

A néma szünetek időtartamának hatása az érzelmi állapot észlelésére

Szabó Eszter¹

¹ Budapesti Műszaki és Gazdaságtudományi Egyetem, Kognitív Tudományi Tanszék
eszabo@cogsci.bme.hu

Kivonat: Korábbi kutatások alapján a szomorú érzelmi állapotot a néma szünetek megnyúlása, míg a vidám érzelmi állapotot ezek lerövidülése jellemzi. Jelen vizsgálat azt kutatja, hogy a spontán monológokban a szünetek hossza hogyan befolyásolja a beszéd érzelmi töltetének észlelését. Semleges tartalmú monológokat a Praat program segítségével úgy módosítottunk, hogy az eredetileg is meglévő szüneteket mesterségesen megnyújtottuk, illetve lerövidítettük, korábbi kutatások adatai alapján. Hipotézisünk az volt, hogy ha minden más feltétel azonos, a hosszabb szüneteket tartalmazó semleges témájú monológokat szomorúbbnak, míg a rövidebb szüneteket tartalmazó monológokat vidámabbnak fogják észlelni a megítélők, mint az eredeti beszédeket. A kutatás hozzájárulhat azoknak a szoftvereknek a fejlesztéséhez, amelyek a beszéd érzelmi töltetét elemzik.

1 Bevezetés

Az érzelmek és a beszéd kapcsolata az elmúlt években nem csak a pszicholingvisztikában, hanem az informatikában, a szintetizált beszédet előállító, illetve az automatikus beszédfelismeréssel foglalkozó kutatók körében is fontos szerephez jutott. Mivel a legújabb kutatások célja, hogy egyre inkább az emberéhez hasonló beszédet tudjunk előállítani, illetve hogy a természetes nyelv mondatait is fel tudja ismerni programunk, egyre több kutatás foglalkozik a beszéd prozódijával. A beszéd gyakran érzelmekkel telített, és a természetesnek ható szintetizált beszéd elkészítéséhez, illetve a természetes környezetben elhangzott érzelemmenteli beszéd felismerésének eléréséhez fontos, hogy tudjuk, a beszéd mely tulajdonságai hogyan változnak egy-egy érzelem esetében.

A beszédet legtöbbször fonetikai szempontból elemzik, és olyan fizikai változókat vesznek figyelembe, mint a hangmagasság, a hangerő, az időtartamok. Több kutatás alapján úgy tűnik, hogy az érzelemfelismerés szempontjából a hangmagasság a legfontosabb paraméter. Ugyanakkor pl. Kienast és munkatársai [1] kiemelik, hogy a beszédtempó, az időtartam, az artikuláció pontossága mind hasznos paraméterek a beszélő érzelmi állapotának megállapításához.

2 Hosszú idejű változók vizsgálata

2.1 Korábbi kutatások

A legtöbb az érzelem és a beszéd kapcsolatát vizsgáló kutatás rövid, néhány szavas, vagy egymondatos beszédmintákat elemez, amelyek általában színészekről származnak. Ugyanakkor a természetes beszédben gyakoriak a monológok, a több mondatos, hosszabb beszédek. A hosszabb szakaszok elemzésének szükségességét, ennek hiányát a mesterséges beszédészleléses kutatások is kiemelik (ld. Ververidis és Kotropoulos, [2]).

Elmesélt élettörténetekben a szomorúság és reménytelenség érzésének kifejezésekor kutatók azt figyelték meg, hogy mélyebb, halkabb, erőtlenebb, monoton vagy intonáció nélkülivé vált a beszéd. Gyakoribbá váltak a szünetek – különösen a hosszú (egy másodpercnél hosszabb) szünetek, akár egy szintaktikai egységen belül is. Vidám érzelmek esetén ezekkel ellentétes módon változik a beszéd: magasabb, hangozósabb, dallamosabb és gyorsabb lesz (Deppermann és Lucius-Hoene [3]; Scherer [4]).

Egy korábbi kutatásunkban (Szabó [5]) a szomorú és a vidám érzelmi állapot hatását vizsgáltuk kísérleti fonetikai módszerekkel. Itt az önéletrajzi emlékezés módszerét és zenét használva sikerült egy olyan kísérleti elrendezést létrehozni, amelyben a kísérleti személyek szomorú, illetve vidám érzelmi állapotokat éltek át, és ez hatással volt a beszédükre: a beszédtempóra, a beszédbeli szünetek hosszára, a szünetek arányára a teljes időtartamhoz viszonyítva, valamint a hangerőre. Félperces szakaszokat elemezve a hangmagasságban nem mutatkoztak különbségek, és az idői jegyek tűntek fontosnak.

2.2 Jelen kutatás

Jelen vizsgálatunkban arra kerestük a választ, hogy a szünetek hosszának milyen hatása van az érzelemazonosításra. Tehát ha minden más feltétel azonos, akkor csupán a szünetek hosszának megváltozása hogyan befolyásolja a hallgatókat abban, hogy milyen érzelmi állapotot tulajdonítanak a beszélőnek.

A spontán beszédben is elhangozható monológok modellálására semleges témájú beszédrészeket használtunk fel, amelyeket nem hivatásos színészek mondtak el egy-egy beszélgetés során. A beszédekről felvételek készültek, amelyeket aztán a Praat program segítségével módosítottunk. A monológokban egyébként is benne lévő néma szüneteket az irodalomban megtalálható adatok alapján mesterségesen megnyújtottuk, illetve lerövidítettük, és percepció tesztelésnek vetettük alá. A megítélőknek arra kellett választ adniuk, hogy a hallott beszédrészeket milyen érzelmi állapotban mondhatta a beszélő.

Hipotézisünk az volt, hogy a hosszabb néma szüneteket tartalmazó monológokat szomorúbbnak, míg a rövidebb szüneteket tartalmazókat vidámabbnak fogják megítélni, mint az eredeti formájában hagyott beszédrészeket. A kutatás nagyban hozzájárulhat a beszéd érzelmi töltetét felismerni kívánó szoftverek fejlesztéséhez.

Hivatkozások

1. Deppermann, A. & Lucius-Hoene, G.: Trauma erzählen – kommunikative, sprachliche und stimmliche Verfahren der Darstellung traumatischer Erlebnisse. *Psychotherapie und Spezialwissenschaft. Zeitschrift für Qualitative Forschung und klinische Praxis*. 1. (2005) 35–73
2. Kienast, M., Paeschke, A., & Sendlmeier, W.: Articulatory reduction in emotional speech. *Proceedings of Eurospeech 1999, Budapest, Hungary (1999)* 117–120
3. Scherer, K. R.: Vocal communication of emotion: A review of research paradigms. *Speech Communication* 40. (2003) 227–256
4. Szabó E.: A szomorú és a vidám érzelmi állapot megjelenése a beszédben. *Magyar Pszichológiai Szemle*, 63, 4, (2008) 651–668
5. Ververidis, D. és Kotropoulos, C.: Emotional speech recognition: Resources, features, and methods. *Speech Communication*, 48 (2006) 1162–1181