the number of speech disfluencies increases. Our results correspond to these studies. Looking at Levelt’s model (1989) we would say that in our study alcohol firstly had an effect on the motor-articulation level. As in this experiment the primary task was to repeat sentences, we can not talk about conceptualization or formulation. Errors were related to the executive level. Santamaria (1989, see Chin & Pisoni, 1997) reports that even a BAC level of 0.05-0.08% has an effect on muscle control. As speech production needs precise muscle control, where this is lacking the number of speech errors increases. So, we would state that although alcohol might have a decreasing effect on self-monitoring, if the task is difficult – which is the case when repeating tongue-twisters – the extra attention and the effect of alcohol on motor processes together lead to an increased number of speech errors.

Previous studies produced different results regarding the pitch change in an alcohol influenced state. Some studies found increases in pitch (D. M. Behne & Rivera, 1990; Hollien et al., 2001; Künzel et al., 1992; Trojan & Kryspin-Exner, 1968), some measured no or ambiguous changes (Gyarmathy, 2008; Johnson et al., 1993; Sobell et al., 1982) or even decreases in pitch (Schiel & Heinrich, 2009; Watanabe et al., 1994). In our study we measured the fundamental frequency of all words and sentences of the sober and the intoxicated conditions. Our results showed that there is no significant difference between the two states.

It is quite difficult to compare the amount of alcohol used in the different studies, because some of the studies measured BAC (Blood Alcohol Concentration), some measured BrAC (Breath Alcohol Concentration), and some gave the amount of the consumed alcohol (ethanol) in grams or in g/kg body weight. When trying to use the work of Chin and Pisoni (1997) to convert the different measures¹⁰ we find that studies which found an effect of alcohol on fundamental frequency had a maximum BAC from approximately 0.1 to 0.2%, while studies which did not report such differences had 0.079 to 0.13% BAC. In our study BAC was

¹⁰ We only have taken into account studies which measured BAC or the amount of alcohol because there is no agreement in the case of the BrAC (Chin & Pisoni, 1997; Haffner, Graw, Dettling, Schmitt, & Schuff, 2003; Jones, 2010; Widmark, 1981).

approximately 0.08%, which does not rate as a high value. Thus one could argue that with more alcohol we also would have the increased pitch effect.

The study of Watanabe et al. (1994) is the only one that found decreases in pitch for women as well as for men. This is also the only study that was not conducted in Europe but in Japan. It might therefore be possible that other language or culture specific effects are stronger or co-effect with the physiological effect of the alcohol. We also investigated whether there are gender differences. Regarding the pitch change, we didn't find any difference between men and women. The increase in the pitch in the alcohol influenced state was approximately the same in the case of the men and the women. These results contradict Schiel and Heinrich (2009) who found that fundamental frequency typically increases for women and decreases for men.

In connection with the fundamental frequency we can summarize that in our study there was no difference between the alcohol and the control condition. We neither found any gender differences, nor any tendency to a change in fundamental frequency in one direction in the case of men or women. Reviewing previous studies we can suppose that there is a threshold value in blood-alcohol level and fundamental frequency below this value does not change. In our study, BAC level might not have exceeded this threshold.

The data are relevant regarding temporal changes in the alcohol influenced state. Many studies stated that speech rate decreases in the intoxicated state (Dawn M. Behne et al., 1991; Hollien et al., 2001; Künzel et al., 1992; Pisoni et al., 1985; Pisoni & Martin, 1989; Sobell et al., 1982). At the same time, apart from one study (Künzel et al., 1992), no one analyzed in detail whether this happens because of the decrease in the articulation rate or because of the increased number of the pauses. Künzel et al. (1992) investigated speech rate, articulation rate and the number and length of pauses in the alcohol influenced state. Their results showed that in intoxicated states, speech rate and articulation rate decreased while the mean length of the pauses increased. The increase in the number of the pauses was not significant.

Contradicting these results, we found that there was no significant difference regarding the articulation rate. Subjects did not utter the words and sentences more slowly in the alcohol influenced than in the sober state, with the case of the syntactic tongue-twisters being the only exception. In our study, the decrease in speech rate was due to the increase in the number of pauses. This might be a task-related effect. While Künzel's subjects had to read a passage, in our research subjects had to repeat tongue-twisters, or make conjugations. Although alcohol has an effect in both cases, and forces speakers to slow down the tempo, in the case of reading

it is due to the decrease in the articulation rate, while in the case of repeating it is due to the increase in the number of pauses. Slowing down speech tempo in the intoxicated state might be a form of compensation, whereby subjects try to counterweight the negative effects of alcohol on speech production. The largest compensation was observed in the case of the syntactic tongue-twisters, where not only the number of pauses increased but the articulation rate decreased as well. These tongue-twisters were the longest and probably the most difficult, which necessitated this extra compensation.

The question arises of how the expectation effect can influence the error rate and whether people who feel more intoxicated have the expectation that they will make more speech errors and thus perform accordingly, or conversely, take more care and make fewer errors. Subjects rated themselves at a 4.19 (SD=1.38) level of intoxication on the 1-to-7 scale. When we correlated the subjective perceived rate of intoxication and the number of the errors, we found that there was no relation between these two variables. Subjects who felt more intoxicated did not make more or fewer errors than subjects who rated themselves more sober. Self perceived intoxication did not correlate with the articulation rate or with the number of speech pauses either. So, subjects' expectations might not affect their performance. This has practical importance regarding the behavioral content of self intoxication reports e.g. in criminal cases.

CONCLUSION

In this paper we presented the results of a Hungarian research experiment, in which the results partly correspond with, partly complement, and even partly contradict findings from earlier (primarily German and British) studies. We can state that changes in speech were observed in alcohol influenced states compared to the sober state. At the same time, there are many limits to investigating the effect of alcohol on speech. Depending on the amount of alcohol, weight and physique of the person, and previously consumed food, the effects of alcohol can vary. Communication skills are also variable. Although tongue-twisters are a good instrument for investigating the effect of alcohol on speech, their practical application is not recommended, as people make speech errors in them even in a sober state, and would therefore be more appropriate for measuring within-subject changes. However, the need for more research investigating the effect of alcohol on speech and voice cannot be questioned. One of the most

important issues would be applications on telephone based tests, where for example it would be possible to test the trustworthiness of people on duty.

Acknowledgement

We thank Dora Somogyi for the language correction and Katinka Imréczi for her assistance.

REFERENCES

- Babor, T. F., Biddle-Higgins, J. C., Saunders, J. B., & Monteiro, M. G. (2001). AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care
- Behne, D. M., & Rivera, S. M. (1990). Effects of alcohol on speech: Acoustic analysis of spondees. *Research on Speech Perception* 16, 263–291.
- Behne, D. M., Rivera, S. M., & Pisoni, D. B. (1991). Effects of alcohol on speech: Durations of isolated words, sentences, and passages in fluent speech. *Journal of the Acoustical Society of America*, 90, 2311.
- Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1.05.). Computer program
- Chin, S. B., & Pisoni, D. B. (1997). *Alcohol and Speech*. San Diego: Academic Press.
- Christenfeld, N., & Creager, B. (1996). Anxiety, alcohol, aphasia, and ums. *Journal of Personality and Social Psychology*, 70(3), 451-460.
- Collins, P. J. (1980). A Comparison of the Oral Syntactic Performance of Alcoholic and Non-Alcoholic Adults. *Language and Speech*, 23, 281-288.
- Crystal, T. H., & House, A. S. (1990). Articulation Rate and the Duration of Syllables and Stress Groups in Connected Speech. *Journal of the Acoustical Society of America*, 88(1), 101-112.
- Forney, R. B., & Hughes, F. W. (1961). Delayed auditory feedback and ethanol: Effect on verbal and arithmetical performance. *Journal of Psychology*, 52, 185-192.

- Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey* (Vol. 3, pp. 69–96). Cambridge: Cambridge University Press.
- Givens, B. S., & Breese, G. R. (1990). Electrophysiological Evidence That Ethanol Alters Function of Medial Septal Area without Affecting Lateral Septal Function. *Journal of Pharmacology and Experimental Therapeutics*, 253(1), 95-103.
- Gósy, M. (2003). A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. [Co-occurrence and frequency of disfluencies in Hungarian spontaneous speech]. *Magyar Nyelvőr*, 127, 257–277.
- Gósy, M. (2004). *Fonetika, a beszéd tudománya*. Budapest: Osiris Kiadó.
- Gyarmathy, D. (2008). Hogyan hat az alkohol a spontán beszédre? . In Á. Kuna & Á. Veszelszki (Eds.), *Félúton 3 Konferenciakötet* (pp. 81-99).
- Haffner, H. T., Graw, M., Dettling, A., Schmitt, G., & Schuff, A. (2003). Concentration dependency of the BAC/BrAC (blood alcohol concentration/breath alcohol concentration) conversion factor during the linear elimination phase. *International Journal of Legal Medicine*, 117(5), 276-281. doi: DOI 10.1007/s00414-003-0384-5
- Hollien, H., DeJong, G., Martin, C. A., Schwartz, R., & Liljegren, K. (2001). Effects of ethanol intoxication on speech suprasegmentals. *Journal of the Acoustical Society of America*, 110(6), 3198-3206.
- Hollien, H., Harnsberger, J. D., Martin, C. A., Hill, R., & Alderman, G. A. (2009). Perceiving the Effects of Ethanol Intoxication on Voice. *Journal of Voice*, 23(5), 552-559. doi: DOI 10.1016/j.jvoice.2007.11.005
- Hollingworth, H. L. (1923). The influence of alcohol. *Journal of Abnormal Psychology and Social Psychology*, 18, 204-237.
- Hunt, W. A. (1985). *Alcohol and biological membranes*. New York: Guilford Press.
- Johnson, K., Southwood, M. H., Schmidt, A. M., Mouli, C. M., Holmes, A. T., Armstrong, A. A., . . . Wilson, A. S. (1993). *A physiological study of the effects of alcohol on speech and voice*. Paper presented at the 22nd annual Symposium on the Care of the Professional Voice at the Voice Foundation.

- Jones, A. W. (2010). The Relationship between Blood Alcohol Concentration (BAC) and Breath Alcohol Concentration (BrAC): A Review of the Evidence. *Road Safety Web Publication*, 15. Retrieved from <http://www.dft.gov.uk/pgr/roadsafety/research/rsrr/theme3/report15.pdf> website:
- Klingholz, F., Penning, R., & Liebhardt, E. (1988). Recognition of Low-Level Alcohol-Intoxication from Speech Signal. *Journal of the Acoustical Society of America*, 84(3), 929-935.
- Künzel, H. J., Braun, A., & Eysholdt, U. (1992). *Einfluss von Alkohol auf Sprache und Stimme*. Heidelberg: Kriminalistik Verlag.
- Lester, L., & Skousen, R. (1974). The phonology of drunkenness. In A. Bruck, R. Fox & M. W. LaGaly (Eds.), *Papers from the parasession on natural phonology* (pp. 233-239). Chicago: Linguistic Society.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Moskowitz, H., & Roth, S. (1971). Effect of alcohol on response latency in object naming. *Quarterly journal of studies on alcohol*, 32(4), 969-975.
- Pisoni, D. B., Hathaway, S. N., & Yuchtman, M. (1985). *Effects of alcohol on the acoustic-phonetic properties of speech: Final report to GM Research Laboratories. (SRL Technical Note No. 85-03)*. Bloomington: Indiana University.
- Pisoni, D. B., & Martin, C. S. (1989). Effects of Alcohol on the Acoustic-Phonetic Properties of Speech - Perceptual and Acoustic Analyses. *Alcoholism-Clinical and Experimental Research*, 13(4), 577-587.
- Schiel, F., & Heinrich, C. (2009). *Laying the Foundation for In-car Alcohol Detection by Speech*. Paper presented at the INTERSPEECH, Brighton, UK.
- Smith, R. C., Parker, E. S., & Noble, E. P. (1975). Alcohol's effect on some formal aspects of verbal social communication. *Archives of General Psychiatry*, 32, 1394-1398.
- Sobell, L. C., & Sobell, M. B. (1972). Effects of alcohol on the speech of alcoholics. *Journal of speech and hearing research*, 15(4), 861-868.

- Sobell, L. C., Sobell, M. B., & Coleman, R. F. (1982). Alcohol-induced disfluency in nonalcoholics. *Folia Phoniatrica*, 34, 316–323.
- Starmer, G. A. (1989). Effects of low to moderate doses of ethanol on human driving-related performance. In K. C. Crow & R. D. Batt (Eds.), *Human metabolism of alcohol* (Vol. 1, pp. 101-130). Boca Raton, FL: CRC Press.
- Suzdak, P. D., & Paul, S. M. (1987). Ethanol stimulates GABA receptor-mediated Cl⁻ ion flux in vitro: possible relationship to the anxiolytic and intoxicating actions of alcohol. *Psychopharmacology Bulletin*, 23(3), 445-451.
- Szabó, E., Rossu, R., Varga, V., & Imréczi, K. *Az alkohol hatása a rövid távú emlékezetre és a beszédprodukciós folyamatokra.*
- Trojan, F., & Kryspin-Exner, K. (1968). The decay of articulation under the influence of alcohol and paraldehyde. *Folia Phoniatrica (Basel)*, 20(4), 217-238.
- Wallgren, H., & Barry, H. (1970). *Actions of alcohol*. Amsterdam: Elsevier.
- Watanabe, H., Shin, T., Matsuo, H., Okuno, F., Tsuji, T., Matsuoka, M., . . . Matsunaga, H. (1994). Studies on vocal fold injection and changes in pitch associated with alcohol intake. *Journal of Voice*, 8(4), 340-346.
- Widmark, E. M. P. (1981). *Principles and Applications of Medicolegal Alcohol Determination*. Davis, CA: Biomedical Publications.
- Zaimov, K. (1969). Die Sprachstörungen als Kriterium der Bewusstseinsstörungen. *Psychiatrie, Neurologie, und Medizinische Psychologie*, 21, 218–225.

APPENDIX

List of the tongue-twisters

A. Phonetic tongue-twisters

- a. ha a hörcsög szörpöt szörcsög, rátörnek a hörcsög görcsök
- b. pepita bóbitás babapókok pattognak a bobbályán
- c. Tóth Tódor doktora Tóth Dorka doktora is
- d. rálel-e e remek lelet lereagálására
- e. szétzúzott cicaszántás
- f. sóska zsúri csacska zsúri
- g. pocakon pöckölt icike-picike pocok
- h. Lenin mauzóleumának lelinóleumozása
- i. zabra zebra zsebre zabra
- j. fekete bikapata kopog a patika pepita kövén

B. Phonologic tongue-twisters (minimal pairs)

- a. már mér kér két lét
- b. hány lány lát fát
- c. pót sót szó Óz, ez íz így egy kegy
- d. két kék fék fér
- e. száz szász szesz lesz

C. Morphologic tongue-twisters (conjugation)

- a. lassan látni
- b. két kérdést kérdezni
- c. felfogni, hogy felfogtátok
- d. várni a váram

D. Syntactic tongue-twisters

- a. Nem lehet a Márta másé, mert a Márta már Tamásé.
- b. Szemben a nagy teátrummal, iszik egy úr teát rummal, de mire a tea kész lett eltörött a teás készlet.
- c. Nem minden fajta szarka farka tarka-barka, csak a tarka fajta szarka farka tarka-barka.
- d. Az igazság a gazság igazsága?
- e. “Aj, e nő-kebelű Lidi óta, a Jenőke belül idióta.”
- f. Már itt volt ősz szinte, mikor egy őszinte ősz inte, hogy legyek őszinte, mert ő szinte őszinte.

E. Foreign words

- a. kalazantinum
- b. mezateriitisz
- c. patriszticizmus
- d. emplaiztrum
- e. kromatipográfia
- f. szekvesztráció
- g. nomonogáfia
- h. nummulitok
- i. xeroftalmia
- j. vucsicsevázok

Tanulmány [4]

Szabó, E., Rossu, R., Varga, V., Imrécz, K. (2012) Az alkohol hatása a rövid távú emlékezetre és a beszédprodukciós folyamatokra. *Magyar Pszichológiai Szemle*, 67, 3, 409-430. DOI: 10.1556/MPSzle.67.2012.3.1

(a megjelent változattal tartalmilag megegyező kézirat változat)

Absztrakt

Kutatásunkban az alkohol beszédre és rövid távú emlékezetre gyakorolt hatását vizsgáltuk jól kontrollált feltételek mellett. Ennek vizsgálatára a Jacobs-féle számterjedelemtesztet és nyelvi szintek szerint csoportosított nyelvtörőket használtunk, és összehasonlítottuk alkoholos és kontroll feltételben a számterjedelemteszten nyújtott teljesítményt, illetve a nyelvtörők kimondása során elkövetett hibák számát és típusát. Különválasztottuk és elemeztük az emlékezeti teljesítmény romlásából fakadó és a motoros-artikulációs folyamatokból eredő hibákat. Az eredmények azt mutatják, hogy az alkohol nincs hatással a számterjedelemre, azonban hatással van a nyelvtörők közvetlen felidezésére. Ugyanakkor a nyelvtörők visszamondásában elkövetett hibák 58-61 %-a a pontatlan artikulációból fakad. Elemeztük az artikulációs tempó, a hangmagasság és a szünetek számának változását is. Az eredmények nagyrészt egybevágóak korábbi angol és német kutatási eredményekkel, de részben ki is egészítik azokat.

Kulcsszavak: alkohol, beszéd, megakadások, munkamemória, hangmagasság, tempó

Bevezető

Bár számos módszer létezik arra, hogy az alkoholt kimutassák szervezetünkben (vérvétel, kilélegzett levegő alkoholtartalmának mérése), mégis érdekes kérdés, hogy a beszédünkben hogyan mutatkoznak meg az alkoholos befolyásoltság jegyei. Pszichológiai kísérletek

segítségével egyrészt jobban megismerhetjük az alkohol pszichés folyamatokra gyakorolt hatását, másrészt olyan módszereket fejleszthetünk ki, amelyek segítségével egy korábban felvett hanganyag alapján megállapítható a beszélő józansága illetve ittassága. (Erre tettek például kísérletet Johnson, Pisoni és Bernacki (1989, 1990) valamint Pisoni, Johnson és Bernacki (1991), amikor az Exxon Valdez tankhajó balesetének körülményeinek vizsgálatakor a kapitánytól származó hanganyagot elemezték.)

Felmerül tehát kérdésként, hogy a beszédbeli jegyeket felhasználhatjuk-e az ittasság megállapításához, és ha igen, melyek azok a jegyek, amelyek beszélőtől függően, és melyek azok, amelyek beszélőfüggetlenül felhasználhatók. Manapság Amerikában például egy jól ismert és a rendőrök által használt, bár nem konkrét kutatáson alapuló józansági teszt az ábécé nagy sebességű elmondatása az alanyokkal (ld. Karch, 2008).

Az alkoholnak az emlékezetre gyakorolt hatásával kapcsolatban a legismertebb jelenség, amikor egy átmulatott éjszaka után másnap nem emlékszünk, hogy kivel és miről is beszélgettünk előző nap. Az alkohol és a rövid távú emlékezet kapcsolata már nem ennyire ismert. A kutatások egy része gátló, a másik része épphogy serkentő hatást mutatott ki, a harmadik részbe tartozó vizsgálatok pedig alkoholmennyiség- és feladattípus-függővé teszik, hogy milyen hatással van az alkohol a rövid távú memóriára. Kutatásunkban az alkohol nyelvtörők visszamondására gyakorolt hatását vizsgáltuk. Ez a feladat egyrészt emlékezeti másrészt beszédprodukciós folyamatokat is tartalmaz, az elemzésben pedig kísérletet tettünk a két jelenség különválasztására.

Szakirodalmi áttekintés

Az alkohol hatása a rövid távú emlékezetre

Több kutatás is kimutatta, hogy az alkohol negatív hatással van a rövidtávú emlékezetre: romlik a képek, a betűk, a szavak, a szövegrészletek felidézése vagy felismerése (ld. Miller és Dolan, 1974; Maylor és Rabbitt, 1987; Nordby, Watten és mtsaik, 1999). Jones és Jones (1977) nemi különbséget is kimutatott: viszonylag kis mennyiségű alkohol a nők rövid távú munkamemória-teljesítményére nagyobb hatással van, mint a férfiakéra.

Jones és Jones (1980) nőknél vizsgálták az alkohol hatását a rövid távú emlékezetre a kor valamint az alkoholfogyasztási szokások függvényében. A kísérleti személyeknek egy szótagú szavakból álló listák elemeit kellett visszamondaniuk megadott mennyiségű alkohol elfogyasztása után. Az úgynevezett „közvetlen emlékezeti pontszámot” a listánként visszamondott szavak számának összege adta. Az összes lista lejátszása után a kísérleti személyeknek még 5 perc állt a rendelkezésére, hogy az összes megtekintett listából felidézzenek szavakat, ez adta az ún. „rövid távú emlékezeti pontszámot”. Az alkoholnak a vér-alkohol görbe emelkedő ívén volt a legnagyobb a hatása, ezért az itt nyújtott teljesítményt vették figyelembe és hasonlították össze a józan állapotban nyújtott teljesítménnyel. Eredményeik szerint az életkor és az alkohol mennyisége befolyással vannak a teljesítményre. Az alkoholnak szignifikánsan nagyobb hatása van a rövid távú emlékezetre a középkorúaknál, mint a fiataloknál, valamint nagyobb a hatása a több alkoholt fogyasztóknál. A közvetlen emlékezet esetében azonban nem tudtak kimutatni ilyen hatást.

Nordby, Watten és munkatársaik (1999) kutatásukban kis és közepes mennyiségű alkohol hatását vizsgálták nyolc szám közvetlen felidézésére, visszamondására. A kutatás azt a jelenséget kívánta modellezni, amikor kimondott illetve leírt telefonszámokat kell ittas állapotban rövid ideig megjegyezni és feltárcsázni. A kísérletben a résztvevőknek nyolc számot kellett egy a telefonéhoz hasonló billentyűzet segítségével visszamondani (-mutatni) abban a sorrendben, ahogy látták illetve hallották őket, és a jól visszaidézett számsorozatok arányát mérték a kutatók. Az eredmények azt mutatták, hogy alkohol hatással volt a számok közvetlen felidézésére: alkoholos állapotban vizuális bemutatás esetén 9, auditoros bemutatás esetén pedig 15 %-kal csökkent a helyesen felidézett sorozatok aránya.

Nuotto és Mattila (1990) áttekintése alapján az etanol mennyiségétől függ, hogy az alkohol rontja-e az emlékezeti feladatokban a teljesítményt. A rövid távú vagy közvetlen emlékezet viszonylag ellenálló az etanol hatásának: 0,4 g/kg-nak nincs hatása a közvetlen felidézésre, a szótanulásra vagy a szófelismerésre. 0,8 g/kg esetében már különböző hatásokat mutattak ki: csökken az újonnan elsajátított információ felidézésének képessége, azonban a már korábban tárolt információ felidézésére, valamint a számterjedelemteszten nyújtott teljesítményre nincs hatással ez a mennyiség. 1,2 g/kg etanol hatására azonban már ez utóbbi feladatban is teljesítménycsökkenés figyelhető meg.

Az alkoholfogyasztás és az emlékezeti feladat egymáshoz viszonyított időpontja is befolyásolja az emlékezeti teljesítményt. Jones (1973) kísérletében a személyek különböző emlékezeti feladatokat végeztek el józan és alkoholos állapotban. Az eredmények azt

mutatták, hogy alkoholos állapotban rosszabbul teljesítettek a kísérleti személyek, mint józanul: csökkent a szavak közvetlen, rövidtávú és hosszú távú helyes felidézésének aránya. Ugyanakkor az alkoholnak nem volt hatása a józan állapotban megtanult szavak felidezésére – sem az alkoholos állapotban, se másnap nem romlott a teljesítmény. Goodwin és munkatársai vizsgálatában (1969) azok a személyek, akik józanul tanulták meg a szavakat és alkoholos állapotban kérdezték ki őket, jobban teljesítettek, mint azok, akik alkoholos állapotban tanultak és józan állapotban kérdezték ki őket.

Egyes kutatások kis mennyiségű alkohol esetében még serkentő hatást is kimutattak (Baum-Baicker, 1985). A korábban (előző nap) józan állapotban megtanult a szövegrészek felidézése szignifikánsan javult alkohol hatására. A szavak visszamondása ugyanakkor alkohol hatása alatt nem változik.

Másrészt ismert az „állapotfüggő tanulás” jelensége (Eich, 1980), amely kimondja, hogy ha valamilyen állapotban (pl. alkohol hatására) valamilyen eseményt élünk át, akkor akkor a legjobb a felidézési teljesítmény az eseményről, ha újból hasonló állapotba kerülünk – tehát javul a teljesítmény, ha a felidezés közbeni állapot megegyezik a kódolás közbeni állapottal. Goodwin és munkatársai vizsgálatukban (1969) kimutatták, hogy 24 órával a tanulás után a személyek jobban teljesítenek, ha a tanulás és a kikérdezés is alkoholos állapotban történt, mint amikor a tanulás alkoholos, de a kikérdezés józan állapotban történt.

Annak, hogy az alkohol bizonyos esetekben javítja a rövidtávú emlékezeti teljesítményt, Tyson és Schirmuly (1994) szerint több magyarázata lehet. A konszolidációs hipotézis szerint az alkoholnak rövid távon neurokémiaileg serkentő hatása van, hogyha a megjegyzendő anyagot az alkohol fogyasztása előtt mutatták be. Az alkoholt fogyasztó kísérleti személyek több szót tudtak felidézni függetlenül attól, hogy a visszakérdezés azonnal történt vagy 40 perc késéssel. Egy olyan feladatban, ahol szó-kép párokat kellett megjegyezniük a kísérleti személyeknek, javult a teljesítmény, ha a tanulás az alkoholfogyasztás előtt történt, és romlott a teljesítmény, ha a tanulás alkoholos állapotban történt. Egy másik magyarázat szerint az alkohol depresszáns hatása rontja az új információk elsajátításának képességét, és javítja az emlékezeti teljesítményt azzal, hogy csökkenti a később érkező információ interferáló hatását.

Az alkohol hatása a beszédre

Általában már viszonylag kis mennyiségű alkohol hatása is hallatszik beszédünkön.

Megváltozhat a beszéd dallama, módosulhat a hangerő és a beszédtempó, elnagyoltta válhat a beszédhangok képzése, végül teljesen felborulhatnak a szünettartási jellemzők. Mindezek következtében a beszéd töredezetté válik, ritmusossága sérül: nő a beszúráások, a kihagyások, a cserék, az ismétlések száma (összefoglalást ld. Chin és Pisoni, 1997). A változásokat a beszélgetőpartnerek is meghallják, és különbséget tudnak tenni az alkoholos és a józan állapotban kimondott mondatok között (Pisoni, Hathaway és Yuchtman, 1985; Pisoni és Martin, 1989).

Pisoni és Martin (1989) összehasonlították kísérleti személyeik beszédét józan és ittas állapotban, és azt találták, hogy alkoholos állapotban nőtt a mondatok hossza, és a hangmagasság variabilitása. A legnagyobb hatás az artikulációban mutatkozott: a felpattanó zöngés hangok (pl. b vagy d) előtti zárások meghosszabbodtak, a záróhangok (pl. cs vagy dzs) előtti zárások pedig teljesen elmaradtak.

Behne, Rivera és Pisoni (1991) kutatásában 9 kísérleti személy vett részt, akikkel egy és két (hosszú) szótagú szavakat, mondatokat és beszédrészleteket mondatnak el józanon, illetve alkoholos befolyásoltság alatt. Eredményeik szerint az egy szótagú szavak, a mondatok és a beszédrészletben lévő mondatok kiejtésének időtartama alkohol hatása alatt szignifikáns növekedést mutatott, két szótagú szavak esetében viszont nem volt különbség a kiejtés időtartama, tehát az artikulációs tempó tekintetében. Pisoni, Hathaway és Yuchtman (1985) vizsgálatában alkohol hatására szintén csökkent az artikulációs tempó. Künzl, Braun és Eyshold (1992) vizsgálatában az elfogyasztott alkohol mennyiségével csökkent a beszélők artikulációs tempója. Ugyanakkor a kis mennyiségű alkohol hatására a személyek beszéd közben kevesebb szünetet tartottak, mint józanul, és csak 1,6 ezrelék véralkoholszint felett nőtt a szünetek száma.

Hollien, Dejong és munkatársaik (2001) azt találták, hogy ahogy nő az ittasság mértéke, vele párhuzamosan emelkedik az alaphangmagasság, nő a közlések ideje és a megakadások száma. Schiel és Heinrich (2009) vizsgálata szerint alkohol hatására megváltozik a beszéd ritmusa és általában nő a hangmagasság. Ez leginkább a nőkre jellemző, a férfiak esetében pedig az is előfordul, hogy csökken a hangmagasságuk.

Összegezve tehát azt mondhatjuk, hogy a legtöbb kutatás szerint alkoholos állapotban a beszéd szünetei hosszabbá és gyakoribbakká válnak, a beszédtempó lelassul, nő az alaphangmagasság, pontatlanabbá válik az artikuláció, egybemosódnak a szóhatárok.

Kérdésfeltevés, hipotézisek

Kutatásunkban egy általunk kifejlesztett teszttel kívántuk mérni, vajon találunk-e különbségeket alkohol hatása alatt és józan állapotban a személyek beszédében, rontja-e az alkoholos befolyásoltság az egyének munkamemóriáját és beszédprodukciónak. Vizsgálni kívántuk még a beszéd tempóját, az elakadások és szünetek számát, továbbá a hangmagasság változását.

Hipotézisünk szerint alkoholos állapotban romlik a személyek számterjedelemteszten nyújtott teljesítménye; több hibát ejtenek beszéd közben, amelyek egy része a rövid távú emlékezet romlásából, másik része pedig az artikuláció pontatlanságából, a beszédhez szükséges izmok működésének megváltozásából fakad. Feltételezzük ezen kívül, hogy alkoholos állapotban a személyek artikulációs tempója lassul, több szünetet tartanak a beszédben, és hangmagasságuk megemelkedik.

Módszer

Vizsgálati személyek

A kísérletben 15 felsőfokú tanulmányokat folytató fiatal vett részt: 7 lány és 8 fiú. Átlagéletkoruk 20,73 ($s=1,79$). A kísérletre önkéntesen jelentkeztek. A vizsgálatban való részvételért semminemű juttatásban nem részesültek.

A vizsgálatot két próbateszt előzte meg, mind a két esetben 2-2 fővel: egy fiúval és egy lánnyal. Átlagéletkoruk 19,25 év.

Eszközök

A kísérlet mérőeszköze egy általunk összeállított feladatsor volt. Ezt kiegészítette egy adatfelvétel, amely tartalmazta az AUDIT (Babor, Biddle-Higgins és mtsaik, 2001) néhány kérdését, mellyel a diákok ivási szokásairól kívántunk tájékozódni. Az AUDIT (Alcohol Use Disorders Identification Test) egy tíz kérdésből álló teszt, melyet a WHO fejlesztett ki 1989-

ben, hogy azonosíthassa vele azon alkoholfogyasztókat, akiknek ivási szokásai veszélyesek, károsak. Kutatásunkban ezen teszt 3 kérdését fordítottuk magyarra és használtuk fel: milyen gyakran isznak alkoholt, egyszerre hány egységnyit, és milyen gyakran isznak 5 egységnél többet. Egy egység italnak számított fél üveg sör, 1 dl bor vagy 4 cl rövidital.

A kiegészítő adatoknál nemem, életkoron, végzettségen kívül rákérdeztünk még a vizsgálati személyek testsúlyára, illetve arra, hogy a vizsgálat előtt mikor és pontosan mit ettek utoljára. Arra a kérdésre is válaszolniuk kellett, hogy az alkohol általában hogyan hat aktivitásukra. Továbbá be kellett írniuk, hogy a felkínált alkoholok közül melyiket választották, és egy hétfokú skálán jelölniük kellett, hogy teszt felvételekor mennyire érezték magukat ittasnak

Maga a feladatsor tartalmazott egy számterjedelemtesztet, nyelvtörőket és idegen szavakat. A számterjedelemtesztet eredetileg Jacobs dolgozta ki 1887-ben iskolai tanulók emlékezeti képességének vizsgálatára (ld. Racsmány, Lukács és mtsaik, 2005). A tesztben a kísérletvezető számokat olvas fel a kísérleti alanynak, 1 másodperc szünetet tartva minden szám után. A feladat a számsorozat pontos megismétlése. A kísérleti személy rövid távú emlékezetének terjedelmét az a sorozathosszúság jelöli, amelyből két számsorozatot pontosan megismétel. A teszt növekvő terjedelmű számsorozatokot használ. A magyarországi életkori standardokat Racsmány és munkatársai felmérése alapján ismerjük (2005).

A nyelvtörőket a Pázmány Péter Katolikus Egyetem Bölcsészettudományi Kar nyelvész hallgatóinak segítségével osztályoztuk, és négyféle kategóriába soroltuk őket. 10 fonetikai (pl. *pepita bóbitás babapókok pattognak a bobbályán*), 5 fonológiai (a magyar nyelvben használatos minimálpárok pl. *két kék fék fér,*), 4 morfológiai (pl. a *lassan látni* felszólító módú ragozási paradigmája) és 6 szintaktikai (*Már itt volt ősz szinte, mikor egy őszinte ősz inte, hogy legyek őszinte, mert ő szinte őszinte*) nyelvtörőt használtunk. A feladatot 10 idegen szó zárta (pl. *vucsicsevázok*).

A nyelvtörőket és az idegen szavakat a vizsgált alanyok hangfelvételtől hallgatták. Válaszaikat számítógépen egy jó minőségű dinamikus mikrofonnal (Pioneer DM-DV5) rögzítettük a PRAAT program segítségével (www.praat.org). A PRAAT egy ingyenesen hozzáférhető és szabadon használható számítógépes program a beszéd fonetikai vizsgálatára. Paul Boersma és David Weenink (2009) fejlesztették ki. A hangfelvételeket a későbbiekben ezzel programmal is elemeztük.

Eljárás

A vizsgálat a Budapesti Műszaki és Gazdaságtudományi Egyetem Kognitív Tudományi Tanszékének kutatási-etikai engedélyével készült.

A vizsgálatot két előkísérlet előzte meg, amelyek alapján meghatároztuk, hogy mennyi idő teljen el az alkoholfogyasztás és a feladatsor felvétele között, mikor tudunk valamilyen hatást megfigyelni, továbbá kiszűrtük a konvencionális nyelvtörőket, melyeket sokan ismernek. Az előkísérlet tapasztalatai alapján határoztuk meg a vizsgálat pontos menetét is. A két előkísérlet eredményeit kutatásunk során nem vettük figyelembe.

Minden kísérleti személy részt vett mind az alkoholos mind a kontroll feltételben, az egy személyen belüli változásokra voltunk kíváncsiak. Először az alkoholos vizsgálatot végeztük el. A vizsgálatok késő délután zajlottak egy csendes helyiségben. A résztvevőket megkértük, hogy a kísérlet előtt minimum 4 órával ne egyenek. Ezzel kívántuk biztosítani, hogy a vizsgálat éhgyomorra történjen, ezzel is fokozva az alkohol hatását (Tringer, 2005). Miután a személyeket tájékoztattuk a kísérlet menetéről, és aláírták a beleegyező nyilatkozatot, három alkohol közül választhattak: tequila, pálinka vagy vodka. Azért erre a három típusú italra esett választásunk, mert színük azonos, alkoholtartalmuk közel egyenlő, emiatt felszívódásuk is hasonló sebességgel megy végbe. Összesen háromszor 5 cl-t kellett elfogyasztaniuk a lehető legrövidebb idő alatt (maximum 5 perc). Ezt követően egy órát vártunk, hogy az alkohol felszívódjon szervezetükbe, hiszen a vér alkoholkoncentrációja a csúcértéket körülbelül ennyi idő múlva éri el (Tringer, 2005).

A feladatsor felvétele a következő módon zajlott: a kísérleti személy fejhallgatón keresztül hallotta az instrukciót, mely szerint hangosan el kellett ismételnie a mikrofonba a hallott mondatokat, szavakat, illetve el kellett végeznie az adott ragozásokat (ld. morfológiai nyelvtörők). Ezek után a fejhallgatón keresztül nyelvtörőket illetve idegen szavakat hallott randomizált sorrendben. A szintaktikai nyelvtörők előtt mindig jeleztük neki, hogy az adott mondatot kétszer fogja hallani. Mindeközben válaszait számítógépen a PRAAT program segítségével rögzítettük. A kísérleti személyek felével először az adatokat és a számterjedelemtesztet vettük fel, a másik felükkel a sorrend megcserélődött.

A kísérlet után két héttel a diákok újra felkerestek bennünket, és most józanul mondták vissza az adott nyelvtörőket, idegen szavakat, és oldották meg a számterjedelemtesztet. Azt feltételeztük, hogy két hét elegendő idő lesz ahhoz, hogy a kísérleti személyek elfelejték a

nyelvtörőket, és ez ne befolyásolja teljesítményüket a második alkalommal. Mivel mindkét feltételben ugyanaz volt az ingeranyag, az „alkoholos→józan” sorrendválasztással az volt a célunk, hogy ne növeljük az alkoholos feltételben nyújtott teljesítményt azáltal, hogy elősegítjük a korábban már hallott mondatok felidézését. Így ha különbség mutatkozik a mondatok visszamondásának pontosságában alkoholos és józan feltétel között, akkor az inkább az alkohol hatásának tulajdonítható, mint az emlékezetének, mint ha a sorrend fordított lenne.

Eredmények

Az alkoholfogyasztásra vonatkozó adatok

A vizsgált személyek átlagosan 5,67 órával ($s=1,68$) ettek utoljára a kísérlet megkezdése előtt. Átlagosan heti két-három alkalommal fogyasztottak alkoholt. Ilyenkor 5-6 egységnyi italt isznak. Nők és férfiak között egyik esetben sem volt szignifikáns különbség. Általában havonta szoktak öt egységnél többet inni. Saját állítása szerint 12 alanyt általában felpörget az alkohol, míg 3-at inkább lelassít. Egy hétfokú skálán 4,19 ($s=1,38$) mértékben érezték ittasnak magukat. Kilenc ember fogyasztott vodkát, négy tequilát és kettő pálinkát. A férfiak átlag testsúlya 77,88 kg ($s=9,14$) a nőké pedig 59,14 kg ($s=7,71$) volt. Vizsgálatunkban kiszámoltuk kísérleti alanyaink véralkoholszintjét is (elfogyasztott alkohol (etanol) mennyisége/testtömeg g/kg-ban kifejezve, jelen esetben 47,25 g etanol, ld. AUDIT (Babor, Biddle-Higgins és mtsaik, 2001). Mivel a férfiak és a nők különböznek tekintetben, hogy mennyi testüknek a vízszázaléka, és ez befolyásolja a véralkoholszintet, ezért 1,12%-os szorzót használtunk a nők véralkoholszintjének meghatározásakor (ld. Kalapos, 2007). Így kutatásunkban a férfiak véralkohol szintje átlagosan 0,62 g/kg ($s=0,07$) volt, nőké pedig 0,9 g/kg ($s=0,1$).

A számterjedelem alkoholos és józan állapotban

Az egymintás t-próba eredménye alapján a kontroll feltételben a számterjedelemteszten elért eredmények (Átlag: 6,13, s=1,25) megfeleltek az életkori standardnak ($t=-0,363$, $p=0,722$) (ld. Racsmány, Lukács és mtsaik, 2005).

Annak eldöntésére, hogy vajon az alkoholos feltételben tényleg romlott-e a teljesítmény, páros (összetartozó mintás) t-próbát alkalmaztunk. Hipotézisünkkel ellentétben nem volt szignifikáns különbség a kontroll (Átlag: 6,13, s=1,25) illetve az alkoholos (Átlag: 5,93, s=1,1) feltételben mért számterjedelem között ($df=14$, $t=-0,764$, $p=0,458$). Tehát alkohol hatása alatt nagyjából ugyanannyi számra emlékeztek a kísérleti személyek, mint kontroll feltételben.

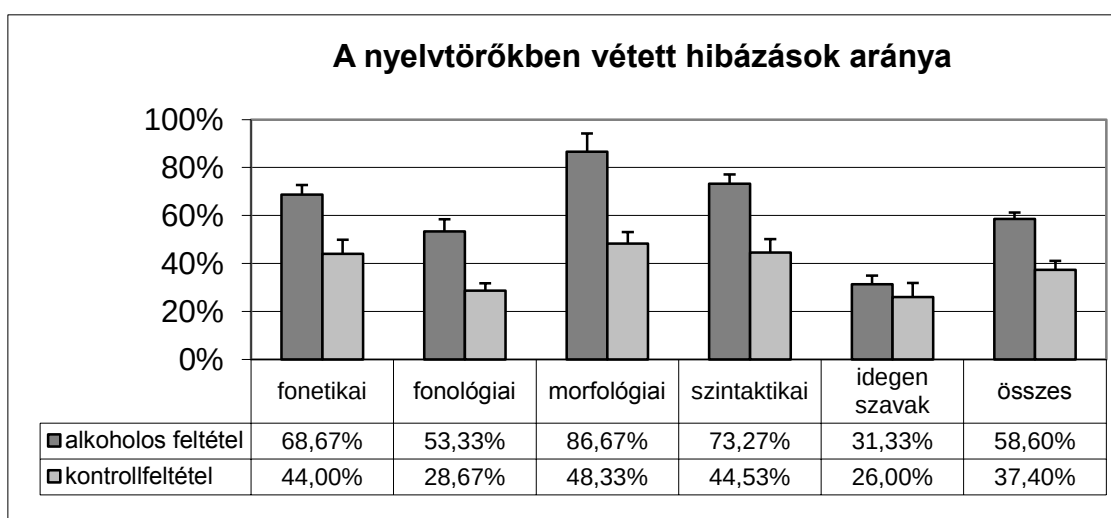
A nyelvtörőkben mutatott hibázások

Annak eldöntésére, hogy a nyelvtörőkben többet hibáznak-e a kísérleti személyek alkoholos befolyás alatt, mint józanul, szintén páros t-próbát alkalmaztunk. Kutatásunkban hibának tekintettünk minden egyes rontást (pl. nyelvbtlás, elakadás stb.). Az eredmények alátámasztották elvárásainkat: szignifikáns különbség volt a kontroll (Átlag: 13,07, s=4,93) és az alkoholos (Átlag: 20,53, s=3,58) feltétel között. Alkoholos állapotban a kísérleti személyek több hibát ejtettek ($t=8,456$, $df=14$, $p<0,05$). A vizsgálati alanyok mind a négy nyelvi szinten többet hibáztak alkohol hatása alatt (ld. 1. táblázat). Az idegen szavak esetében nem volt kimutatható különbség.

	Alkoholos feltétel	Kontrollfeltétel	t (df=14)	p
Fonetikai nyelvtörő	6,87 (0,4)	4,4 (0,58)	t=5,819	p<0,001
Fonológiai nyelvtörő	2,67 (0,25)	1,47 (0,17)	t=5,392,	p<0,001
Morfológiai nyelvtörő	3,47 (0,19)	1,93 (0,3)	t=4,38,	p<0,05
Szintaktikai nyelvtörő	4,4 (0,24)	2,67 (0,33)	t=5,773	p<0,001
Összesen	20,53 (0,93)	13,07 (1,27)	t=8,456	p<0,001

1. táblázat: A nyelvtörőkben elkövetett hibázások átlagos száma (zárójelben az átlag szórása – a SEM) típusonként alkoholos és józan állapotban.

A hibaátlagok százalék szerinti megoszlását az 1. ábra mutatja.



1. ábra: Hibaátlagok százalékos megoszlása és a SEM az egyes nyelvtörőtípusoknál és idegen szavak esetén alkoholos és kontroll feltételben. A 100 %-ot az adott típusba tartozó összes nyelvtörő száma jelenti.

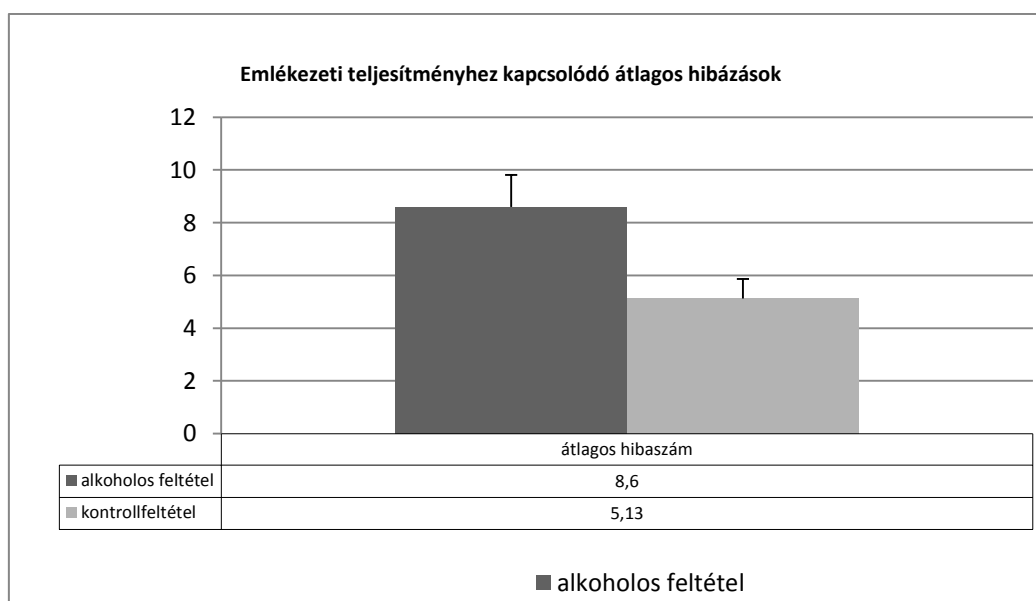
Kutatásunkban azt is megvizsgáltuk, hogyan van-e eltérés az egyes nyelvtörők között a hibázások mértékében. Ennek eldöntésére az alkoholos feltétel hibázásaiból kivontuk a kontroll feltétel hibázásait, ez a szám lett a hibázás mértéke minden esetben. A páros t-próba

alapján idegen szavaknál a vizsgálati alanyok kevesebbet hibáztak, mint fonetikai ($t=2,527$, $df=14$, $p<0,05$), fonológiai ($t=4,005$, $df=14$, $p<0,05$), morfológiai ($t=3,498$, $df=14$, $p<0,05$) vagy szintaktikai nyelvtörők ($t=3,032$, $df=14$, $p<0,05$) esetén.

A rövid távú emlékezeti teljesítményhez és a pontatlan artikulációhoz kapcsolódó hibázások alkoholos befolyás alatt

A további elemzésben különválasztottuk a hibatípusokat az alapján, hogy azokat az emlékezeti terjedelem romlása okozta vagy az artikuláció pontatlanságából eredtek-e. A **nyelvtörők felidézésében mutatkozó munkamemória-teljesítmény** vizsgálatára három mutatót hoztunk létre. Az *elakadás* alatt azt értettük, amikor a kísérleti személy nem tudta végigmondani, illetve befejezni a nyelvtörőt. *Rontásnak* azokat a hibázásokat vettük, amikor a kísérleti személy más szót mondott, vagyis hibásan idézte fel az adott szót. A *kihagyások* közé azt soroltuk, amikor valaki kifelejtett egy vagy több szót a mondatból.

Ezt követően mind az alkoholos, mind a kontrollfeltételben összesítettük a hibázásokat. Az eredményeket a 2. ábra mutatja.

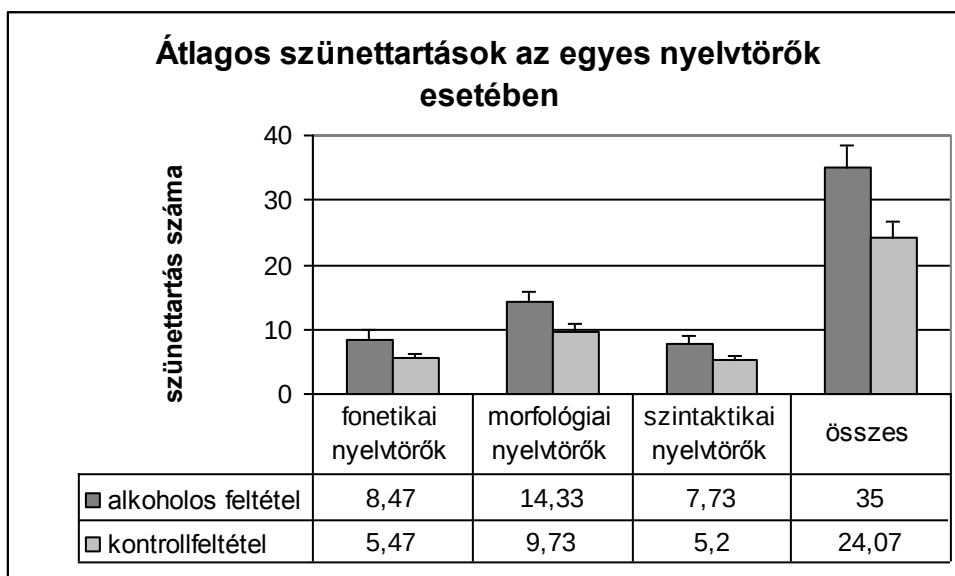


2. ábra: Az emlékezeti teljesítményhez kapcsolódó átlagos hibázások és a SEM alkoholos állapotban és a kontrollfeltételben

A páros t-próba alapján ittas állapotban rosszabbul teljesítettek a kísérleti személyek a nyelvtörők felidőzésében, mint józan állapotban (nyelvtörők felidőzésében vétett hibák átlaga ittasan: 8,60 és józanul: 5,13; $t=3,57$, $df=14$, $p<0,05$).

A szünetek száma

Kísérletünkben vizsgáltuk a szünetek számát is, azaz minden egyes szünettartás függetlenül attól, hogy helyesen vagy helytelenül mondja-e a személy az adott nyelvtörőt. A páros t-próba szignifikáns különbséget mutatott a szünetek számában az alkoholos és kontroll feltétel között: alkoholos feltételben nőtt a szünetek száma ($t=-3,081$, $df=14$, $p<0,05$).



3. ábra Az átlagos szünettartások száma és a SEM alkoholos és kontroll feltételben (csak a szignifikáns különbségeket ábrázoltuk)

Egyes részfeladatokra lebontva a következő eredményt kaptuk: a szünetek száma szignifikánsan eltért a fonetikai nyelvtörők ($t=2,717$, $df=14$, $p<0,05$), a morfológiai ($t=3,123$, $df=14$, $p<0,05$) és szintaktikai feladatok ($t=2,172$, $df=14$, $p<0,05$) esetén, míg a fonológiai ($t=1,219$, $df=14$, $p=0,243$) nyelvtörőknél illetve az idegen szavaknál nem volt kimutatható különbség (ld. 3. ábra).

Emlékezeti teljesítménybeli összefüggések

Megnéztük, hogy a számterjedelemteszten nyújtott teljesítmény és a hibázások száma között van-e korreláció. Alkohol hatása alatt a számterjedelem negatívan korrelált az összes hiba számával ($r=-0,625$, $p<0,05$), a szintaktikai nyelvtörőkben elkövetett hibák számával ($r=-0,542$, $p<0,05$) és az idegen szavak visszamondásában ejtett hibák számával ($r=-0,548$, $p<0,05$). Tehát alkohol hatása alatt azon vizsgálati alanyok, akiknek számterjedelme nagyobb volt, kevesebb hibát vétettek összességében, kevesebb hibát vétettek a szintaktikai nyelvtörőkben (habár megjegyezzük, hogy ezen nyelvtöröket a kísérleti személyek kétszer is hallották), és az idegen szavakat is pontosabban vissza tudták mondani. Ezzel szemben a kontroll feltételben a számterjedelem csak a fonetikai nyelvtörőkben elkövetett hibák számával korrelált negatívan ($r=-0,629$, $p<0,05$).

Ezen felül vizsgáltuk, hogy vajon a nyelvtörök elmondása alatt tartott szünetek száma korrelál-e a számterjedelemmel. Alkohol hatása alatt nem találtunk korrelációt sem összesítve ($r=0,134$, $p=0,633$), sem az egyes nyelvtörőtípusokra lebontva. A kontroll helyzetben egyedül a fonetikai nyelvtöröknél a szünetek száma negatívan korrelált a számterjedelemmel ($r=-0,542$, $p<0,05$). Tehát minél kisebb volt a kísérleti alany számterjedelme, annál több szünetet tartott a fonetikai nyelvtöröknél. Összesítve kontroll helyzetben sem korrelált a számterjedelem és a szünetek száma ($r=-0,196$, $p=0,484$).

A számterjedelemteszten mutatott teljesítményromlás és a hibázásokkal mért munkamemória-teljesítményromlás közötti összefüggést is megvizsgáltuk. A teljesítményromlást a józan és az alkoholos állapotban nyújtott teljesítmények különbségei adták mindkét mutatóra, ezek azonban nem korreláltak egymással. Úgy tűnik tehát, hogy nincs összefüggés az alkohol hatására bekövetkezett számterjedelem-csökkenés és a nyelvtörök felidézését érintő munkamemória-teljesítmény csökkenése között ($r=-0,28$, $p=0,31$).

A pontatlan artikulációból fakadó hibák

Az összes hiba és az emlékezeti teljesítményből fakadó hibák száma alapján a százalékos megoszlás azt mutatja, hogy alkoholos állapotban a hibák 58%-áért a pontatlan artikuláció

felel és 42 %-uk fakad az emlékezeti teljesítmény romlásából, józan állapotban ehhez hasonlóan 61 és 39 % az arány.

A részletesebb elemzésben a fonetikai típusú nyelvtörőket vettük csak figyelembe, és megállapítottuk a legtipikusabb magánhangzó- illetve mássalhangzó-cseréket. Ennek bemutatása inkább kvalitatív, mint kvantitatív.

Mássalhangzókkal kapcsolatos hibázások a fonetikai feladatokban

A mássalhangzókkal kapcsolatos hibázásokat a zöngés-zöngétlen mássalhangzó párok felcserélése szerint elemeztük (ld. 2. táblázat). A fonetikai nyelvtörőkben az alábbi hibatípusok fordultak elő¹¹:

B-P hangok cseréje: **P**epita **b**óbítás **b**abapókok **p**attognak a **b**obpályán. (10 kritikus msh.)

D-T hangok cseréje: **T**óth **T**ódor **d**oktora **T**óth **D**orka **d**oktora is. (8 kritikus msh.)

Z-SZ és *ZS-S* hangok cseréje: **Z**abra **z**ebra **z**sebre **z**abra. (4 kritikus msh.)

S-ZS, DZS-CS hangok: **S**óska **z**súri **cs**acska **z**súri. (összesen 6 kritikus msh.)

A nyelvtörőkben nem minden esetben a zöngés-zöngétlen mássalhangzó párok szerinti cserét követtek el a kísérleti személyek. Így bár a „zabra zebra zsebre zabra” nyelvtörőben a *z* hang és *zs* hang gyakori cseréje figyelhető meg; ugyanakkor pl. a „sóska zsúri csacska zsúri” nyelvtörőben nemcsak a *zs* és *s* hangok, hanem ezek és a *dzs*-vel való cseréje is előfordult.

Mássalhangzócsere	Alkoholos állapot	Józan állapot
B-P	5	7
D-T	7	4
Z-ZS	9	8
ZS-DZS-S	5	1

2. táblázat Mássalhangzócserek előfordulása (db)

¹¹ Az egyszerűség kedvéért a hangokat a magyar ábécé betűivel és kurzívan jelöltük

A mássalhangzók tekintetében nem találtunk olyan hibatípust, ami bejósolná az alkoholos hatást. Az egyes cserék alkoholos és józan állapotban is bekövetkeztek. Bár a leggyakoribb a zöngés-zöngétlen csere volt, ez magukból a nyelvtörőkből is fakadt.

Magánhangzókkal kapcsolatos hibázások a fonetikai feladatokban

A magánhangzókkal vétett hibák elemzéséhez az alább olvasható nyelvtörőket vizsgáltuk meg, és a következő hibatípusokat találtuk:

Sóska zsúri csacska zsúri → *Ú* hang helyett *Ó* ejtése

Ű hang helyett *Ő* hang ejtése

Ú hang helyett *Ű* hang ejtése

Zabra zebra zsebre zabra → *E* hang helyett *A* hang ejtése

A hang helyett *E* hang ejtése

	<i>Hiba mértéke</i>	
	<i>ALKOHOL</i>	<i>KONTROLL</i>
<i>Ú</i> → <i>Ó</i> csere	53.3%	46.7%
<i>Ű</i> → <i>Ő</i> csere	46.7%	20%
<i>Ú</i> → <i>Ű</i> csere	6.7%	0%
<i>E</i> → <i>A</i> csere	66.7%	32.3%

3. táblázat A magánhangzócserekre vonatkozó hibázások aránya alkoholos és józan állapotban

Páros t-próbát végeztünk annak ellenőrzésére, hogy alkoholos és nem alkoholos hatás alatt van-e különbség a magánhangzók felcserélésének aránya között. Az *A-E* hangok felcserélése alkohol hatása alatt szignifikánsan nagyobb mértékű volt, mint józan állapotban ($t=2,65$, $df=14$, $p<0,05$). *Ű* hang helyett *Ő* hang ejtése alkohol hatása alatt szintén szignifikánsan nagyobb mértékben fordult elő, mint a kontrollfeltételben ($t=2,26$, $df=14$, $p<0,05$). A többi magánhangzókkal kapcsolatos hibázás nem mutatott szignifikáns különbséget az alkoholos és nem alkoholos állapot között (ld. 3. táblázat)

Hangmagasság

A hangmagasság vizsgálatánál a kísérleti személyek összes válaszát felhasználtuk, minden egyes nyelvtörő hangmagasságát egyenként megmértük az esetleges nevetések, köhintések kiszűrésével.

A kérdés eldöntésére, hogy a hangmagasság változik-e alkohol hatására, páros t-próbát alkalmaztunk, tehát a személyen belüli változásra voltunk kíváncsiak. Vizsgálatunkban nem találtunk szignifikáns különbséget a két feltétel között hangmagasság tekintetében ($t=1,267$, $df=14$, $p=0,226$).

Az artikulációs tempó módosulása

Az artikulációs tempó mérésekor a kísérleti személyek minden egyes válaszát figyelembe vettük, a nyelvtörőket egyenként vizsgáltuk. Artikulációs tempónak tekintettük az egy másodperc alatt kimondott hangok számát szünetek nélkül. A páros t-próba itt sem mutatott szignifikáns különbséget, ha a nyelvtörőket és az idegenszavakat összességében tekintettük ($t=1,165$, $df=14$, $p=0,264$). Azonban a részfeladatokra való lebontásnál a szintaktikai nyelvtörők esetén szignifikáns eltérést kaptunk: alkoholos befolyás alatt az artikulációs tempó lassult a kontroll feltételhez képest (átlag alkoholos feltételben: 8,61, kontroll feltételben: 9,04; $t=-2,264$, $df=14$, $p<0,05$).

A véralkoholszint és a fenti mutatók összefüggései

Vizsgáltuk, hogy vajon a véralkoholszint mutat-e a fenti mutatókkal illetve ezek változásával valamilyen összefüggést. Egyrészt korreláltattuk a véralkoholszintet és az alkoholos állapotban mért hibázások számát, szünetek számát, számterjedelem teszten elért pontszámot, a hangmagasságot és az artikulációs tempót. Másrészt ezekre a mutatókra kiszámoltunk egy ún. „változás mutatót”, amely minden esetben az alkoholos feltétel mínusz a kontroll feltétel eredménye adta, és ezzel a változás mutatóval is korreláltattuk a véralkoholszintet (Spearman-féle korrelációt használtunk minden esetben).

A véralkoholszint és az alkoholos feltételben mért mutatók közül szignifikáns korreláció mutatkozott a hangmagasság esetében (az összes hanganyagot figyelembe véve: $r=0,805$, $p<0,001$). Ez tükrözi a nemi különbségeket: a nőknek, akik testsúlyuk miatt is magasabb véralkoholszinttel rendelkeztek, mint a férfiak (nők átlag: 0,9 g/kg, férfiak átlag: 0,62 g/kg; $t=-6,529$, $df=13$, $p<0,001$), magasabb volt a hangmagasságuk, mint a férfiaknak (nők átlag: 219,3 Hz, férfiak átlag: 126,4 Hz; $t=-11,433$, $df=13$, $p<0,001$).

Ezen kívül a véralkoholszint szignifikánsan korrelált alkoholos feltételben a szintaktikai nyelvtörök elmondásakor mutatott artikulációs tempó mértékével ($r=-0,59$, $p<0,05$). Minél magasabb volt valakinek a véralkoholszintje, annál lassabban artikulált.

A változásra vonatkozóan nem találtunk semmilyen szignifikáns korrelációt.

Diszkusszió

Kutatásunk célja egy olyan vizsgálati eljárás kialakítása volt, amelynek a segítségével esetlegesen a jövőben kiszűrhetőek lennének az alkoholos befolyás alatt álló személyek. Vizsgálatunk középpontjába a beszédprodukción és a rövidtávú memóriát helyeztük: azaz célunk olyan beszédprodukción és memóriazavarok felfedése volt, amelyek általánosságban jellemzők az ittas egyénekre. Ehhez különböző nyelvtörőket, idegen szavakat, továbbá a számterjedelemtesztet alkalmaztuk. A nyelvtörőket a nyelv szintjei (fonetika, fonológia, morfológia, szintaxis) szerint is osztályoztuk, hogy ezáltal is jobban differenciálhassuk az esetlegesen megjelenő hatásokat.

A kísérleti feltételben a vizsgálati személyeknek 3x5 cl égetett szeszesitalt kellett elfogyasztaniuk. Bár a legtöbb kutatás vízbe vagy narancslébe keverve, hígítva kínálja fel az alkoholt a vizsgálati személyek számára, a tömény ital fogyasztása sem áll ellentétben a kutatási előzményekkel (ld. pl. Künzel, Braun, Eyshold, 1992). Jelen kutatásban az alkohol mennyisége minden kísérleti személy számára azonos volt, így ez a testsúlybeli és a nemi különbségek miatt eltérő véralkoholszintet eredményezett. Ennek korrigálására a véralkoholszinttel mint független változóval való számításokat végeztünk. Ugyanakkor mivel a személyek háromféle alkohol (tequila, pálinka és vodka) közül választhattak, az ezekben lévő egyéb anyagok illetve a minimális ethanoltartalombeli különbségük miatt a véralkoholszint meghatározása is pontatlanabb lehet.

A vizsgálat nagyobb részét a nyelvtörők visszamondása során ejtett hibák elemzése tette ki. Feltételezéseinket alátámasztották eredményeink: mind a négy nyelvi szinten szignifikáns különbséget találtunk: ittas állapotban nőtt a hibázások száma. Ezzel ellentétben idegen szavak esetén nem találtunk szignifikáns eltérést. Kitértünk arra is, hogy vajon eltértek-e az egyes nyelvtörők illetve az idegen szavak annak tekintetében, hogy a vizsgálati alanyok milyen mértékben hibáztak bennük. Eredményeink alapján a kísérleti személyek az idegen szavaknál szignifikánsan kevesebbet hibáztak, mint a többi nyelvtörőnél. Úgy tűnik tehát, hogy ez a feladat tehát nem jó mutatója az ittasságnak.

Az alkoholnak a beszédre gyakorolt hatásának okairól kevés magyarázat született. Christenfeld és Creager (1996) kutatásában alkoholos befolyás alatt a személyek beszédében csökkent egy bizonyos típusú megakadás („the ums”) aránya. A jelenséget a szerzők azzal magyarázták, hogy alkoholos állapotban a személyek kevésbé monitorozzák önmagukat és kevésbé öntudatosak, ami ahhoz vezet, hogy beszédük automatikusabbá, folyamatosabbá válik, és ennek következtében kevesebb lesz benne a hezitáció. Ugyanakkor más kutatások (Forney és Hughes, 1961; Hollien, Dejong és mtsaik, 2001; Künzel, Braun, Eyshold, 1992; Zaimov, 1969) azt találták, hogy alkoholos állapotban pont hogy nő a megakadások, hibák száma. A mi vizsgálatunk is ezek közé a kutatások közé tartozik. Jelen esetben a hibák a beszédprodukciónak a kivitelezés részéhez, a motoros-artikulációs szinthez kapcsolhatók (Levelt, 1989), nem beszélhetünk a megfogalmazás vagy a szavakba öntés hibájáról, mivel a feladat mondatok, kifejezések megismétlése volt. Santamaria (1989, ld. Chin és Pisoni, 1997) arról számol be, hogy már 0,05-0,06%-os véralkoholszint is befolyásolja az izomműködés kontrollját. Mivel a beszédprodukcióhoz pontos izomműködésre van szükség, ha ez hiányzik, nő a beszéd pontatlansága, a nyelvbtlások száma. Tehát azt mondhatjuk, hogy bár alkohol hatására csökken az önmonitorozás mértéke, ha a feladata nehéz, és koncentrációt, figyelmet igényel, valamint jelen van az alkohol artikulációs folyamatokra gyakorolt negatív hatása, akkor ezek összességében növelik a beszédbeli hibák előfordulásának arányát.

Megvizsgáltuk, hogy a hibázások száma összefügg-e alkoholos és józan állapotban, de csak a fonetikai nyelvtörők esetében találtunk pozitív korrelációt. Tehát azok a személyek, akik józanul többet hibáztak, alkoholos állapotban is több hibát vétettek a fonetikai nyelvtörők visszamondása esetén. Összességében és a többi nyelvi szinten azonban nem találtunk ilyen együttjárást.

Az emlékezeti teljesítményre gyakorolt hatás

A számterjedelemtesztrel mért rövid távú emlékezet esetleges romlását vizsgáltuk alkoholos befolyás alatt. Azonban kísérleti alanyainknál nem volt megfigyelhető szignifikáns különbség a két feltétel között: feltételezéseinkkel ellentétben az alkohol nem rontotta a résztvevők számterjedelemtesztrel mért emlékezeti teljesítményét, sőt számos kísérleti személynek a számterjedelme nőtt ittas állapotban. Kísérletünkben valószínűleg az alkohol mennyisége nem volt hatással a számterjedelemre, esetleg nagyobb mennyiség esetén várhatnánk változást, hiszen Nuotto és Mattila (1990) összefoglalója alapján a 0,8 g/kg etanolnak még nincs hatása a számterjedelemre, a jelen kísérletben résztvevők átlagos véralkoholszintje pedig ekörül az érték körül volt, illetve férfiak esetében ennél alacsonyabb volt. Mivel a számterjedelem alkoholos állapotban nem mutatott összefüggést a véralkoholszinttel, ezért valószínűsíthető, hogy egy küszöbértékről van szó: egy bizonyos véralkoholszint felett mutatható csak ki az alkohol hatása a rövid távú memóriára.

A következőkben részletesebben megvizsgáltuk, hogy a hibázásoknak mi lehet az oka. Az egyik feltételezésünk az volt, hogy ezek a munkamemória-teljesítmény romlásából következnek, a másik pedig, hogy az alkohol hatására bekövetkező izomlazulás okozta artikulációs pontatlanságból. A nyelvtörőkben produkált, munkamemória-teljesítményt érintő hibázások között előfordultak elakadások, rontások és kihagyások. Az eredmények azt mutatták, hogy alkoholos állapotban többet hibáztak a vizsgálati személyek nyelvtörök felidézésében, mint józanul.

Szintén emlékezeti mutatónak vettük a visszamondás során tartott szünetek számát. Alkoholos állapotban a személyek a legtöbb nyelvi szinten több szünetet tartottak, mint józan állapotban. A szünetek száma azonban alkoholos állapotban nem korrelált a számterjedelemmel. Józan állapotban pedig csak a fonetikai nyelvtörök esetében találtunk negatív korrelációt, akárcsak a hibázások esetében: tehát azok a személyek, akiknek jobb volt a számterjedelmi teljesítménye, józanul kevesebbet hibázott és kevesebb szünetet tartott a fonetikai nyelvtörök visszamondásakor.

Az emlékezeti teljesítménnyel kapcsolatos teljesítmény romlása (elakadás, rontás vagy kihagyás) és a számterjedelemteszten mutatott teljesítményváltozás között nem találtunk együttjárást. Tehát bár az alkohol nem volt hatással a számterjedelemteszten nyújtott teljesítményre, hatással volt a nyelvtörökre való emlékezésre és azonnali visszamondásra.

Úgy tűnik tehát, hogy a számterjedelemteszt és a nyelvtörők visszamondása más emlékezeti kapacitást igényel, hiszen másképp hat a két folyamatra az alkohol. A munkamemória-modellben (ld. Baddeley, 2003) értelmezve valószínű, hogy míg a számterjedelemteszt pusztán a fonológiai hurok kapacitását veszi igénybe, addig a hosszabb mondatok fejbentartása és visszamondása a végrehajtó funkciókat is terheli, amelyre az alkoholnak már negatív hatása van. Ennek ellenőrzésére érdemes lenne a számterjedelem visszafelé tesztet elvégezni alkoholos befolyásoltság alatt, hiszen ez a teszt a végrehajtó funkciókat is igénybe veszi (Gathercole és McCarthy, 1994, Németh és Pléh, 2001 összefoglalója alapján). Nyberg és Tulving (1996) is hangsúlyozzák, hogy bizonyos drogok másképp hathatnak a különböző emlékezeti folyamatokra, lehetséges, hogy míg az egyik emlékezeti feladaton romlik a teljesítmény, addig ugyanannak a drognak a hatására a másik emlékezeti feladaton nem változik a teljesítmény.

Felmerül, hogy a hibázásokat valamiféle elvárási hatás is befolyásolja, tehát hogy azok a személyek, akik ittasabbnak érzik magukat, már eleve úgy gondolják, hogy többet fognak hibázni, és ez meg is mutatkozik a teljesítményükben, vagy pont fordítva: jobban odafigyelnek, és ezért kevesebbet hibáznak. Ezt cáfolják azonban eredményeink: a szubjektívan megítélt részegség és a rontások száma nem korrelált egymással ($r=-0,023$, $p=0,935$). Tehát azon kísérleti személyek, akik józanabbnak érezték magukat, nem hibáztak kevesebbet vagy többet azon alanyoknál, akik jobban érezték az alkohol hatását.

Pontatlan artikuláció: Fonetikai feladatokban elkövetett hibázások

Az összes hiba száma és az emlékezeti teljesítményből fakadó hibák száma alapján kiszámítottuk, hogy a pontatlan artikuláció józan állapotban a hibák 61 %-áért, alkoholos állapotban pedig a hibák 58 %-áért felel. A mássalhangzók tekintetében nem találtunk olyan hibatípust, ami bejósolná az alkoholos hatást. Bár a leggyakoribb a zöngés-zöngétlen csere volt, ez magukból a nyelvtörőkből is fakadt. A vizsgált magánhangzók közül az A-E felcserélése és az Ű-Ő felcserélésének aránya különbözött alkoholos és józan állapotban, alkoholos állapotban gyakoribbak voltak ezek a típusú hibák.

Kutatásunkban azt is megvizsgáltuk, hogy a hangmagasság illetve az artikulációs tempó változik-e alkohol hatására. Több kutatásai szerint (pl. Behne és Rivera, 1990; Hollien, Dejong és mtsaik, 2001; Künzel, Braun, Eyshold, 1992; Trojan és Kryspin-Exner, 1968) alkohol hatására a hangmagasság növekszik. Más kutatások semmilyen vagy kétértelmű változást találtak (Johnson, Southwood és mtsaik, 1993; Sobell, Sobell és Coleman, 1982), megint mások (Schiel és Heinrich, 2009; Watanabe, Shin és mtsaik, 1994) a hangmagasság csökkenéséről számolnak be alkoholos állapotban. Mi kutatásunkban nem találtunk szignifikáns különbséget az alkoholos és a kontrol feltétel között a hangmagasságra vonatkozóan.

Viszonylag nehéz összehasonlítani a különböző kutatásokban mért véralkoholszintet, mivel a kutatók a rendelkezésre álló eszközöktől függően különböző módszereket választanak ennek meghatározására (vérvétellel mért véralkoholszint, a kilélegzett levegőben mért koncentráció, az elfogyasztott alkoholmennyiség és a testtömeg hányadosa), és ezeknek egymásba való átszámításának módja mai napig nem teljesen egyértelmű (ld. erről Chin és Pisoni, 1997; Haffner, Graw és mtsaik, 2003; Jones, 2010; Widmark, 1981). Ha csak a véralkoholszint és a testtömeget felhasználó kutatásokat vesszük figyelembe, azt mondhatjuk, hogy azokban a kutatásokban, amelyek kimutatták az alkoholnak a hangmagasságra gyakorolt hatását, a kísérleti személyek véralkoholszintje 0,1 és 0,2% között volt. Míg azoknál a kutatásoknál, ahol nem találtak ilyen változást, a személyek véralkoholszintje 0,079 és 0,13% között mozgott. Jelen kutatás vizsgálati személyeinek véralkoholszintje (Chin és Pisoni, 1997 konvertálási mutatója alapján számítva) kb. 0,08% volt. Ez az érték, és a hangmagasságbeli változás elmaradása tehát nem mond ellent korábbi kutatásoknak.

Künzel, Braun és Eyshold (1992) az elfogyasztott alkohol mennyiségétől teszi függővé, hogy emelkedik-e a hangmagasság, így elképzelhető, hogy a mi kísérleti személyeink esetében is bekövetkezett volna a hangmagasság-emelkedés, ha több alkoholt fogyasztanak. Valószínűsíthető tehát, hogy akárcsak az emlékezeti teljesítmény esetében, itt is van egy küszöbérték a véralkoholszintben, ami felett bekövetkezik a hangmagasság változása, míg ez alatt az érték alatt nem.

Számos korábbi kutatás (Künzel, Braun és Eyshold, 1992; Behne, Rivera és Pisoni, 1991; Hollien, Dejong és mtsaik, 2001; Pisoni, Hathaway, Yuchtman, 1985; Pisoni és Martin, 1989;

Sobell, Sobell és Coleman, 1982) azt találta, hogy alkoholos állapotban a beszédtempó lassul. Ugyanakkor ezek a kutatások általában nem részletezik, hogy ez a változás a lassabb artikulációból vagy a több szünettartásból fakad-e. Habár saját magunk azt tapasztaltuk, hogy az alkoholos felvétel sokkal lassabban zajlott, a kísérleti személyek lassabban beszéltek, az adatok ezt a megfigyelést nem támasztották alá, nem volt különbség a beszélők artikulációs tempója között alkoholos és józan állapotban. Egyedül a szintaktikai nyelvtörőknél figyelhettük meg, hogy az alkohol valóban befolyással van az artikulációs tempóra: az emberek lassabban mondták vissza a szintaktikai feladatokat a kontroll feltételhez képest. Ez a változás a véralkoholszinttel volt kapcsolatban: minél magasabb volt egy személy véralkoholszintje, annál lassabban artikulált.

Ittas állapotban a beszéd lassulásának szubjektív érzetét a szünetek megnövekedett száma okozhatta. Kutatásunk ugyanis alátámasztotta, hogy az emberek alkohol hatására beszédükben több szünetet tartanak. Különösen igaz volt ez a fonetikai, morfológiai és szintaktikai szintekre. Ebben a három típusban vétettek egyébként a személyek a legtöbb hibát. Valószínű, hogy a szünettartás (és szintaktikai nyelvtörők esetében az artikuláció lelassítása) egyfajta kompenzáció lehet: a személyek ezzel próbálják meg az alkohol munkamemóriára és beszédprodukcióra gyakorolt negatív hatásait ellensúlyozni.

Eredményünk tehát részben megegyezik a korábbi eredményekkel, hiszen a személyek beszédtempója lassult alkohol hatására. Ugyanakkor ki is egészíti azokat azzal, hogy ez nem az artikuláció gyorsulásából, hanem a hosszabb szünettartásokból fakad.

Ki kell még térnünk a kísérleti és a kontrollfeltétel sorrendjének esetleges hatására. Kutatásunkban a vizsgálati személyek előbb az alkoholos feltételben vettek részt, majd két hét múlva józan állapotban is elvégezték a feladatokat. Azt feltételeztük, hogy két hét elegendő idő arra, hogy a személyek elfelejtsék a nyelvtörőket, és hogy az ismerősség nem fogja befolyásolni teljesítményüket. Ugyanakkor jelen kísérleti elrendezésben nem zárhatjuk ki egyértelműen az idői hatást. Ha az időnek volt hatása, akkor az facilitáló hatás volt: az ismerősség és a nyelvtörőkre való emlékezés növelhette a visszamondás pontosságát, csökkenthette a nyelvtörők visszamondásakor ejtett hibák számát (az emlékezeti teljesítményből fakadókét), és gyorsabb visszamondást (gyorsabb artikuláció és kevesebb szünetek) eredményezhetett a második alkalommal. Ugyanakkor az időnek nincs hatása az artikulációs pontatlanságból fakadó hibákra, a fonológiai hurok teljesítményét mérő munkamemóriateszten mért teljesítményre, valamint a hangmagasságra. Annak ellenőrzésére, hogy hatással van-e az idő a második alkalombeli teljesítményre, érdemes lett volna egy olyan

kontrollcsoporttal is dolgozni, amelyik mind az első, mind a második alkalommal alkohol fogyasztása nélkül végzi el a feladatokat. Illetve a hatás kiszűrésére olyan elrendezést lehetett volna még használni, amelyben a személyek egyik fele az alkoholos és a józan feltételekben fordított sorrendben vesz részt.

A kapott eredmények a számterjedelemteszt, az artikulációs-motoros hibák és a hangmagasság esetében egyértelműen az alkohol hatását tükrözik, míg az emlékezeti teljesítményből fakadó hibák, az artikulációs tempó és a szünetek száma esetében az alkohol és esetlegesen az idő nem ismert hatásának kombinációjával kell számolnunk.

Összegzés

Ebben a tanulmányban egy magyar kísérlet eredményeiről számoltunk be, amely az alkohol beszédre és munkamemóriára gyakorolt hatását vizsgálja. Eredményeink nagyrészt összeegyeztethetők korábbi angol illetve német nyelven készült kutatások eredményeivel. Az emlékezeti teljesítménnyel kapcsolatban fontos megállapításunk, hogy az alkohol elsősorban a végrehajtó funkciót is igénylő műveletekre hat, hiszen a számterjedelemteszten mutatott teljesítményre nem volt hatással, ugyanakkor bonyolultabb szószerkezetek visszamondására igen. Ugyanakkor az eredmények még pontosabb értelmezéséhez érdemes lett volna bevonni egy placebo feltételt, ahol a vizsgálati személyek második alkalommal is alkohol nélkül végzik el a feladatokat.

Bár a nyelvtörők jó eszköz arra, hogy az artikulációs pontosságot és a munkamemóriát egyszerre ellenőrizzük velük, gyakorlati, az ittasság ellenőrzésének megállapítására való alkalmazásuk megfontolandó, mivel a személyek általában józan állapotban is hibáznak bennük. Így az egy személyen belüli változás kimutatására alkalmas lehet, de a beszélőfüggetlen tesztként valószínűleg nem alkalmazható. Az alkohol és a beszéd kapcsolatának vizsgálatakor számos korláttal találkozunk: az alkohol mennyiségétől, a személy súlyától, testalkatától, a korábbal elfogyasztott ételtől függően a hatás más és más lehet. A személyek beszéde közötti akusztikai-fonetikai jellemzők, valamint a beszédkésztség tekintetében is nagy különbségeket találunk. Az alkoholos befolyásoltság biztos megállapításához felhasználható „beszédteszt” elkészítéséhez még további kutatásokra van szükség.

IRODALOMJEGYZÉK

- Babor, T. F., Biddle-Higgins, J. C., Saunders, J. B., Monteiro, M. G. (2001) *AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care*. World Health Organization, Geneva, Switzerland
- Baddeley, A. (2003) Working memory: looking back and looking forward. *Neuroscience*, 4, 10, 829–839.
- Baum-Baicker, C. (1985) The psychological benefits of moderate alcohol consumption: A review of the literature. *Drug and Alcohol Dependence*, 15, 4, 305–322.
- Behne, D. M., & Rivera, S. M. (1990). Effects of alcohol on speech: Acoustic analysis of spondees. *Research on Speech Perception*, 16, 263–291.
- Behne, D. M., Rivera, S. M., Pisoni, D. B. (1991) Effects of alcohol on speech. *Journal of the Acoustical Society of America*, 90, 4, 2311-2311.
- Boersma, P., Weenink, D. (2009) Praat: doing phonetics by computer (Version 5.1.05.). Computer program.
- Chin, S. B., Pisoni, D.B. (1997) *Alcohol and Speech*. Academic Press, San Diego
- Christenfeld, N., Creager, B. (1996) Anxiety, alcohol, aphasia, and ums. *Journal of Personality and Social Psychology*, 70, 3, 451-460.
- Eich, J. (1980) The cue-dependent nature of state-dependent retrieval. *Memory and Cognition*, 8, 157–173.
- Forney, R. B., Hughes, F. W. (1961) Delayed auditory feedback and ethanol: Effect on verbal and arithmetical performance. *Journal of Psychology*, 52, 185-192.
- Gathercole S. E., McCarthy R. A. (eds.) (1994) *Memory Tests and Techniques*. Lawrence Erlbaum Associates, Hove, UK (A Special Issue of Memory)
- Goodwin, D. W., Powell, B., Bremer, D., Hoine, H., Stern, J. (1969) Alcohol and recall: State dependent effects in man. *Science*, 163, 1358–1360.
- Haffner, H. T., Graw, M., Dettling, A., Schmitt, G., & Schuff, A. (2003) Concentration dependency of the BAC/BrAC (blood alcohol concentration/breath alcohol concentration) conversion factor during the linear elimination phase. *International Journal of Legal Medicine*, 117, 5, 276-281.

- Hollien, H., Dejong, G., Martin, C. A., Schwartz, R., Liljegren, K. (2001) Effects of ethanol intoxication on speech suprasegmentals. *Journal of the Acoustical Society of America*, 110, 3198-206.
- Johnson, K., Pisoni, D. B., és Bernacki, R. H. (1990) Do voice recordings reveal whether a person is intoxicated? A case study. *Phonetica*, 47, 215-237.
- Johnson, K., Pisoni, D. B., és Bernacki, R. H. (1989) Report to the NTSB: Analysis of speech produced by the captain of the Exxon Valdez. *Research on Speech Perception*, 15, 1-9.
- Johnson, K., Southwood, M. H., Schmidt, A. M., Mouli, C. M., Holmes, A. T., Armstrong, A. A., . . . Wilson, A. S. (1993). *A physiological study of the effects of alcohol on speech and voice*. Paper presented at the 22nd annual Symposium on the Care of the Professional Voice at the Voice Foundation.
- Jones, A. W. (2010) The Relationship between Blood Alcohol Concentration (BAC) and Breath Alcohol Concentration (BrAC): A Review of the Evidence. *Road Safety Web Publication*, 15. Elektronikus dokumentum: <http://www.dft.gov.uk/pgr/roadsafety/research/rsrr/theme3/report15.pdf>
- Jones, B. M. (1973) Memory impairment on the ascending and descending limbs of the blood alcohol curve. *Journal of Abnormal Psychology*, 82, 1, 24-32.
- Jones, B. M., Jones, M. K. (1977) Alcohol and memory impairment in male and female social drinkers. In Birnbaum, I. M. and Parker, E. S. (eds.) *Alcohol and human memory*. 127–138. Erlbaum, Hillsdale, N.J.
- Jones, M. K., Jones, B. M. (1980) The Relationship of Age and Drinking Habits to the Effects of Alcohol on Memory in Women. *Journal of Studies on Alcohol*, 41, 1, 179–186.
- Kalapos, M. P. (2007) *Bevezetés az alkoholológiába*. Orvostanhallgatóknak és szakorvosjelölteknek. Medicina, Budapest.
- Karch, S. B. (2008) *Forensic issues in alcohol testing*. CRC Press, U.S.A.
- Künzel, H. J., Braun, A., Eyshold, U. (1992) *Einfluss von Alkohol auf Sprache und Stimme*. Kriminalistik Verlag, Heidelberg
- Levelt, W. J. M. (1989) *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

- Maylor, E. A., Rabbitt P. M. A (1987) Effect of alcohol on rate of forgetting. *Psychopharmacology*, 91, 2, 230–235.
- Miller, L. L., Dolan, M. P. (1974) Effects of alcohol on short term memory as measured by a guessing technique. *Psychopharmacology*, 35, 4, 353–364.
- Németh D., Pléh Cs. (2001) Nyelvfeldolgozás, munkamemória, fordítás. *Fordítástudomány*, 3, 1, 40–53.
- Nordby, K., Watten, R. G., Raanaas, R. K., Magnussen, S. (1999) Effects of Moderate Doses of Alcohol on Immediate Recall of Numbers: Some Implications for Information Technology. *Journal of Studies on Alcohol*, 60, 6, 873–878.
- Nuotto, E., Mattila, M. J. (1990) *Actions and interactions of alcohol* on performance tests. In S. Parvez & H. Parvez (Series Eds.) & H. Ollat, *Alcohol and Behavior*, 117–130. VSP Press, Utrecht
- Nyberg, L., Tulving, E. (1996) Classifying human long-term memory: evidence from converging dissociations. *European Journal of Cognitive Psychology*, 8, 163–183.
- Pisoni D. B., Martin C. S. (1989) Effects of alcohol on the acoustic-phonetic properties of speech: perceptual and acoustic analyses. *Alcoholism: Clinical Experimental Research*, 13, 4, 577–587.
- Pisoni, D. B., Hathaway, S. N., Yuchtman, M. (1985) *Effects of alcohol on the acoustic-phonetic properties of speech: Final report to GM Research Laboratories*. (SRL Technical Note No. 85-03). Bloomington: Indiana University.
- Pisoni, D. B., Johnson, K., Bernacki, R. H. (1991) Effects of alcohol on speech. In *Proceedings of the Human Factors Society* (pp. 694-698). Santa Monica, CA: Human Factors Society.
- Racsmány, M., Lukács, Á., Németh, D., Pléh, Cs. (2005) A verbális munkamemória magyar nyelvű vizsgálóeljárásai. *Magyar Pszichológiai Szemle*, 60, 4, 479–505.
- Santamaria, J. N. (1989). The effects of ethanol and its metabolism on the brain. In K. E. Crow & R. D. Batt (Eds.), *Human metabolism of alcohol: Vol. Ill. Metabolic and physiological effects of alcohol* (pp. 35-47). Boca Raton, FL: CRC Press.
- Schiel, F., Heinrich, C. (2009) Laying the Foundation for In-car Alcohol Detection by Speech. *Proceedings of the INTERSPEECH 2009*, Brighton, UK, pp. 983–986.

- Sobell, L. C., Sobell, M. B., & Coleman, R. F. (1982) Alcohol-induced disfluency in nonalcoholics. *Folia Phoniatrica*, 34, 316–323.
- Tringer, L. (2005) *A pszichiátria tankönyve*. Semmelweis Kiadó, Budapest.
- Trojan, F., Kryspin-Exner, K. (1968) The decay of articulation under the influence of alcohol and paraldehyde. *Folia Phoniatrica (Basel)*, 20, 4, 217-238.
- Tyson, P. D., Schirmuly, M. (1994) Memory enhancement after drinking ethanol: consolidation, interference, or response bias? *Psychology & Behavior*, 56, 5, 933–937.
- Watanabe, H., Shin, T., Matsuo, H., Okuno, F., Tsuji, T., Matsuoka, M., . . . Matsunaga, H. (1994) Studies on vocal fold injection and changes in pitch associated with alcohol intake. *Journal of Voice*, 8, 4, 340-346.
- Widmark, E. M. P. (1981) *Principles and Applications of Medicolegal Alcohol Determination*. Davis, CA: Biomedical Publications.
- Zaimov, K. (1969) Die Sprachstörungen als Kriterium der Bewusstseinsstörungen. *Psychiatrie, Neurologie, und Medizinische Psychologie*, 21, 218–225.

Diszkusszió

A fentiekben olyan kutatásokat ismertettünk, amelyek mindegyike azt vizsgálta, hogy különböző belső állapotok hogyan nyilvánulhatnak meg a beszédben, illetve hogy a hallgatók bizonyos beszédparaméterek együttes megjelenéséből milyen belső állapotokra következtetnek, hogyan ismerik fel ezeket, melyek a következtetésben leginkább szerepet játszó beszédparaméterek.

Az elméleti bevezetésben részletesen írtunk olyan nyelvészeti és pszichológiai irányzatokról, amelyek foglalkoznak az érzelmeknek a beszédbeli megjelenésével. A nyelvészeti társalgáselemzés a dialógus szintjén, a tartalomelemzés a szavak, mondatok jelentését vizsgálva, a vokálisérzelem-kutatások pedig a beszédhang sajátosságait figyelembe véve teszik ezt. Tudomásunk szerint nem csak magyar nyelven, de angolul sem áll rendelkezésünkre olyan tanulmány, amely e három irányzat módszereit és kutatásait együtt ismertetné.

A beszédpszünetek és a beszédbeli egyéb megakadások vizsgálata, ami a mi kutatásaink középpontjában is állt, részben a társalgáselemzés, részben a vokális érzelmkifejezés kutatásai körébe tartozik. A beszédpszünetek érzelmetulajdonításban játszott szerepével kapcsolatos kísérleteink rámutattak arra, hogy a pszünetek fontos szerepet játszanak abban, hogy milyen érzelmeket tulajdonítunk a beszélőnek. Ugyanazokat a beszédeket szomorúbbnak és ijedtebbnek ítéljük, ha hosszabbak a pszünetek. Mivel a mindennapi szituációkban gyakori, hogy egy-egy beszélő hosszabb ideig beszél megszakítás nélkül, ezért mindenképpen indokolt az ilyen, több mondaton átívelő beszédrészek vizsgálata. Egy korábbi kutatásunk (Szabó, 2008) kimutatta, hogy szomorú érzelmi állapotban hosszabb pszüneteket tartva, nagyobb pszünetarányal beszélnek a személyek, mint vidám érzelmi állapotban. A mostani kutatásunk pedig arra mutat rá, hogy a pszünetek proximálisan is jelzésként szolgálnak a beszélő állapotára vonatkozóan.

Többféle megközelítés létezik arra nézve, hogy mi is az oka és a funkciója a beszédbeli megakadásoknak, milyen változók befolyásolják a megakadások arányát. Goldman-Eisler (1958, 1968) valamint Rochester (1973) összefoglalója szerint a megakadások arányának növekedését a kognitív feldolgozási igény megnövekedése is okozhatja. Így egy nehezebben hozzáférhető szó vagy egy szintaktikailag nehezebb szerkezet előtt, illetve ha több erőfeszítést

igényel a verbális feladat, több a szünet. Mahlra vezethető vissza az az elképzelés, hogy a megakadások a beszélő érzelmi állapotát tükrözik (Mahl, 1956): minél jobban szorong valaki, annál több megakadás lesz a beszédében; Hofman és munkatársai (1997) eredményei alapján a főbiás személyek előadásában több a kitöltött szünet, hosszabbak és gyakoribbak a néma szünetek, mint a kontrollcsoport személyei beszédében. Christenfeld és Creager (1996) az önmonitorozáshoz mutatójának tartják megakadások arányának megváltozását. Kutatásuk szerint a kitöltött szünetek (és csak ezek) nem a szorongás, hanem az önmonitorozás mutatói. Ha a személyek kevésbé öntudatosak, csökken az önmonitorozás – például alkohol hatása alatt –, és ezért folyamatosabb lesz beszédük, kevesebb kitöltött szünet lesz beszédükben. A legtöbb kutatás (ld. Rochester, S. R. (1973), Mahl (1987), Christenfeld és Creager (1996) kimutatta, hogy a kitöltött szünetek (hezitációk) másképp viselkednek, mint a többi megakadásféle, így pl. a néma szünet, tehát elképzelhető, hogy míg az egyik aránya nő, a másiké csökken egy adott érzelmi állapotban.

Látjuk, hogy a szünetek aránya, a többi megakadással együtt, függ belső állapotainktól, és az is bebizonyosodott, hogy a szünetek proximális jelzésként információt szolgáltatnak ezekről az állapotokról a hallgató számára. Mivel továbbra is csak kevés olyan kutatás áll a rendelkezésünkre, amely 1) természetes helyzetben előforduló mondatokat, 2) hosszabb beszédrészleteket, 3) szüneteket és megakadásokat vizsgál érzelmi vagy más belső állapot függvényében, ezért további vizsgálatok szükségesek ezeken a területeken. Arról, hogy az érzelmek pontosan hogyan befolyásolják a beszédhez szükséges kognitív folyamatokat, nincsenek pontos ismereteink.

A tanulmányok megírása és benyújtása óta több olyan beszédtechnológiai cikk is megjelent, amely magyar nyelvű beszédmintákon vagy magyar nyelvű megítélők segítségével vizsgálja a gépi és humán érzelmfelismerés lehetőségeit (Vicsi 2010, Vicsi et al., 2010, Tóth et al., 2008, Riviello et al., 2012). Ezek nagy és nemzetközi szinten is kevés előzményt tartalmazó újítása, hogy érzelmeket tartalmazó spontánbeszéd adatbázisokat hoztak létre, és ezek segítségével vizsgálják, hogy melyek az egyes érzelmekre jellemző prozódiai mintázatok, valamint hogy milyen jegyek segítenek a gépi érzelmfelismerés során. A korábbi szó- vagy egyszerű mondatlappú egységek helyett pedig a frázisokat (intonációs egységeket) veszik alapul (Vicsi és Sztahó, 2012). Mivel tudjuk, hogy az érzelem és a beszéd kapcsolatának vizsgálatakor az univerzális jegyeken túl nyelvspecifikus meghatározókkal is számolnunk kell, ezért a hazai kutatások segíthetnek a hasonlóságok és a különbségek feltérképezésében.

A szüneteknek az érzelemtulajdonításban játszott szerepének vizsgálata mellett azt kutattuk, hogy az alkohol hogyan hat a beszédre és a rövid távú emlékezetre. Ismert, hogy az alkohol hat érzelmi állapotunkra, a bennünk végbemenő kognitív és motoros folyamatokra, ezért feltételeztük, hogy nyelvtörők és számsorozatok visszamondásakor is megmutatkozik ez a hatás. Az egy személyen belüli változásokat elemeztük alkoholos és józan állapotban. A nyelvtörők visszamondásakor azt találtuk, hogy a vizsgálati személyek több hibával beszélnek és több szünetet tartanak, mint józan állapotban. A hibázások megnövekedését az alkohol motoros-artikulációs folyamatokra gyakorolt hatása okozhatja: az izmok lazulása és a pontatlan artikuláció beszédhangok összekeveréséhez vezet. A szünetek és az artikulációs tempó különválasztott választott elemzése azt mutatta, hogy a korábbi kutatásokkal megegyezően a beszédtempó csökkent alkoholos befolyásoltság alatt, ugyanakkor ezt a lassulást nem az artikulációs sebesség lassulása, hanem a megnövekedett szünethossz okozza. A hangmagassággal kapcsolatos eredményeink szerint a hangmagasság nem változik 0,08% alkoholszint mellett, ugyanakkor korábbi kutatások alapján feltételezzük, hogy magasabb alkoholszint felett a hangmagasság nő.

A hibázások típusát részletesen is megvizsgáltuk, és azt találtuk, hogy a hibák körülbelül fele az artikuláció pontatlanságából, másik fele pedig az emlékezeti teljesítmény csökkenéséből ered. Az a tény, hogy alkohol hatására több ilyen utóbbi, a rövid távú emlékezeti teljesítményhez kapcsolódó hibát követtek el a kísérleti személyek, ugyanakkor a szintén rövid távú memóriát vizsgáló számterjedelemteszten nem volt ilyen különbség a két feltétel között, arra utal, hogy a két feladat némileg más kapacitást igényel, másfajta folyamatokat foglal magában. A Baddeley-féle munkamemória modellben (2003) értelmezve: míg a számterjedelemteszt a fonológiai hurok kapacitását, addig a bonyolult mondatok visszamondása a végrehajtó funkciókat is terheli, így lehetséges, hogy mérsékelt mennyiségű alkohol hatása csak a végrehajtó funkciót is igénylő bonyolultabb feladatoknál jelentkezik, és az egyszerű ismételtetésre nincs hatással. A folyamat pontos megismeréséhez további kutatásokra van szükség.

Alkalmazhatóság, konklúzió

Az ismertetett kutatások közelebb visznek bennünket ahhoz, hogy teljes képet kapjunk a belső állapotok, az érzelmek, a személyiség beszédbeli megmutatkozásáról. Ezekkel az

ismeretekkel egyrészt tudatosíthatjuk például pszichológusokban, orvosokban, politikaibeszéd-elemzőkben, hogy melyek azok a változások, amelyek figyelmet érdemelnek. Másrészt az egyes állapotokhoz tartozó disztális és proximális jelzések vizsgálatával és megállapításával lehetővé válik például érzelmes beszéd előállítása beszédszintetizátorral (pl. Türk és Schröder, 2008), vagy az érzelmi állapot (pl. Vicsi és Sztahó, 2011) és az alkoholos befolyásoltság (pl. Schiel és Heinrich, 2009) automatikus detektálása számítógépes alkalmazások segítségével.

Ahogy a bevezetőben is említettük, a jövőbeli kutatások egyik célja minden valószínűség szerint az lesz, hogy a ma még egy-egy irányzaton belüli kutatásokat összekapcsolja, és olyan modelleket majd programokat alkosson, amelyek egyszerre tudják feldolgozni a tartalom, a hang és a társalgási struktúra jellegzetességeit.

Összesített irodalomjegyzék

- Albas, D.C., Mccluskey, K.W., Albas, C.A (1976) Perception of Emotional Content of Speech - Comparison of 2 Canadian Groups. *Journal of Cross-Cultural Psychology*, 7(4): p. 481-490.
- Apple, W., Streeter, L. A., Krauss, R. M. (1979) Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, 37, 715-727.
- Aubergé, V. és Lemaître, L. (2000) The prosody of smile. *SpeechEmotion-2000*, 122–126.
- Audibert, N., Aubergé, V. & Rilliard, A. (2005) The prosodic dimensions of emotion in speech: the relative weights of parameters. In: *Interspeech'2005 – Eurospeech*. Lisszabon, Portugália, 525–528.
- Babor, T. F., Biddle-Higgins, J. C., Saunders, J. B., Monteiro, M. G. (2001) *AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care*. World Health Organization, Geneva, Switzerland
- Baddeley, A. (2003) Working memory: looking back and looking forward. *Neuroscience*, 4, 10, 829–839.
- Banos, R. M., Botella, C., Liaño, V., Rey, B., Guerrero, B., Alcaniz, M. (2003) Virtual reality as mood induction procedure. Presented at PRESENCE 2003, *6th Annual International Workshop on Presence*.
- Banse, R. és Scherer, K. R. (1996) Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology* 70 (3), 614–636.
- Bänzinger, T. és Scherer, K. R. (2005) The role of intonation in emotional expressions. *Speech Communication* 46, 252–267.
- Battacchi, M. W., Suslow, T., Renna, M. (1996) *Emotion und Sprache. Zur Definition der Emotion und ihren Beziehungen zu kognitiven Prozessen, dem Gedächtnis und der Sprache*. Frankfurt am Main: Lang.
- Baum-Baicker, C. (1985) The psychological benefits of moderate alcohol consumption: A review of the literature. *Drug and Alcohol Dependence*, 15, 4, 305–322.
- Beaupre, M.G., Hess, U. (2005) Cross-cultural emotion recognition among Canadian ethnic groups. *Journal of Cross-Cultural Psychology*, 36(3), 355-370.

- Behne, D. M., Rivera, S. M. (1990). Effects of alcohol on speech: Acoustic analysis of spondees. *Research on Speech Perception*, 16, 263–291.
- Behne, D. M., Rivera, S. M., Pisoni, D. B. (1991). Effects of alcohol on speech: Durations of isolated words, sentences, and passages in fluent speech. *Journal of the Acoustical Society of America*, 90, 2311.
- Beier, E.G., Zautra, A.J. (1972) Identification of vocal communication of emotions across cultures. *J Consult Clin Psychol*, 39(1): p. 166.
- Bergmann, G., Goldbeck, T., Scherer, K.R. (1988) [Emotional impression of prosodic speech markers]. *Z Exp Angew Psychol*,. 35(2): p. 167-200.
- Biehl, M., Matsumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T., Ton, V. (1997) Matsumoto and Ekman's Japanese and Caucasian Facial Expressions of Emotion (JACFEE): Reliability data and cross-national differences. *Journal of Nonverbal Behavior*, 21(1), 3-21.
- Boersma, P., Weenick, D. (2009) Praat: doing phonetics by computer (Version 5.1.05.) Computer program, Insitute of Phonetic Sciences, University of Amsterdam, Hollandia. www.praat.org Letöltés ideje: 2009.05.01.
- Böhm, T. (2009) *Analysis and modeling of speech produced with irregular phonation*, PhD disszertáció. BME Villamosmérnöki és Informatikai Kar, Budapest.
- Bolla, K. (1992) Érzelmek kifejezése a megnyilatkozás szuprasegmentális hangszerkezetében. Szuprasegmentális elemzések. *Egyetemi Fonetikai Füzetek* 7. 90–116.
- Bosch, L. T. (2003) Emotions, speech and the ASR framework. *Speech Communication*, 40, 1, 213-225.
- Breitenstein, C., D. Van Lancker, I. Daum (2001) The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Cognition & Emotion*, 15, 1, 57-79.
- Breitenstein, C., et al. (1996) Erfassung der Emotionswahrnehmung bei zentralnervösen Läsionen und Erkrankungen: Psychometrische Gütekriterien der "Tübinger Affekt Batterie". *Neurologie & Rehabilitation*,. 2(93-101).

- Brewer, D., Doughtie E. B., Lubin, B. (1980) Induction of mood and mood shift. *Journal of Clinical Psychology*, 36, 215-226.
- Bucci, W. és Freedman, N. (1981) The language of depression. *Bulletin of Menninger Clinic*, 45, 334–58.
- Burkhardt, F. and W.F. Sendlmeier (2000) Verification of Acoustical Correlates of Emotional Speech using Formant-Synthesis, in *ISCA Workshop (ITRW) on Speech and Emotion2000*: Belfast.
- Burkhardt, F., Audibert, N., Malatesta, L., Turk, O., Arslan, L., Auberge, V. (2006) Emotional prosody - does culture make a difference? *Proceedings of Speech Prosody2006*: Dresden, Germany.
- Cahn, J. E. (1990) The Generation of Affect in Synthesized Speech. *Journal of the American Voice I/O Society*, 1-19.
- Campbell, R. S. és Pennebaker, J. W. (2002) The secret life of pronouns: flexibility in writing style and physical health. *Psychological Science*, 14, 60–65.
- Carlson, R., (1992) Synthesis - Modeling Variability and Constraints. *Speech Communication*, 11(2-3): p. 159-166.
- Chin, S. B., Pisoni, D. B. (1997). *Alcohol and Speech*. San Diego: Academic Press.
- Christenfeld, N., & Creager, B. (1996). Anxiety, alcohol, aphasia, and ums. *Journal of Personality and Social Psychology*, 70(3), 451-460.
- Cohen, A. S., Hong, S. L., Guevara, A. (2010) Understanding emotional expression using prosodic analysis of natural speech: Refining the methodology. *Journal of Behavior Therapy and Experimental Psychiatry*, 41, 150–157.
- Collins, P. J. (1980). A Comparison of the Oral Syntactic Performance of Alcoholic and Non-Alcoholic Adults. *Language and Speech*, 23, 281-288.
- Cooney, O. M., McGuigan, K. G., Murphy, P. J. P., Conroy, R. M. (1998). Acoustic analysis of the effects of alcohol on the human voice. *Journal of the Acoustical Society of America*, 103, 2895.
- Cowie, R. R. és Cornelius, R. (2003) Describing the emotional states that are expressed in speech. *Speech Communication*, 40, 5–32.

- Crystal, T. H., & House, A. S. (1990). Articulation Rate and the Duration of Syllables and Stress Groups in Connected Speech. *Journal of the Acoustical Society of America*, 88(1), 101-112.
- Deppermann, A. és Lucius-Hoene, G. (2005) Trauma erzählen – kommunikative, sprachliche und stimmliche Verfahren der Darstellung traumatischer Erlebnisse. *Psychotherapie und Spezialwissenschaft. Zeitschrift für Qualitative Forschung und klinische Praxis*. 1. 35–73.
- Douglas-Cowie, E. és Cowie, R. (2002). *The vocal correlates of stress: a summary*. HTML dokumentum, URL: http://www.image.ece.ntua.gr/oresteia/private/contributed_docs/reports/Voice-and-Stress.pdf Letöltés ideje: 2008. március 13.
- Douglas-Cowie, E., Campbell, N., Cowie, R., Roach, P. (2003) Emotional speech: Towards a new generation of databases. *Speech Communication*, 40, 33–60.
- Drahota, A., Costall, A., Reddy, V. (2008) The vocal communication of different kinds of smile. *Speech Communication*, 50, 278–287.
- Dyck, M. J., Farrugia, Ch., Shochet, I. M., Holmes-Brown, M. (2004) Emotion recognition/understanding ability in hearing or vision-impaired children: do sounds, sights, or words make the difference? *Journal of Child Psychology and Psychiatry*, 45, 4, 789–800.
- Ehmann, B. (2002) *A szöveg mélyén*. Budapest, Új Mandátum Könyvkiadó.
- Eibl-Eibesfeldt, I. (1973) The expressive behavior of the deaf-and-blind-born. In Von Cranach M. & Vine I. (eds.) *Social communication and movement* (163–194). New York, Academic Press.
- Eich, J. (1980) The cue-dependent nature of state-dependent retrieval. *Memory and Cognition*, 8, 157–173.
- Ekman, A., J. Ström, 1969 [Habit forming drugs (7): somatic complications]. *Lakartidningen*, 66(48), 5021-4.
- Ekman, P. (1999) Basic emotions. In T. Dalgleish & M. Power, *Handbook of cognition and emotion*. Pp 45-60. Sussex, UK, John Wiley & Sons.

- Ekman, P., Friesen, W.V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., Lecompte, W.A., Pitcairn, T., Ricci-Bitti, P.E., Scherer, K.R., Tomita, M., Tzavaras A. (1987) Universals and cultural differences in the judgments of facial expressions of emotion. *J Pers Soc Psychol*, 53(4): p. 712-7.
- Ekman, P., Friesen, W.V. (1971) Constants across cultures in the face and emotion. *J Pers Soc Psychol*, 17(2): p. 124-9.
- Ekman, P., Sorenson, E.R., Friesen, W.V. (1969) Pan-cultural elements in facial displays of emotion. *Science*, 164(3875): p. 86-8.
- Eldred, S.H., Price, D.B. (1958) A linguistic evaluation of feeling states in psychotherapy. *Psychiatry*, 21(2): p. 115-21.
- Elfenbein, H.A., Ambady, N. (2003) Cultural similarity's consequences - A distance perspective on cross-cultural differences in emotion recognition. *Journal of Cross-Cultural Psychology*, 34(1): p. 92-110.
- Elfenbein, H.A., Beaupre, M., Levesque, M., Hess, U. (2007) Toward a dialect theory: cultural differences in the expression and recognition of posed facial expressions. *Emotion*, 7(1), 131-46.
- Fairbanks, G., Hoaglin, L.W. (1941) An experimental study of the durational characteristics of the voice during the expression of emotion. *Speech Monographs*, 8(1), 85-90.
- Fék M., Olaszy, G., Németh, G., Gordos, G. (2005). Érzelem kifejezése gépi beszéddel. *Beszéd kutatás 2005*. MTA Nyelvtudományi Intézet, 134–144.
- Fiehler, R. (1990) *Kommunikation und Emotion. Theoretische und empirische Untersuchungen zur Rolle von Emotionen in der verbalen Interaktion*. Walter de Gruyter, Berlin.
- Fiehler, R. (2001) Emotionalität im Gespräch. In Klaus, B., Gerd, A., Wolfgang, H., Sven, S. F. (Hrsg) *Text- und Gesprächslingistik. Ein internationales Handbuch zeitgenössischer Forschung*. Halbband 2. Walter de Gruyter, Berlin, New York, 1425–1438.
- Földi, É. (1996) Az érzelemkifejezés szupraszegmentális formái és percepciója. *Egyetemi Fonetikai Füzetek* 20.

- Fónagy I. (1981) Emotions, voice and music. In: J. Sundberg (ed.) *Research aspects on singing. Papers given at a seminar organized by the Committee for the Acoustics of Music*, 51-79. Stockholm: The Royal Swedish Academy of Music.
- Fónagy I. (2001) *Languages within language: an evolutive approach*. Amsterdam; Philadelphia: John Benjamins.
- Fónagy, I. (1998) *A költői nyelvről*. Corvina, Budapest.
- Fónagy, I., Magdics, K. (1963) Emotional patterns in intonation and music. *Zeitschrift für Phonetik*, 16, 293–326.
- Fónagy, I., Magdics, K. (1967) A magyar beszéd dallama. Akadémiai Kiadó, Budapest.
- Fontaine, J., Scherer, K., Roesch, E., Ellsworth, P. (2007) *The world of emotions is not two-dimensional*. *Psychological Science*, 18 (12), 1050-1057.
- Forgács, J. (2003) *Az érzelmek pszichológiája*. Kairosz Könyvkiadó, Budapest.
- Forney, R. B., Hughes, F. W. (1961) Delayed auditory feedback and ethanol: Effect on verbal and arithmetical performance. *Journal of Psychology*, 52, 185-192.
- Frolov M. V., Milovanova G. B., Lazarev N. V., Mekhedova A. Y. (1999) Speech as an Indicator of the Mental Status of Operators and Depressed Patients. *Hum. Psychol.* 66, 398–412.
- Frye, N. (1998) *A kritika anatómiája*. Helikon, Budapest.
- Fülöp, É. (2008) Az érzelmek reprezentációja történelmi regényekben és történelemkönyvekben. *Megismerés, reprezentáció, értelmezés, 10 éves a PTE Pszichológiai Doktori Iskolája*, 99-114.
- Fussell, S. R., Ed. (2002). *The verbal communication of emotion: Interdisciplinary perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey* (Vol. 3, pp. 69–96). Cambridge: Cambridge University Press.
- Gathercole S. E., McCarthy R. A. (eds.) (1994) *Memory Tests and Techniques*. Lawrence Erlbaum Associates, Hove, UK (A Special Issue of Memory)

- Givens, B. S., & Breese, G. R. (1990). Electrophysiological Evidence That Ethanol Alters Function of Medial Septal Area without Affecting Lateral Septal Function. *Journal of Pharmacology and Experimental Therapeutics*, 253(1), 95-103.
- Gobl, C. és Ni Chasaide, A. (2003) The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, 40 (1-2), 189-212.
- Goldman-Eisler, F. (1958) Speech production and the predictability of words in context. *Quarterly Journal of Experimental Psychology*, 10, 96-106.
- Goldman-Eisler, F. (1968) *Psycholinguistics: experiments in spontaneous speech*. London: New York, Academic Press.
- Goodwin, D. W., Powell, B., Bremer, D., Hoine, H., Stern, J. (1969) Alcohol and recall: State dependent effects in man. *Science*, 163, 1358–1360.
- Gósy, M. (2003) A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar Nyelvőr*, 127, 257–277.
- Gósy, M. (2003). A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. [Co-occurrence and frequency of disfluencies in Hungarian spontaneous speech]. *Magyar Nyelvőr*, 127, 257–277.
- Gósy, M. (2004). *Fonetika, a beszéd tudománya*. Budapest: Osiris Kiadó.
- Gósy, M. (2008) Magyar spontán beszéd adatbázis - BEA. *Beszéd kutatás 2008*, 116–128.
- Gósy, M. and Kovács, M. (2008) Virtual sentences of spontaneous speech, in Gardner-Bonneau, D., Blanchard, H. E. (Eds.) *Human Factors and Voice Interactive Systems*, p. 193-207. Springer: New York ; London
- Gottschalk, L. A., Gleser, G. C. (1969) *The measurement of psychological states through the content analysis of verbal behavior*. Univ. Of California Press, Berkeley. MONITOR, 4. sz., 1972. OIEI, Kézirat.
- Gottschalk, L. A., Winget, C. N., Gleser, G. C. (1969) *Manual of Instructions for Using the Gottschalk-Gleser Content Analysis Scales: Anxiety, Hostility, Social Alienation- Personal Disorganization*. Los Angeles, Berkeley, University of California Press.
- Grandjean, D., Bänziger, T., Scherer, K. R. (2006) Intonation as an interface between language and affect. *Progress in Brain Research*, 156, 235–268.

- Gyarmathy, D. (2008). Hogyan hat az alkohol a spontán beszédre? . In Á. Kuna & Á. Veszelszki (Eds.), *Félúton 3 Konferenciakötet* (pp. 81-99).
- Haffner, H. T., Graw, M., Dettling, A., Schmitt, G., & Schuff, A. (2003) Concentration dependency of the BAC/BrAC (blood alcohol concentration/breath alcohol concentration) conversion factor during the linear elimination phase. *International Journal of Legal Medicine*, 117, 5, 276-281.
- Hart, R. P. (2001) Redeveloping DICTION: theoretical considerations. *See West*, pp. 43–60.
- Henderson, A.I., Goldman-Eisler, F., Skarbek, A. (1966) Sequential temporal patterns in spontaneous speech. *Language and Speech*, 9, 207-216.
- Hofmann, S.G., Gerlach, A.L., Wender, A., Roth, W.T. (1997) Speech disturbances and gaze behavior during public speaking in subtypes of social phobia. *J Anxiety Disord*, 11(6), 573-85.
- Hollien, H., & Martin, C. A. (1996) Conducting research on the effects of intoxication on speech. *Forensic Linguistics: The International Journal of Speech, Language and the Law*, 3, 107–127.
- Hollien, H., DeJong, G., Martin, C. A., Schwartz, R., & Liljegren, K. (2001) Effects of ethanol intoxication on speech suprasegmentals. *Journal of the Acoustical Society of America*, 110(6), 3198-3206.
- Hollien, H., Harnsberger, J. D., Martin, C. A., Hill, R., & Alderman, G. A. (2009) Perceiving the Effects of Ethanol Intoxication on Voice. *Journal of Voice*, 23(5), 552-559.
- Hollien, H., Liljegren, K., Martin, C. A., DeJong, G. (1999) Prediction of intoxication levels by speech analysis. In A. Braun (Ed.), *Advances in Phonetics* (pp. 40–50). Stuttgart: Steiner Verlag.
- Hollingworth, H. L. (1923) The influence of alcohol. *Journal of Abnormal Psychology and Social Psychology*, 18, 204-237.
- Hunt, W. A. (1985) *Alcohol and biological membranes*. New York: Guilford Press.
- Iványi, Zs. (2001) A nyelvészeti konverzációelemzés. *Magyar Nyelvőr* 125, 74–93.

- Izard, C. E. (1993) Organizational and motivational functions of discrete emotions. In M. Lewis & J. M. Haviland, *Handbook of emotions*, pp. 631-41. New York, Guilford Press.
- Izard, C.E. (1971) *The Face of Emotion.*, New York: Appleton-Century-Crofts.
- Izard, C.E. (1994) Innate and universal facial expressions: evidence from developmental and cross-cultural research. *Psychol Bull*, 115(2), 288-99.
- Johannes, B., Salnitski, P., Gunga, H-C., Kirsch, K. (2000) Voice stress monitoring in space – possibilities and limits. *Aviation, Space and Environmental Medicine*, 71, 9 (11), A58-A65,
- Johnson W. F., Emde, R. N., Scherer K. R., Klinnert M. D. (1986) Recognition of Emotion From Vocal Cues. *Archives from General Psychiatry*, 43(3), 280-283.
- Johnson, K., Pisoni, D. B., és Bernacki, R. H. (1990) Do voice recordings reveal whether a person is intoxicated? A case study. *Phonetica*, 47, 215-237.
- Johnson, K., Pisoni, D. B., és Bernacki, R. H. (1989) Report to the NTSB: Analysis of speech produced by the captain of the Exxon Valdez. *Research on Speech Perception*, 15, 1-9.
- Johnson, K., Southwood, M. H., Schmidt, A. M., Mouli, C. M., Holmes, A. T., Armstrong, A. A., . . . Wilson, A. S. (1993). *A physiological study of the effects of alcohol on speech and voice*. Paper presented at the 22nd annual Symposium on the Care of the Professional Voice at the Voice Foundation.
- Johnstone, T. és Scherer, K. R. (2000). Vocal communication of emotion. In Lewis, M. és Haviland-Jones, J. (Eds.), *Handbook of Emotions, Second Edition*. Guilford Press, New York, 220–235.
- Jones, A. W. (2010) The Relationship between Blood Alcohol Concentration (BAC) and Breath Alcohol Concentration (BrAC): A Review of the Evidence. *Road Safety Web Publication*, 15. Elektronikus dokumentum: <http://www.dft.gov.uk/pgr/roadsafety/research/rsrr/theme3/report15.pdf>
- Jones, B. M. (1973) Memory impairment on the ascending and descending limbs of the blood alcohol curve. *Journal of Abnormal Psychology*, 82, 1, 24-32.
- Jones, B. M., Jones, M. K. (1977) Alcohol and memory impairment in male and female social drinkers. In Birnbaum, I. M. and Parker, E. S. (eds.) *Alcohol and human memory*. 127–138. Erlbaum, Hillsdale, N.J.

- Jones, M. K., Jones, B. M. (1980) The Relationship of Age and Drinking Habits to the Effects of Alcohol on Memory in Women. *Journal of Studies on Alcohol*, 41, 1, 179–186.
- Jovicic, S.T., Kasic, Z., Dordevic, M., Rajkovic, M. (2004) Serbian Emotional Speech Database: Design, Processing and Evaluation, in *SPECOM-2004*, 77-81.
- Juslin, P. N., Laukka, P. (2001) Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion*, 1, 4, 381-412.
- Juslin, P. N., Scherer, K. R. (2005) Vocal expression of affect. In J. A. Harrigan, R. Rosenthal, K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 65-135). New York: Oxford University Press.
- Kalapos, M. P. (2007) *Bevezetés az alkoholológiába*. Orvostanhallgatóknak és szakorvosjelölteknek. Medicina, Budapest.
- Karch, S. B. (2008) *Forensic issues in alcohol testing*. CRC Press, U.S.A.
- Karczag, J. (1988) A Gottschalk-Gleser-féle tartalomelemzési módszer. In Mérei, F. és Szakács, F. (szerk.) *Pszichodiagnosztikai vademecum*. Budapest, Tankönyvkiadó, 273–353.
- Kasl, S.V., Mahl, G.F. (1965) The Relationship of Disturbances and Hesitations in Spontaneous Speech to Anxiety. *J Pers Soc Psychol*, 95, 425-33.
- Kienast, M., Sendlmeier, W. F. (2000) Acoustical analysis of spectral and temporal changes in emotional speech. *SpeechEmotion-2000*, 92-97.
- Klingholz, F., Penning, R., Liebhardt, E. (1988) Recognition of Low-Level Alcohol-Intoxication from Speech Signal. *Journal of the Acoustical Society of America*, 84(3), 929-935.
- Künzel, H. J., Braun, A., Eyshold, U. (1992) *Einfluss von Alkohol auf Sprache und Stimme*. Kriminalistik Verlag, Heidelberg
- Ladd, D.R., Silverman, K., Tolkmitt, F., Bergmann, G., Scherer, K.R. (1985) Evidence for the independent function of intonation contour type, voice quality, and f0 range in signalling speaker affect. *Journal of the Acoustic Society of America*, 78(2), 435-444.

- Lalouschek, J. (2002) Frage-Antwort-Sequenzen im ärztlichen Gespräch. In: Brünner, G., Fiehler, R. & Kindt, W. (Hrsg.) *Angewandte Diskursforschung*, Band 1. Verlag für Gesprächsforschung.
- Lalouschek, J. (2002b) Ärztliche Gesprächsausbildung. Eine diskursanalytische Studie zu Formen des ärztlichen Gesprächs. Radolfzell, Verlag für Gesprächsforschung.
- Lalouschek, J. és Menz, F. (2002) Empirische Datenerhebung und Authentizität von Gesprächen. In Brünner, G., Fiehler, R., Kindt, W. (Hsg.) *Angewandte Diskursforschung*, Band 1., 46 – 68. Radolfzell, Verlag für Gesprächsforschung.
- Lalouschek, J., Menz, F., Wodak, R. (1990) *Alltag in der Ambulanz. Gespräche zwischen Ärzten, Schwestern und Patienten*. Gunter Narr Verlag: Tübingen.
- Lang, P. J., Bradley, M. M., Cuthbert, B. N. (2005) *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-6. Gainesville, FL: University of Florida.
- László, J. (2005a) A narratív pszichológiai tartalomelemzés. *Magyar Tudomány*, 11, 1366–1376.
- László, J. (2005b) A történetek tudománya: Bevezetés a narratív pszichológiába. Pszichológiai Horizont; Budapest, Új Mandátum Könyvkiadó.
- Laukka, P. (2004) Vocal expression of emotion: Discrete-emotions and dimensional accounts. In *Comprehensive summaries of Uppsala dissertations from the faculty of social sciences*, No. 141, pp. 1-80. Uppsala, Sweden: Acta Universitatis Upsaliensis.
- Laukka, P., Linnman, C., Åhs, F., Pissioti, A., Frans, Ö., Faria, V., Michelgård, Å., Appel, L., Fredrikson, M., Furmark, T. (2008) In a Nervous Voice: Acoustic Analysis and Perception of Anxiety in Social Phobics' Speech. *Journal of Nonverbal Behavior*, 32(4), 195-214.
- Laukka, P., Juslin, P. N., Bresin, R. (2005) A dimensional approach to vocal expression of emotion. *Cognition and Emotion*, 19, 633-653.
- Laukka, P., Neiberg, D., Forsell, M., Karlsson, I., Elenius, K. (2011) Expression of affect in spontaneous speech: Acoustic correlates and automatic detection of irritation and resignation. *Computer Speech and Language*, 25, 84–104.

- Lester, L., Skousen, R. (1974). The phonology of drunkenness. In A. Bruck, R. Fox & M. W. LaGaly (Eds.), *Papers from the parasession on natural phonology* (pp. 233-239). Chicago: Linguistic Society.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Lorenz, M., Cobb, S. (1952) Language behavior in manic patients. *Arch. Neur. Psych.*, 67, 763–70
- Mahl, G.F. (1956) Disturbances and silences in the patient's speech in psychotherapy. *J Abnorm Psychol*, 53(1), 1-15.
- Markó, A. (1999) Intonációról az ezredfordulón, avagy minden így marad? *Egyetemi Fonetikai Füzetek* 25., 29–41.
- Marslen-Wilson, W.D., Tyler L.K. (1980)The temporal structure of spoken language understanding. *Cognition*, 8, 1–7.
- Matsumoto, D. (1993) Ethnic-Differences in Affect Intensity, Emotion Judgments, Display Rule Attitudes, and Self-Reported Emotional Expression in an American Sample. *Motivation and Emotion*, 17(2), 107-123.
- Maylor, E. A., Rabbitt P. M. A (1987) Effect of alcohol on rate of forgetting. *Psychopharmacology*, 91, 2, 230–235.
- McCluskey, K.W. and Albas, D.C. (1981) Perception of the Emotional Content of Speech by Canadian and Mexican Children, Adolescents, and Adults. *International Journal of Psychology*, 16(2), 119-132.
- McCluskey, K.W., Albas, D.C., Niemi, R.R., Cuevas, C., Ferrer, C.A. (1975) Cross-Cultural Differences in Perception of Emotional Content of Speech - Study of Development of Sensitivity in Canadian and Mexican Children. *Developmental Psychology*, 11(5), 551-555.
- Menz, F., Lalouschek, J., Gstettner, A. (2008) *Effiziente ärztliche Gesprächsführung. Optimierung kommunikativer Kompetenz in der ambulanten medizinischen Versorgung. Ein gesprächsanalytisches Trainingskonzept*. LIT-Verlag: Wien/Münster.

- Mergenthaler, E. (1996) Emotion-abstraction patterns in verbatim protocols: a new way of describing psychotherapeutic processes. *Journal of Consulting and Clinical Psychology*, 64, 1306–15
- Miller, L. L., Dolan, M. P. (1974) Effects of alcohol on short term memory as measured by a guessing technique. *Psychopharmacology*, 35, 4, 353–364.
- Mitchell, C. J., Menezes, C., Williams, J. C., Pardo, B., Erickson, D., Fujimura, O. (2000) Changes in syllable and boundary strengths due to irritation. *SpeechEmotion-2000*, 98-103.
- Moskowitz, H., & Roth, S. (1971). Effect of alcohol on response latency in object naming. *Quarterly journal of studies on alcohol*, 32(4), 969-975.
- Németh D., Pléh Cs. (2001) Nyelvfeldolgozás, munkamemória, fordítás. *Fordítástudomány*, 3, 1, 40–53.
- Nordby, K., Watten, R. G., Raanaas, R. K., Magnussen, S. (1999) Effects of Moderate Doses of Alcohol on Immediate Recall of Numbers: Some Implications for Information Technology. *Journal of Studies on Alcohol*, 60, 6, 873–878.
- Nuotto, E., Mattila, M. J. (1990) Actions and interactions of alcohol on performance tests. In S. Parvez & H. Parvez (Series Eds.) & H. Ollat, *Alcohol and Behavior*, 117–130. VSP Press, Utrecht
- Nyberg, L., Tulving, E. (1996) Classifying human long-term memory: evidence from converging dissociations. *European Journal of Cognitive Psychology*, 8, 163–183.
- Oatley, K., Jenkins (2001) *Érzelmek*. Osiris, Budapest.
- Osgood, C.E., Suci, G., & Tannenbaum, P. (1957) *The measurement of meaning*. Urbana, IL: University of Illinois Press.
- Pell, M.D., Paulmann, S., Dara, C., Allasseri, A., Kotz, S.A. (2009) Factors in the recognition of vocally expressed emotions: A comparison of four languages. *Journal of Phonetics*, 37(4), 417-435.
- Pell, M.D., Skorup, V. (2008) Implicit processing of emotional prosody in a foreign versus native language. *Speech Communication*, 50(6), 519-530.
- Pennebaker, J. W., Francis, M. E. (1996) Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion*, 10, 601–26.

- Pennebaker, J. W., Francis, M. E., Booth, R. J. (2001) *Linguistic Inquiry and Word Count (LIWC): LIWC 2001*. Mahwah, NJ: Erlbaum
- Pereira, C. (2000) Dimensions of emotional meaning in speech. *Proceedings of the ISCA Workshop on Speech and Emotion*, 25–28.
- Pisoni D. B., Martin C. S. (1989) Effects of alcohol on the acoustic-phonetic properties of speech: perceptual and acoustic analyses. *Alcoholism: Clinical Experimental Research*, 13, 4, 577–587.
- Pisoni, D. B., Hathaway, S. N., Yuchtman, M. (1985). *Effects of alcohol on the acoustic-phonetic properties of speech: Final report to GM Research Laboratories. (SRL Technical Note No. 85-03)*. Bloomington: Indiana University.
- Pisoni, D. B., Johnson, K., Bernacki, R. H. (1991) Effects of alcohol on speech. In *Proceedings of the Human Factors Society* (pp. 694-698). Santa Monica, CA: Human Factors Society.
- Pittam, J. és Scherer, K. R. (1993). Vocal expression and communication of emotion. In Lewis, M. & Haviland, J. M. (Eds.) *Handbook of emotions*. Guilford Press, New York, 185–197.
- Pléh, Cs., Czigler, I., (1979) Kísérlet politikai kifejezések megítélésére alkalmas szemantikus differenciál skálák kidolgozására és alkalmazására. In: Hunyadi György (szerk.), *Pszichológiai Tanulmányok 15*. Budapest: Akadémiai Kiadó, 479-533.
- Pohárnok M. (2004) A térben való mozgás narratív dimenziójának vizsgálata borderline és depressziós betegek élettörténeti epizódjaiban. In: Erős, F. (szerk.), *Az elbeszélés az élmények kulturális és klinikai elemzésében*. Akadémiai, Budapest, 153-166.
- Pólya T., Kis B., Naszódi M., László J. (2005) Az érzelmi tapasztalat minősége az élettörténeti elbeszélésben. A LAS-verticum perspektíva modulja. *Pszichológia*, 25, 2, 143-155.
- Pólya, T. (2004) A narratív perspektíva hatása az elbeszélő személy észlelésére. In: Erős, F. (szerk.), *Az elbeszélés az élmények kulturális és klinikai elemzésében*. Akadémiai, Budapest, 89-107.
- Pólya, T., Kis, B., Naszódi, M., László, J. (2007) Narrative perspective and the emotion regulation of a narrating person. *Empirical Text and Culture Research* 3, 50-61.

- Pope, B., Blass, T., Siegman, A.W., Rahe, J. (1970) Anxiety and depression in speech. *J Consult Clin Psychol.* 35(1), 128-33.
- Putnoky, J. (1978) A „motoritás” a szójelentés három főbb működési körében. *Magyar Pszichológiai Szemle*, 35, 203-213.
- Racsmány, M., Lukács, Á., Németh, D., Pléh, Cs. (2005) A verbális munkamemória magyar nyelvű vizsgálóeljárásai. *Magyar Pszichológiai Szemle*, 60, 4, 479–505.
- Riviello, M. T., Esposito, A., Vicsi, K (2012) A Cross-Cultural Study on the Perception of Emotions: How Hungarian Subject Evaluate American and Italian Emotional Expressions. *Lecture Notes In Computer Science* 7403, pp. 424-433.
- Rochester, S.R. (1973) The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research*, 2(1), 51-81.
- Roessler, R., Lester J. W. (1976) Voice predicts affect during psychotherapy. *Journal of Nervous and Mental Disease*, 163, 3, 166-76.
- Santamaria, J. N. (1989) The effects of ethanol and its metabolism on the brain. In K. E. Crow & R. D. Batt (Eds.), *Human metabolism of alcohol: Vol. Ill. Metabolic and physiological effects of alcohol* (pp. 35-47). Boca Raton, FL: CRC Press.
- Sator, M. (2003) *Zum Umgang mit Relevanzmarkierungen im ArztInnen-PatientInnen-Gespräch. Eine konversationsanalytische Fallstudie eines Erstgesprächs auf der onkologischen Ambulanz.* Wien, Univ., Dipl.-Arb.
- Scherer K. R. (2013) *Vocal markers of emotion: Comparing induction and acting elicitation.* *Computer Speech & Language* 27(1): 40-58
- Scherer, K. R. (2000) Psychological models of emotion. In Borod, J. (Ed.). *The neuropsychology of emotion* (pp. 137–162). Oxford/New York: Oxford University Press.
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication* 40. 227–256.
- Scherer, K. R. és Zentner, M. R. (2001) Emotional effects of music: production rules. In Juslin, P. N. & Sloboda, J. A. (ed) *Music and emotion: theory and research.* Oxford University Press, Oxford, New York, 361–392.

- Scherer, K. R., Banse, R., Wallbott, H.G., Goldbeck, T. (1991) Vocal Cues in Emotion Encoding and Decoding. *Motivation and Emotion*, 15(2), 123–148.
- Scherer, K.R., Banse, R., Wallbott, H.G. (2001) Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, 32(1), 76-92.
- Scherer, K.R., Ladd, D.R., Silverman K.E.A. (1984) Vocal Cues to Speaker Affect - Testing 2 Models. *Journal of the Acoustical Society of America*, 76(5), 1346-1356.
- Schiel, F., Heinrich, C. (2009) Laying the Foundation for In-car Alcohol Detection by Speech. *Proceedings of the INTERSPEECH 2009*, Brighton, UK, pp. 983–986.
- Schönpflug, U. (2008) Pauses in elementary school children's verbatim and gist free recall of a story. *Cognitive Development*, 23, 385-394.
- Schröder, M., (2003) Experimental study of affect bursts. *Speech Communication*, 40(1-2), 99-116.
- Schuller, B., Müller, R., Land, M., Rigoll, G. (2005) Speaker independent emotion recognition by early fusion of acoustic and linguistic features within ensembles. In *Proceedings of INTERSPEECH-2005*. Lisbon, 805–808.
- Selting, M., Auer, P., Barden, B., Bergmann, J., Couper-Kuhlen, E., Günther, S., Meier, C., Quasthoff, U., Schlobinski, P., Uhmman, S. (1998) Gesprächsanalytisches Transkriptionssystem. *Linguistische Berichte* 173. 91–122.
- Smith, R. C., Parker, E. S., Noble, E. P. (1975). Alcohol's effect on some formal aspects of verbal social communication. *Archives of General Psychiatry*, 32, 1394-1398.
- Sobell, L. C., Sobell, M. B. (1972). Effects of alcohol on the speech of alcoholics. *Journal of speech and hearing research*, 15(4), 861-868.
- Sobell, L. C., Sobell, M. B., & Coleman, R. F. (1982) Alcohol-induced disfluency in nonalcoholics. *Folia Phoniatica*, 34, 316–323.
- Sobin, C., Alpert, M. (1999) Emotion in Speech: The Acoustic Attributes of Fear, Anger, Sadness, and Joy. *Journal of Psycholinguistic Research*, 28, 4, 347–365.

- Starmer, G. A. (1989). Effects of low to moderate doses of ethanol on human driving-related performance. In K. C. Crow & R. D. Batt (Eds.), *Human metabolism of alcohol* (Vol. 1, pp. 101-130). Boca Raton, FL: CRC Press.
- Stirman, S. W. és Pennebaker, J. W. (2001) Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine*, 63, 517–22.
- Stone, L. D., Pennebaker, J. W. (2002) Trauma in real time: Talking and avoiding online conversations about the death of Princess Diana. *Basic and Applied Social Psychology*, 24, 172–82.
- Suzdak, P. D., & Paul, S. M. (1987). Ethanol stimulates GABA receptor-mediated Cl⁻ ion flux in vitro: possible relationship to the anxiolytic and intoxicating actions of alcohol. *Psychopharmacology Bulletin*, 23(3), 445-451.
- Szabó, E. (2008) A szomorú és a vidám érzelmi állapot megjelenése a beszédben. *Magyar Pszichológiai Szemle*, 63, 4, 651–668.
- Szabó, E., Rossu, R., Varga, V., Imrécz, K. (2012). Az alkohol hatása a rövid távú emlékezetre és a beszédprodukciós folyamatokra. *Magyar Pszichológiai Szemle*, 67(3), 409–430.
- Tisljár-Szabó, E. (2014) Érzelem és beszéd. In: Pléh Cs. és Lukács Á. (szerk.) A pszicholingvisztika magyar kézikönyve. Budapest, Akadémiai Kiadó. (megjelenés alatt)
- Tisljár-Szabó, E., Pléh, C. (2014) Ascribing emotions depending on pause length in native and foreign language speech. *Speech Communication*, 56, 35–48.
- Tisljár-Szabó, E., Rossu, R., Varga, V., Pléh, Cs. (2014) The effect of alcohol on speech production. *Journal of Psycholinguistic Research*. (Online first, DOI 10.1007/s10936-013-9278-y)
- Thompson, W.F., Balkwill, L.L. (2006) Decoding speech prosody in five languages. *Semiotica*, 158(1-4), 407-424.
- Tóth, Sz. L., Sztahó, D., Vicsi, K. (2008) Speech Emotion Perception by Human and Machine. In: Proceeding of COST Action 2102 International Conference: Revised Papers in Verbal and Nonverbal Features of Human-Human and Human-Machine Interaction. Springer, pp. 213-224.
- Tringer, L. (2005) *A pszichiátria tankönyve*. Semmelweis Kiadó, Budapest.

- Trojan, F., Kryspin-Exner, K. (1968) The decay of articulation under the influence of alcohol and paraldehyde. *Folia Phoniatrica (Basel)*, 20, 4, 217-238.
- Türk, O., Schröder, M. (2008) A comparison of voice conversion methods for transforming voice quality in emotional speech synthesis. *Proc. Interspeech 2008*, Brisbane, Australia.
- Tüske, Z., Simon, M., Mihajlik, P., Fegyő, T. (2007) Érzelmek automatikus felismerése a beszéd akusztikus jellemzői alapján. In: Gósy M (ed.) *Beszéd kutatás 2007*. Budapest: MTA Kiadó, pp. 151-161.
- Tyson, P. D., Schirmuly, M. (1994) Memory enhancement after drinking ethanol: consolidation, interference, or response bias? *Psychology & Behavior*, 56, 5, 933–937.
- VanBezooijen, R., Otto, S.A., Heenan, T.A. (1983) Recognition of Vocal Expressions of Emotion - a 3-Nation Study to Identify Universal Characteristics. *Journal of Cross-Cultural Psychology*, 14(4), 387-406.
- Velten, E. (1968) A laboratory task for induction of mood states. *Behavior Research & Therapy*, 6(4), 473-482.
- Ververidis, D., Kotropoulos, C. (2006) Emotional speech recognition: resources, features, and methods. *Speech Communication*, 48, 1162–1181.
- Vicsi K. (2010) Érzelemfelismerés. In: Németh G, Olaszy G (szerk.) *A magyar beszéd: Beszéd kutatás, beszédtechnológia, beszédinformációs rendszerek*. Budapest, Akadémiai Kiadó, pp. 392-401.
- Vicsi, K., Sztahó, D. (2009) Ügyfél érzelmi állapotának detektálása telefonos ügyfélszolgálati dialógusban. In: Tanács Attila, Szauter Dóra, Vincze Veronika (szerk.) *VI. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged, Magyarország, 2009.12.03-2009.12.04. Szeged: JATEPress, pp. 217-225.
- Vicsi, K., Sztahó, D. (2011) Problems of the Automatic Emotion Recognitions in Spontaneous Speech; An Example for the Recognition in a Dispatcher Center. Toward Autonomous, Adaptive, and Context-Aware Multimodal Interfaces. Theoretical and Practical Issues. *Lecture Notes in Computer Science Volume 6456*, pp 331-339.
- Vicsi, K., Sztahó, D (2012) Recognition of emotions on the basis of different levels of speech segments. *Journal Of Advanced Computational Intelligence And Intelligent Informatics* 16, 2, pp. 335-340.

- Vicsi, K., Sztahó D., Kiss, G., Czira, A. (2010) Spontán beszédben rejlő nem verbális hangjelenségek - érzelmek, hanggesztusok - vizsgálata. In: Tanács A., Vincze V. (szerk.) MSZNY 2010: VII: Magyar Számítógépes Nyelvészeti Konferencia. Szeged, Magyarország, 2010.12.02-2010.12.03. Szeged: Szegedi Tudományegyetem Informatikai Tanszékcsoport, pp. 249-260.
- Vogt, T., André, E., Wagner, J. (2008) Automatic Recognition of Emotions from Speech: a Review of the Literature and Recommendations for Practical Realisation. In Peter, Ch., Beale, R. (szerk.) *Affect and Emotion in Human-Computer Interaction, From Theory to Applications. Lecture Notes in Computer Science*, 4868, 75–91. Springer.
- Wallgren, H., Barry, H. (1970). *Actions of alcohol*. Amsterdam: Elsevier.
- Watanabe, H., Shin, T., Matsuo, H., Okuno, F., Tsuji, T., Matsuoka, M., . . . Matsunaga, H. (1994). Studies on vocal fold injection and changes in pitch associated with alcohol intake. *Journal of Voice*, 8(4), 340-346.
- Weintraub, W. (1981) *Verbal Behavior: Adaptation and Psychopathology*. New York: Springer
- Westermann, R., Spies, K., Stahl, G., Hesse, F. W. (1996) Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, 26, 4, 557-580.
- Widmark, E. M. P. (1981) *Principles and Applications of Medicolegal Alcohol Determination*. Davis, CA: Biomedical Publications.
- Williams, E.J. (1949) Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Scientific Research*,. Ser. A 2, 149-168.
- Wodak, R. (1981) *Das Wort in der Gruppe. Linguistische Studien zur therapeutischen Kommunikation*. Vienna: Akademie der Wissenschaften.
- Yanushevskaya, I., Gobl, C., Ní Chasaide, A. (2005). Voice quality and f0 cues for affect expression: implications for synthesis. In *Interspeech'2005 – Eurospeech*. Lisszabon, Portugália, 1849–1852.
- Zaimov, K. (1969). Die Sprachstörungen als Kriterium der Bewusstseinsstörungen. *Psychiatrie, Neurologie, und Medizinische Psychologie*, 21, 218–225.