

Budapesti Műszaki és Gazdaságtudományi Egyetem
Pszichológia Doktori Iskola – Kognitív Tudomány



Kocsis Zsuzsanna

Egyidejűleg jelentkező hangok szétválasztásának EEG korrelátumai
(EEG signatures of separating concurrent sounds)

PhD téziszfüzet

Témavezető:
Dr. Winkler István

Budapest, 2017

Előzmények és célkitűzések

A dolgozat fő célja az volt, hogy megvizsgáljuk, hogy az emberi agy hogyan képes az egyidejűleg jelentkező hangokat, illetve egy tárgyat a háttértől elválasztani, és hogy eközben milyen agyi korrelátumokat figyelhetünk meg (Bregman, 1990; Ciocca, 2008). A szimultán hangszétválasztás egy alapvető csoportosító folyamat, azonban sok tulajdonsága még ma sem teljesen tisztázott. Két elektromos agyi potenciál (EAP) jelenik meg a szimultán hangszétválasztási folyamat során: a tárgyhoz köthető negatívitás (TKN; Alain és kollégái, 2001) akkor jelenik meg, ha a hallási jelenet során egynél több hallási tárgy fedezhető fel; amennyiben a hangok relevánsak a kísérleti feladat szempontjából egy későbbi pozitív komponens, a P400 is megjelenik. Ebben a dolgozatban egy a szokásosnál a mindennapi hangkörnyezet több tulajdonságát megjelenítő paradigmát tanulmányoztunk a hallási tárgyak szemszögéből, illetve egyszerű kulcsingerek kombinációinak hatását vizsgáltuk a szimultán hangszétválasztásra. A dolgozat az alábbi kérdésekre igyekezett válaszolni: Elsőként, hogyan bírkózik meg az agyunk egy valószínűbb hallási alak-háttér szétválasztási problémával (a sztochasztikus hang-felhő paradigma)? A TKN kiváltódik-e az alak-háttér elkülönítéskor, amennyiben az egyes hangelemeket mind időben, mind frekvencia-összetevők mentén csoportosítani kell? Az I. tézispont így a TKN válasz általánosíthatóságának kérdésével foglalkozik. Ez a tézispont összegzi a sztochasztikus hang-felhő paradigma elektrofiziológiai vizsgálatának eredményeit, melyben a különböző paraméterek hatásait vizsgáltuk. Másodsorban arra voltunk kíváncsiak, hogy a hallórendszer hogyan használja a különböző szimultán hangcsoportosító kulcsingereket és ezek a kulcsingerek hogyan hatnak egymásra? Mit tükröz a TKN EAP korrelátum a kulcsinger kiértékelése tekintetében: a kulcsingerek egyenkénti feldolgozását vagy a kulcsingerek összesített hatását? A további tanulmányok a szimultán hangszétválasztás specifikus kérdéseinek megválaszolására használják a TKN-t,

illetve az avval párhuzamosan jelentkező agyi oszcillációkat. A II. tézispont azokat az eredményeket mutatja be, melyeket a szimultán hangszétválasztás kulcsingereinek és azok kombinációinak vizsgálatakor kaptunk mind passzív, mind aktív figyelmi helyzetben. Harmadsorban arra voltunk kíváncsiak, hogy milyen típusú EEG oszcillációk követik a szimultán hangszétválasztást. A III. tézispont összegzi azokat az eredményeket, melyek a szimultán hangszétválasztás neurális oszcillációkra gyakorolt hatását mutatja be. Végül, milyen EEG korrelátumokat kapunk, ha többféle kulcsingert konvergens (azonos hangkonfigurációra vezető) vagy divergens (különböző hangkonfigurációra vezető) módon alkalmazunk? A IV. tézispont egy olyan paradigmát mutat be, melyben két, illetve három szimultán megjelenő hang észlelését próbáltuk előidézni a kulcsingerek konvergens, illetve divergens manipulációjával.

Összességében arra akartunk választ kapni, hogy a TKN a hallórendszer átfogó, integrált válaszát jelzi-e a jelenlévő kulcsingerek tekintetében, vagy az egyes kulcsingerek számát és a kialakuló hangszerveződéshez való hozzájárulását tükrözi.

Új tudományos eredmények

I. tézispont: Hallási figura-háttér szétválasztás

A mindennapi hallási környezetünkben az alak és a háttér jelei szinte mindig átfedik egymást időben és frekvenciában is. Vagyis, amennyiben az akusztikus alakot, vagy tárgyat el szeretnénk különíteni a háttértől a hangok egyes elemeit mind időben, mind frekvencia alapján kell csoportosítanunk. A szimultán és a szekvenciális csoportosítás eseményhez kötött agyi potenciál korrelátumait korábban már vizsgálták, azonban a kettőt rendszerint külön kezelték. A tanulmány célja az volt, hogy megvizsgáljuk milyen agyi elektromos válaszok jelentkeznek természetesebb szituációkban, ahol mindkét hangcsoportosító folyamat szerepet játszik. Az alak szálenciáját szisztematikusan változtattuk azáltal, hogy a figura észlelését elősegítő szekvenciális, vagy a szimultán kulcsingereket manipuláltuk, egymástól függetlenül. Ez a paradigma lehetővé tette, hogy egy hallási tárgy háttértől való elkülönülésének elektrofiziológiai korrelátumát vizsgáljuk.

Ebben a tanulmányban azt az eredményt kaptuk, hogy az alak helyes azonosítása az TKN és a P400 komponens kiváltását eredményezte, amely alátámasztotta, hogy ezek a komponensek figyelhetőek meg akkor is, ha a szimultán hangszétválasztáshoz a szekvenciális kulcsingerek integrációja is szükséges. A korábbi eredményekkel összhangban azt találtuk, hogy a P400 komponens a detekciós teljesítménnyel korrelált. Emellett a forrás lokalizációs eredmények is a korábbi tanulmányokkal összecsengő eredményt mutattak: mind a TKN, mind a P400 komponens generátorai a temporális kéregben találhatóak (lásd: Alain és McDonald, 2007).

II. tézispont: Többszörös kongruens kulcsingerek hatása a szimultán hangszétválasztásra

A szimultán hangszétválasztás azonnal hozzáférhető kulcsingereken alapszik, mint például a hangmagasság-különbség, a hangkezdeti időben lévő eltérés, a hangforrás helyének eltérése, és akár ezek kombinációja. A legtöbb kulcsingert önmagában vagy más kulcsingerrel kombinálva már többször vizsgálták, azonban a többszörös kulcsingerek szisztematikus vizsgálata passzív és aktív helyzetben még váratott magára.

Ebben a tanulmányban a különböző kulcsingerek kombinációinak hatását vizsgáltuk a TKN és a P400 EAP komponensek tekintetében. A kísérleti személyek összetett hangokat hallottak, melynek fele valamilyen módon manipulált volt: egy vagy két harmonikust elhangoltunk, idői késleltetéssel, vagy eltérő forrás helyszínről mutattunk be, mint a többi harmonikust. Azt vizsgáltuk, hogy az elhangolás kulcsinger szálienciáját tudjuk-e fokozni amennyiben nem egy, hanem két harmonikust hangolunk el kongruens módon, illetve hogy a kombinált kulcsingerek hatása összeadódik, esetleg kisebb vagy nagyobb lesz, mintha az egyes kulcsingerek által kiváltott TKN komponensek összeadódtak volna. Ha mindegyik kulcsinger egy külön TKN-t vált ki, a többszörös kongruens kulcsingerek által kiváltott válasz amplitúdója akkora lesz, mint az egyes kulcsingerek által kiváltott válaszok amplitúdójának összege. Ez azt jelezné, hogy a TKN olyan folyamatokat tükröz, melyek a kulcsingerek kiértékelését követi szorosan, nem pedig az észlelt hangkonfigurációt. Ezzel szemben a TKN a hallórendszer egészleges értékelését is tükrözheti arra vonatkozóan, hogy az bejövő hangfolyam több szimultán hallási tárgyból áll. Ez esetben a TKN az összes jelen lévő kulcsinger kombinált értékelését tükrözi, vagyis a többszörös kulcsingerek által kiválasztott TKN amplitúdója kisebb vagy nagyobb lenne, mint az egyes kulcsingerek által kiváltott TKN-ok amplitúdójának összege.

Az eredmények alapján elmondható, hogy a különböző kulcsingerek nem egyformán hatásosak a TKN kiváltásában (a térbeli különbség kulcsinger a többinél kevésbé volt hatékony), valamint hogy a TKN válaszok nem adódnak össze, ha a szimultán kulcsingereket kombináljuk. Ez alapján elmondható, hogy a TKN komponens nem az egyes kulcsingerek kiértékelését tükrözi. A P400 komponens szintén megfigyelhető volt a legtöbb kondícióban az aktív figyelmi helyzetben, azonban kevésbé érzékenynek bizonyult az egyes kulcsingerekre, illetve azok kombinációira, mint a TKN.

III. tézispont: A szimultán hangszétválasztást kísérő théta oszcillációk

Ebben a tanulmányban a szimultán hangszétválasztáshoz kapcsolódó agyi oszcillációk feltérképezését tűztük ki célul, mivel nem született eddig átfogó beszámoló a témában. Az összetett hangokat egy vagy két szimultán jelenlévő hangként észlelhetik az emberek a összetevők elhangoltságának, a hangkezdeti idő eltérésének függvényében, illetve a térben különböző hangforrások esetében. Különböző kondíciókban a kísérleti személyek vizuális változási detekciós feladatot végeztek (vizuális kontroll feladat), egy lehalkított filmet néztek miközben hallották a feladathoz kapcsolódó hangokat (passzív hallgatási feladat), vagy jelezték minden hang után, hogy egy, vagy két hangot hallottak-e (aktív hallgatási feladat). Célunk az volt, hogy meghatározzuk, milyen neurális oszcillációk köthetőek a szimultán hangszétválasztáshoz azáltal, hogy összehasonlítottuk az oszcillációs tevékenységet a különböző kulcsingerekhez köthető szimultán hangszétválasztás során, illetve a különböző figyelmi feltételek esetében.

A szimultán kulcsingerek megnövekedett théta aktivitást váltottak ki két intervallumban – egy korai és egy későbbi idői ablakban – melyek megegyeznek a TKN és a P400 komponensek idői ablakával. Az eseményhez kötött spektrális perturbációk nagyobbak

voltak az aktív, mint a passzív figyelmi feladatban. Az eredmények alapján elmondható, hogy a TKN és a P400 EAP komponensek kiváltásában részt vevő agyi hálózatok theta hullámok segítségével kommunikálhatnak.

IV. tézispont: Két vagy három szimultán jelen lévő hallási tárgy

Két szimultán jelenlévő tárgy szétválasztását már többször vizsgálták, azonban három szimultán jelenlévő tárgy szétválasztásával korábban még nem vizsgálták. Az összetett hangokon alkalmazott kulcsingereket illetve azok kombinációit a jelen kísérletben úgy alakítottuk ki, hogy azok egy, kettő, illetve három egyszerre jelenlévő hang észelését segítsék elő.

Kutatásunkban a divergens kulcsingerek hatását vizsgáltuk a szimultán hangszétválasztás perceptuális és neurális (TKN) mutatóira. Ha a divergens manipulációk (három tárgy kondíció) által kiváltott TKN amplitúdója akkora, mint az egyes kulcsingerek által kiváltott TKN amplitúdók összege, akkor a TKN a divergens kulcsingerek független értékelését tükrözi (vagyis az egyes kulcsingerek értékeléséhez áll közel). Amennyiben a divergens kulcsingerek által kiváltott TKN amplitúdója általában kisebb, mint az egyes kulcsingerek által kiváltott TKN-ok összegzett amplitúdója, a TKN egyszerre több szimultán hallási tárgy jelenlétét tükrözi, figyelmen kívül hagyva a tárgyak számát, vagy a különböző kulcsingereket. További kérdésünk az volt, hogy a TKN amplitúdó vajon a megjelenő hallási tárgyak számát tükrözi-e, vagy azt, hogy egynél több hallási tárgy van jelen az adott hangkörnyezetben.

Azt találtuk, hogy lehetséges egyszerre három jelenlévő tárgy észlelését előidézni az elhangolt harmonikus paradigmában, azonban a kísérleti személyek nem voltak képesek megbízhatóan elkülöníteni a két, illetve három hallási tárgyat tartalmazó hangokat. Emellett

azt találtuk, hogy a TKN kiváltódik a kulcsingerek kombinációjának legtöbb esetében. Azok a hangok váltották ki a legkisebb amplitúdójú TKN-t, melyben a különbség amplitúdó-, vagy frekvencia-modulációt tartalmazott. Ha két harmonikust divergens módon manipuláltunk, akkor numerikusan nagyobb amplitúdójú TKN-t kaptunk, mint a konvergensen manipulált hangok esetében, de az amplitúdó nem érte el az egyedül alkalmazott kulcsingerek által kiváltott TKN amplitúdók összegét. Ezen eredmények alapján elmondható, hogy a TKN a hallórendszer egészleges értékelését tükrözi arra vonatkozóan, hogy egy vagy több egyidejű hallási tárgy van-e jelen, nem pedig a jelenlévő tárgyak számát.

Összegzés és kitekintés

Az eredmények azt mutatják, hogy a TKN eseményhez kötött agyi potenciál tükrözi egy hallási tárgy háttérből való kiemelését a stochasztikus figura-háttér paradigmában. A TKN a szimultán hallási tárgyak szétválasztását is tükrözi amennyiben egy vagy több kulcsingert alkalmazunk összetett hangokon. A TKN komponens a P400 követi olyan helyzetekben, amikor a hangok relevánsak a feladat szempontjából. A TKN amplitúdójának maximuma az inger kezdete után 150-250 ms-mal figyelhető meg, és egy korábbi, figyelemtől független folyamatot tükröz, míg a P400 amplitúdójának maximuma az inger kezdete után 400 ms-mal figyelhető meg, egy későbbi, figyelmi feldolgozást tükröz.

A II. és IV. tézispontok alapján elmondható, hogy a TKN a hallási rendszer egészleges értékelését mutatja arra vonatkozóan, hogy egyszerre több hallási tárgy van jelen. A P400 pedig a perceptuális döntési folyamathoz kapcsolható. Tehát a TKN egy inger-vezérelt primitív csoportosítási folyamatot jelez, a P400 magában foglalja az akaratlagos hatásokat a perceptuális döntésben.

A TKN és a P400 tulajdonságai egybecsengenek Bregman (1990) kétfázisú elméleti rendszerével. Az első fázisban, melyben a hallórendszer az akusztikus inputot lehetséges perceptuális csoportokká alakítja (vagyis proto-tárgyakká), a TKN egy kétállású kapcsolót reprezentál, mely azt jelzi, hogy valószínűleg egy vagy több hallási tárgy lehet-e jelen. Ha egyetlen kulcsinger illetve azok kombinációja sem ér el egy bizonyos határt, a bejövő jelet egy hallási tárgyként értelmezzük. Amikor az ingerből kiemelt jelzőmozzanatok értékelése átlépi ezt a határt (függetlenül a kongruens kulcsingerek számától), több mint egy hallási tárgy jelenlétét jelzi a rendszer és megkezdődik az alternatívák közötti versengés. Ezt a versengést a figyelem befolyásolhatja, kiválasztva az alternatívát, amelyiket tudatosan is észlelünk.

Egyes elméletek szerint a szimultán hangszétválasztás folyamata megelőzi a szekvenciális hangszétválasztást (lásd Winkler és Schröger (2015) hallási események keletkezésének modelljét). Ez lehetővé teszi a szekvenciális hangszétválasztás számára, hogy integrálja a szimultán hangszétválasztás eredményét, és a hallórendszer így hozzon létre proto-tárgyakat a szimultán megjelenő hangok egy-egy csoportjából.

Emellett, a jelen eredmények a hallási jelenetelemzés idői koherencia modelljével is kompatibilisek (Shamma és kollégái, 2010; 2011). Ebben a rendszerben a TKN mint az idői koherencia detektora venne részt. Teki és kollégái (2011; 2013) azt feltételezik, hogy a feldolgozás első lépésében a tulajdonságok analízise történik meg, míg a második lépésben a hallórendszer az előző lépésből származó output-ot csoportosítja az idői koherencia alapján. Mivel a TKN a kulcsingerek egészleges értékelését tükrözi, úgy tekinthető, mint egy idői koherencia detektor. Teki és kollégái (2013) szerint az első két lépés mellett a figyelemnek is fontos szerep jut a hangláncok formálásában, mivel a figyelem egy bizonyos hangcsoportosítás felé tereli az észlelést, mely a személy aktuális viselkedéses és perceptuális céljaitól függ. Ilyen tekintetben a P400 ennek a befolyásoló folyamatnak az eredménye: a cél

tulajdonságainak detekciója. Fontos megjegyezni, hogy bizonyos neurofiziológiai jelek érzékenyek az idői koherenciára (O'Sullivan és mtsai., 2015). Vagyis elképzelhető, hogy az idői koherencia az agyban ténylegesen kiszámolásra kerül. Azonban ez még nem mérvadó a két elmélet közötti döntésben, mivel az idői koherencia alapú csoportosítás létrejöhet a Bregman-féle hallási jelenetelemzés első lépése során is (vagyis olyan szerepet játszana, mint a közös sors Gestalt-elv heurisztikus implementációja).

További kutatást igényelne, hogy dönteni lehessen a Bregman-i és az idői koherencia modell között. Olyan kísérletek által lehetne betekintést nyerni, melyek más csoportosító mechanizmusokat alkalmaznak, mint az idői koherencia, mint például megjósolható mitázatok beillesztése hallási láncokba (például egy ismerős dallam egy ismeretlen háttérbe ágyazva, lásd: Szalárdy és mtsai., 2014), száliens vagy nem várt ingerek, amelyek nem konzisztensek a hallási környezettel (például kutyaugatás a fürdőszobában), vagy olyan beszédláncok bemutatása, melyekben a szemantikai megfelelés elősegíti a szétválasztást. Másik lehetséges bizonyíték származhatna az alternatív (nem domináns) hangszerveződések reprezentációjának kimutatásából, mivel ezek az idői koherencia modell alapján nem számolhatóak. Olyan idegsejt populációk létezése, melyek aktuálisan nem észlelt hallási tárgyakat kódolnak, alátámasztaná a Bregman-féle hallási jelenetelemzés modellt.

További nyitott kérdések a következők: Elsőként, jelenleg keveset tudunk a szimultán hangszétválasztás idegi hátteréről. fMRI mérések fényt deríthetnének arra, hogy mely agyi régiók vesznek részt a folyamatban. Másodsorban, a sztochasztikus figura-háttér paradigma közelebb vitte a kutatást a realisztikusabb hallási környezetekhez. Azonban még ezen a téren is lehetne finomítani az eljárásokon: a szimultán hangszétválasztást lehetne természetes hangok láncai által tanulmányozni (nem véletlen zajból álló háttér mellett): például olyan környezeti hangok segítségével, melyek tartalmazznak egy azonosítható hallási tárgyat. Az eddig alkalmazott paradigmák leginkább rövid hangokat használtak. A jövőben olyan

kutatásokat lenne célszerű folytatni, melyekben folytonos hangokat hallanak a kísérleti személyek, és például az elhangolás a hosszú hang közben történne. Végezetül, a szimultán hangszétválasztás vizsgálata a beszédláncok szétválasztásának folyamatában segítene jobban megérteni, hogy a hallórendszerünk hogyan oldja meg a koktélparti problémát.

A tézispontokhoz kapcsolódó tudományos közlemények

- I. Tóth, B., Kocsis, Z., Háden, G.P., Szerafin, Á., Shinn-Cunningham, B., & Winkler, I. (2016). EEG signatures accompanying auditory figure-ground segregation. *Neuroimage*, 141, 108-119. DOI: 10.1016/j.neuroimage.2016.07.028.
- II. Kocsis, Z., Winkler, I., Szalárdy, O., & Bendixen, A. (2014). Effects of multiple congruent cues on concurrent sound segregation during passive and active listening: An event-related potential (ERP) study. *Biological Psychology*, 100, 20-33. DOI: 10.1016/j.biopsycho.2014.04.005.
- III. Tóth, B., Kocsis, Z., Urbán, G., & Winkler, I. (2016). Theta oscillations accompanying concurrent auditory stream segregation. *International Journal of Psychophysiology*, 106, 141-151. DOI: 10.1016/j.ijpsycho.2016.05.002.
- IV. Kocsis, Z., Winkler, I., Bendixen, A., & Alain, C. (2016). Promoting the perception of two and three concurrent sound objects: an event-related potential study. *International Journal of Psychophysiology*, 107, 16-28. DOI: 10.1016/j.ijpsycho.2016.06.016.

Felhasznált irodalom

- Alain, C., Arnott, S. R., & Picton, T. W. (2001). Bottom-up and top-down influences on auditory scene analysis: Evidence from brain potentials. *Journal of Experimental Psychology*, 27, 1072-1089.
- Alain, C., & McDonald, K. L. (2007). Age-related differences in neuromagnetic brain activity underlying concurrent sound perception. *Journal of Neuroscience*, 27, 1308- 1314.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: The MIT Press.
- Ciocca, V. (2008). The auditory organization of complex sounds. *Frontiers in Bioscience*, 13, 148-169.
- O'Sullivan, J. A., Shamma, S. A., & Lalor, E. C. (2015). Evidence for neural computations of temporal coherence in an auditory scene and their enhancement during active listening. *Journal of Neuroscience*, 35(18), 7256-7263.
- Shamma, S. A., & Micheyl, C. (2010). Behind the scenes of auditory perception. *Current Opinion in Neurobiology*, 20, 361-366.
- Shamma, S. A., Elhiali, M., & Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. *Trends in Neuroscience*, 34(3), 114–123.
- Szalárdy, O., Bendixen, A., Böhm, T. M., Davies, L. A., Denham, S. L., & Winkler, I. (2014). The effects of rhythm and melody on auditory stream segregation. *Journal of the Acoustical Society of America*, 135(3), 1392-1405.
- Teki, S., Chait, M., Kumar, S., von Kriegstein, K., & Griffiths, T. D. (2011). Brain bases for auditory stimulus-driven figure-ground segregation. *The Journal of Neuroscience*, 31, 164-171.
- Teki, S., Chait, M., Kumar, S., Shamma, S., & Griffiths, T. D. (2013). Segregation of complex acoustic scenes based on temporal coherence. *eLife*, 2: e00699.
- Winkler, I., & Schröger, E. (2015). Auditory perceptual objects as generative models: Setting the stage for communication by sound. *Brain and Language*, 148, 1-22.