

Working paper

Bódog Alexa

Természetesség és multimodalitás az ember–gép interakciókban¹

Fejlődéslélektani megfontolások

Abstract

The basic assumption of the paper is that in order to model quasi–natural human–computer interactions, models of HCI must be based on the features of human–human communication. Therefore, we must focus on such properties which are suitable for analogue comparisons between naturally occurring human–human communication and human–computer interaction. The analogical similarities contribute to build quasi–natural models of human–computer interaction. The aim of my paper is to explore such properties of early social development which could inspire human–computer interaction models.

Keywords: infancy, social development, intermodal coordination, imitative learning, supramodal representation system, proprioception

0 A dolgozat célja

A jelenlegi ember–gép interakciós technológiák nem rendelkeznek az emberi összetevő természetességével és spontaneitásával (Hunyadi 2011). Ez a merevség azt eredményezi, hogy a gépi (számítógépes) rendszerek alapvetően az úgynevezett command–and–control elv alapján egy bizonyos típusú bemenetre (emberi írás, hang, tekintet, kézmozdulat) egy bizonyos típusú kimenetet szolgáltatnak (például egérmutató mozgatása tekintet általi irányítással), s nem használják ki a különböző modalitású információk ötvözése nyújtotta lehetőségeket sem (Seppänen 2009). A legtöbb ember–gép interakciós technológiának ez a rigiditás komoly korlátokat szab a fejlesztés és a fölhasználás területein egyaránt (Hunyadi 2011). Éppen ezért dolgozatom célja olyan elméleti lehetőségek bemutatása, melyek figyelembe vételével a jelenleginél természetesebb ember–gép interakciókat eredményező elméleti és technológiai modellek hozhatók létre.

Az ember–gép interakciók spontaneitásának és természetességének növelése érdekében olyan elméleti megfontolásokból érdemes kiindulni, melyek az ember–ember kommunikációt vizsgálják (Németh T. 2010). Ennek megfelelően azokat a mozzanatokat érdemes vizsgálni,

¹ Jelen cikk a TÁMOP-4.2.2-08/2008-0009 számú *Az ember-gép kommunikáció technológiájának elméleti alapjai* című pályázat támogatásával valósult meg. Köszönettel tartozom Abuczki Ágnesnek, Bársony Olgának, Hunyadi Lászlónak és Németh T. Enikőnek, akik értékes megjegyzéseikkel segítettek és segítenek kutatói munkám során.

melyek az ember–ember közti, úgynevezett személyközi kommunikációt jellemzik, s ezeket átültetni az ember–gép interakciós modellekbe. Azonban figyelembe kell venni azt is, hogy a gép (az emberi összetevővel ellentétben) mindössze korlátozott kontextuális, nyelvi, szemantikai, pragmatikai információ manipulálására képes, így szükségszerűen olyan korlátokat kell húzni az ember–gép interakciós modellek építése során, melyek legitim kihívásokat fogalmaznak meg a technika felé, mindeközben a megvalósíthatóságot nem veszélyeztetik (Hunyadi 2011). Tehát olyan elméleti megfontolásokat kell keresnünk, melyek segítségével a személyközi kommunikáció jellegzetességeit a technológiai modell számára érthetővé tudjuk tenni, és amelyek előreláthatólag kezelhetőek lesznek a technológiai megvalósítás során.

A fentiekből az következik, hogy a személyközi kommunikáció kutatásának azon irányait és eredményeit érdemes vizsgálatunk tárgyává tennünk, melyek a gép által is kezelhető alapegységekkel dolgoznak, és amelyek a produkció és a megértés (a szintézis és az analízis) folyamatait egyaránt vizsgálják.²

1 A fejlődéslélektani perspektíva

Az ember és gép közötti interakció természetesebbé tételében kiemelkedő szerepet játszhatnak a személyközi kommunikáció ontogenetikai vizsgálatai. Ennek oka kettős. A babák (a nyelvvelsajátítást megelőzően) kevés nyelvi, szemantikai információra támaszkodnak a mamával való kommunikáció során, mint ahogy a gépek is minimális verbális (szemantikai) információ földolgozására képesek az emberi fölhasználóval történő interakcióban. A másik ok pedig az, hogy pont a legkisebbek azok, akik a személyközi kommunikáció legtermészetesebb, legspontánabb módjainak használatával kommunikálnak, hiszen a kulturális, társadalmi, normatív aspektusok az egyedfejlődés későbbi szakaszaiban kerülnek csak előtérbe. Ezeknek megfelelően a babák és az ember–gép interakció gépi összetevői analóg módon összehasonlíthatók, s a kapott hasonlóságok fölhasználhatók az egyes ember–gép interakciós modellek létrehozásában.³

Ha követjük a Hunyadi László (2011) által javasolt modellépítési megfontolásokat, akkor olyan mozzanatokat kell keresnünk a kommunikatív képességek ontogenezisében, melyek kulcsként szolgálnak a babák számára az interpretáció (befogadás, analízis) oldaláról:

- az egyes kommunikatív események szegmentálásában (a kommunikáció kezdetének és végének megállapításában),
- a kommunikáció folyamatának strukturálásában,
- a kommunikatív szándék észlelésében és értelmezésében,

² A Hunyadi-modell mutat rá arra, hogy az analízis-szintézis egyetlen keretben történő kezelését a technikai megvalósíthatóság motiválja, mivel lehetővé teszi egyes viselkedés- és eseménytípusok automatikus fölismerését és kategorizációját (Hunyadi 2011).

³ A fejlődésanalógiák föltárása nemcsak elméleti, hanem módszertani szempontokat figyelembe véve is érdekes lehet. A kognitív pszichológiát és fejlődéslélektant ugyanazon statisztikai vs. szabály/megszorítás alapú tanulás vita formálta a nyolcvanas évektől, mint amely a formális vagy automatizált nyelvföldolgozó és –tanulórendszerek mögött húzódó kognitív architektúra modelljének mibenlétét is firtatja (l. Horgan & Tienson 1996, vagy a „The past and future of the past tense”-vita a Trends in Cognitive Sciences 2002-es hasábjain). De ezen módszertani analógia kifejtésére jelen dolgozat keretei között nincs lehetőség.

valamint a produkció (szintézis) oldaláról:

- a kommunikáció kezdeményezésében és lezárásában,
- valamint a már zajló kommunikációba történő becsatlakozásban.

E főnti mozzanatok természetesen nem választhatók le a csecsemőkori megismerés azon sajátosságairól, melyek a kommunikatív képességek kibontakozásán túl más kognitív tartományok ontogenezisét is meghatározzák, befolyásolják, hiszen ahhoz, hogy sikeres kommunikációt hajtsunk végre, általános észlelési és cselekvési képességeinket is latba kell vetnünk. Éppen ezért elsőként az emberi megismerés azon inherens tulajdonságaira kell összpontosítani, melyek az információmanipulációt⁴ alapvetően jellemzik az egyedfejlődés kezdeti szakaszaiban, s melyek megragadhatók, modellezhetők s előreláthatólag implementálhatók is lesznek az ember–gép interakciós technológiák kidolgozása során.

Ne gondoljuk azonban azt, hogy a babák és a gépek azért jó analogonjai egymásnak, mert a babák ugyanúgy rosszul észlelnek, keveset tudnak és keveset értenek a világ dolgaiból, mint a gépek. A fejlődéslélektanban a Piaget-féle pszichológia tényleg szolipszista gyermekképet tárt elénk az 1950-es években (s ezt megelőzően a freudi személyiséglélektan is) (l. Piaget 1952, 1954, Meltzoff 2002), azonban mind ez a gyermekkép, mind a vele járó lélektani metaforák a kognitív fejlődéslélektan fényében már történeti, nem pedig kurrens elméleti jelentőséggel bírnak.⁵ Hipotézisem szerint a babák és a gépek azért állíthatók analógiás viszonyba egymással, mert ugyanazt az alapproblémát kell megoldaniuk: föl kell építeniük azt a megismerési rendszert, mely segítségével interakcióba léphetnek a társas környezettel. Vagyis a feladatom az, hogy olyan mozzanatokot találjak a korai társas megismerésben és információmanipulációban, melyek ezen megfontolásokat figyelembe véve relevánsak lehetnek az ember–gép interakció gépi összetevőjének modellezése során.

2 Multimodalitás és korai társas megismerés

A korai társas megismerés kutatásában az 1970-es évektől kezdve egyre több olyan empirikus munka látott napvilágot, mely az újszülöttek társas percepció és produkciós képességeire hivatott evidenciákat szolgáltatni. Meltzoff és Moore például már 1977-ben kimutatta, hogy a 2–3. héttől kezdve a babák (motoros fejlettségüknek megfelelően) képesek utánozni bizonyos ujjmozgásokat, a száj nyitását és az ajakbiggyesztés különböző formáit (Meltzoff és Moore 1977). A nativizmus kognitív fejlődéslélektanban bevett szemlélete szerint a babák erre azért képesek, mert veleszületett képességük van arra, hogy megfeleltetéseket hozzanak létre (1) az általuk észlelt és végrehajtott cselekvések között (Meltzoff & Prinz 2002), (2) saját nyilvános megnyilatkozásaik és belső érzéseik, mentális állapotaik között, valamint (3) ezen megfeleltetéseket figyelembe véve következtetni tudjanak a másik ember belső állapotaira (Meltzoff 2002: 9).⁶ Ha ezt a nativista megfontolást követjük, akkor a korai utánzás okaként az intermo-

⁴ Jelentése itt: információkezelés.

⁵ A bevett metaforák közé tartozik például az ún. „burok-metafora”, mely az újszülöttet egy számára gyakorlatilag észlelhetetlen és érthetetlen világba helyezi, ahol a baba képtelen társas interakcióba lépni a környezetével, az őt és a környezetét elválasztó kognitív korlátok miatt.

⁶ Vegyük észre, hogy ebben a szemléletben a gyermeknek nem kell rendelkeznie semmiféle kifinomult elméletiával, sem pedig készen kapott innát modulokkal, mint ahogy a fodori (1983) modularitástézis feltételezné!

dális tanulást (Bahrck 1988) s annak nativista modelljét, az ún. aktív intermodális leképezés folyamatát nevezhetjük meg (a továbbiakban AIM) (Meltzoff & Moore 1997). Az AIM szerint a babák esetében az egyes mozgások regisztrációja szupramodális jellegű, a különböző észlelési modalitásokból származó információ fölvétele és földolgozása egy közös reprezentációs rendszerben történik.⁷ Ha pedig létezik ez a rendszer (s erre temérdek empirikus bizonyítékunk van), akkor feltételezhetjük, hogy a babák születésükkor már koordinált észlelési modalitásokkal rendelkeznek.

Az egyes észlelési modalitások közötti koordináció rendkívül fontos a babák számára, az alapvető társas tanulási folyamatok elképzelhetetlenek e koordináció nélkül. E koordináció segítségével például a baba életének 18. hetétől azt az arcot preferálja, amelyik beszél hozzá (MacKain és mtsai 1983, Kuhl & Meltzoff 1984). A gépi analógiára térve, az egyes modalitások közötti koordináció fontos lehet a gépi felismerés számára is, hiszen például egy interaktív reklám esetében a gépnek legelőször föl kell ismernie, hogy a fölhasználó kommunikációt kezdeményez vele, s ez a probléma nem pusztán technikai jellegű, ha természetes ember–gép interakcióban gondolkodunk, hiszen mind az analízis, mind a szintézis oldalát egyszerre kell figyelembe venni, s a különböző modalitású információkat szinergikus egységben kell kezelni (Hunyadi 2011). Például a modalitások koordinációja teszi lehetővé a szájról olvasást, melynek segítségével kiegészítjük a hallásos modalitásból származó hiányos információt (Kuhl & Meltzoff 1982). Ez a fejlődéslélektani eredmény szintén megtermékenyítőleg hathat például a spontán beszéd fölismerésének javítására.

Az egyes modalitások közötti koordináció biztosítja a redundanciát is a megismerésben, melynek több funkciója is releváns számunkra. A redundancia egyrészt megerősíti a motívátlan (önkényes) modalitásközi kapcsolatokat a perceptuális tanulás során (pl. az auditoros és vizuális tartományokban) (Slater és mtsai 1999, Gibson & Pick 2000, Bahrck 2004), másrészt lehetővé teszi multimodális események konzisztens és koherens észlelését különböző amodális relációkon keresztül (Bahrck & Lickliter 2002, 2004). Ez utóbbi azt jelenti, hogy a csecsemők olyan viszonyok észlelésére is képesek, melyek észlelési modalitástól függetlenül reprezentálódnak az őket körülvevő társas és fizikai világban. Ilyen viszonyok tekintjük az időbeli szinkroniát és átfedést, a tempót, a ritmust és az intenzitást (Lewkowicz 2000, Spelke és mtsai 1983). Ezt az amodalitást magunkon is tapasztalhatjuk, amikor például egy embert „intenzív” jellemnek tekintünk, s ez megmutatkozik a nem verbális akusztikus, valamint a vizuális modalitásokban is (Hunyadi, személyes közlés).

Az intermodális leképezés és koordináció, valamint a redundancia olyan tulajdonságai a csecsemőkori megismerésnek, melyek hasznosíthatók az ember–gép interakció modelljeiben, egyrészt azért, mert a koordináció a gépi megismerés számára is kezelhető információs szerkezeten alapszik (időbeli szinkronia, időbeli átfedés, tempó, ritmus, intenzitás), másrészt azért, mert segíti a különböző modalitásokban reprezentálódó események együttállásának

⁷ A közös reprezentációs rendszer meglétét nemcsak babákkal végzett empirikus kutatások támasztják alá. Doreen Kimura (1973a, 1973b) például kimutatta, hogy telefonálás közben azért gesztikulálunk, mert a beszéd és a gesztusok ugyanazon produkciós és interpretációs rendszerben reprezentálódnak, s ha az egyik modalitás produkcióját leállítjuk, a másik modalitás produkciója is megakad. Treffner és munkatársai (2008) kutatásokban amellettszolgáltattak evidenciát, hogy a gesztusok előkészítő fázisának produkciója anticipálja a beszédet, s ugyanígy a beszéd anticipálja a gesztusok előkészítését (ez utóbbi folyamatra jó példa Treffner és Barrett 2004-es kísérletsorozata, mely kimutatta hogy kihangosított mobiltelefon használatkor sokkal jobban lecsökken a vezetési teljesítményhez szükséges koordináció és kontrolltevékenységek minősége, mint kihangosítatlan mobiltelefonálás esetében, ugyanis kezünkben a telefonnal nem anticipálódik a gesztuskezdesi fázis, nem úgy, mint üres kézzel).

fölismerését (például nem verbális akusztikus és vizuális információk illesztését). Az intermodális leképezés és koordináció, az amodalitás és a redundancia olyan természetes és inherens tulajdonságai a korai megismerésnek, melyek lehetővé teszik a különböző események, így a kommunikatív események észlelését, szegmentálását, a kommunikáció folyamatának alapvető strukturálását. Éppen ezért a személyközi kommunikáció multimodalitását nemcsak az egyes modalitások összege adja, hanem ahogy Hunyadi (2011) is rámutat, a földolgozandó információ felszíni megjelenésén túlmenően olyan absztrakt szerkezeti és funkcionális mozzanatokot kell feltételeznünk a multimodális információfeldolgozásban, melyek a társas megismerés és ezen belül a kommunikáció lehetséges formáit szervezik.

3 Utánzás és kommunikáció

A reciprocitás az ember–gép interakciók egyik sarkalatos pontja, mivel ezek hiányában a kommunikáció spontaneitása és természetes folyása sérül. Már Bruner (1983) is megfigyelte, hogy a reciprocitás legegyszerűbb formái is mozgósítják a kommunikációt, mosolyra készítetik a babákat, s ez a kommunikatív igény, valamint a mosoly a babákkal kommunikáló felnőtteken is „tükröződik”.⁸ Az emberi fölhasználóra visszamosolygó avatar (humanoid arcberendezésű, grafikus gépi összetevő) minden bizonnyal ugyanígy fölkelti az ember érdeklődését, és meg-alapozhatja a természetes(nek tűnő) ember–gép interakciót is. A babák esetében a kölcsönöség és a reciprocitás az utánzás korai formáinak aktív használatával fejlődik (Meltzoff 2002). A babák kölcsönös viselkedéseinek (ún. reciprocitási játékaiknak) alapjául szintén a fönt már tárgyalt intermodális koordináció szolgál. E játékok során a baba és a mama ösztönösen utánozza egymás arckifejezését és nem verbális akusztikai megnyilvánulásait. Az utánzásos játék során a babák sajátos társas információ birtokába jutnak: „a másik ember is olyan, mint én” (Meltzoff 2002). E játékok eredményeként alakul ki a 14. hónapra az a fajta strukturális kongruencia, mely a baba–mama társalgás fordulóátvételi és kooperatív természetéért felel (Trevarthen 1980, Carpenter és mstai 1998). Az intermodális leképezések finomodása is elő-segíti azt, hogy a baba az általa észlelt (= a mama által produkált) és létrehozott (= a mama által észlelt) viselkedési formák között strukturális egyezéseket találjon. A kölcsönöség (interszubjektivitás) kibontakozásának eredményeként a babák sikeres diadikus prototársalgást képesek véghezvinni. Ezt a korai társalgást diadikusnak nevezzük, mert a baba és a mama a reciprocitási játékok során közvetlenül egymás szemébe néznek, tekintetük egymásra irányul, nem pedig egy harmadik, külső mozzanatra.⁹ A kommunikáció univerzuma ebben az időszakban a baba és a mama kölcsönös észlelését és figyelmét foglalja magába. Az a kíváncsiság, mely a mama tekintetének irányára helyeződik, majd csak később bontakozik ki, a mamával való strukturális kongruencia eredményeként. Így fognak összekapcsolódni az utánzásos tanulás és a tekintetirány követésének fejlődéses aspektusai.¹⁰

⁸ A tükröződést azért teszem idézőjelbe, mert itt a szót annak legnaivabb, leghétköznapiabb értelmében használok, nem pedig szakkifejezésként.

⁹ Jelen tanulmány lektora fölhívta a figyelmemet arra, hogy a diadikus terminus a fönti kontextusban sokkal szűkebb értelmezésű, mint ahogy a diadikusság általában a személyközi kommunikációban megjelenik.

¹⁰ Hobson (2002), valamint Tomasello és mstai (2004) rámutatnak arra, hogy a prototársalgások sikere nagyban függ az interakcióba vitt érzelmek arci kifejezésének kontingenciájától is. Ennek bővebb kifejtése azonban nem témája jelen dolgozatnak.

Az ember–gép interakció szempontjából azért lehet érdekes az utánzás korai formáinak vizsgálata, mert a mosoly és a reciprocitás mozgósítja a kommunikációt (fölvívhatja az emberi főhasználó figyelmét a gépi interakció lehetőségére például egy információs pultnál). Mindemellett a korai társas interakciók társalgási fordulóinak jobb megértése közelebb vihet minket a gépi összetevőt is tartalmazó társalgások szegmentálási lehetőségeinek föltérképezéséhez is, mivel a baba a számára elérhető perceptuális információk alapján képes kinyerni és értelmezni a társalgás megfelelő szegmentálásához szükséges kommunikációs markereket. Mindemellett azért is érdemes a korai társas interakciókra fókuszálnunk, mert az egyedfejlődés későbbi időszakaiban előtérbe kerülő magasszintű szándéktulajdonító viselkedések a fönnálló gyakorlati nehézségek miatt nem modellezhetők megfelelő alapossággal a gépi összetevő oldaláról (a már említett statikus és dinamikus kontextuális információk korlátozott elérhetősége miatt). A gépi összetevő által is megvalósítható tekintetirány-detekció és –követés fejlődéslélektani aspektusai azonban értékes adalékokat szolgáltatnak az ember–gép interakciós modellek szempontjából.

4 A tekintet szerepe a korai társas fejlődésben

Dolgozatom előző fejezetében a diadikus jellegű prototársalgásokat az ún. triadikus interakciók „előszobájaként” jellemeztem. A diadikus interakciók során a baba és a mama egymásra figyelve sajátos, közvetlen kapcsolatot létesít egymással (Farroni és mtsai 2002). A triadikus interakciókban azonban megjelennek a baba és a mama közvetlen, egymásra irányított figyelmén kívül elhelyezkedő dolgok is, s a kommunikációban résztvevő felek megpróbálják kitalálni, mi az a dolog, amelyre a másik néz (Tomasello és mtsai 2005). Ez a tipikus kognitív működésű felnőttek számára egyszerű feladat, a csecsemő és a gépek számára viszont kemény probléma. A csecsemőnek nemcsak azt kell „belőnie”, hogy merre néz a mama, hanem „ki kell találnia” tekintetének célját is. A triadikus interakciók tehát alapvetően a tekintet alapul vevő monitorozást követelik meg a babától és a mamától (valamint a gépektől), melyhez a tekintet irányának követése elengedhetetlen fontosságú tényező. Ezért dolgozatom ezen fejezetében olyan fejlődéslélektani megfontolásokat veszek sorra, melyek jó támpontokat adhatnak a gépi összetevőt is tartalmazó interakciók modellezéséhez.

A tekintetirány követéséhez vezető út első lépéseiről többféle fejlődéslélektani elképzelés született az elmúlt években. A dolgozatban kifejtendő analógia szempontjából azonban csak azok a fölfogások relevánsak, melyek teret engednek a viselkedésbeli markerek fölismerésének és interpretációjának, mivel a gépi összetevő is csak a számára érzékelhető, észlelhető stimulusokra hagyatkozhat a kommunikatív interakciókban való részvétele során. Ezért fontos kiemelni azt a megállapítást, hogy a tekintetnek van markere, a figyelemnek viszont nincs (Meltzoff & Brooks 2007: 237). A kommunikáció fókuszában elhelyezkedő dolgokra irányuló figyelmi komponens monitorozása a tekintetirány monitorozásának következménye, s a kettő egymástól elválasztható, elválasztandó.¹¹ A kommunikáció figyelmi komponensének (komponenseinek) monitorozása babák esetében csak a 18. hónaptól jelenik meg, addig a baba a viselkedésekben tetten érhető markerek észlelésére és interpretációjára hagyatkozik a

¹¹ Meltzoff és Brooks (2007) Magic Johnson kosárlabdázó becsapós passzaival szemlélteti a különbséget: Johnson arról volt ugyanis híres, hogy képes volt tekintetét és figyelmét teljesen szétválasztani, s rendkívül pontos passzokat adni társának úgy, hogy közben másfelé nézett.

kommunikáció fókuszának megállapításában. E fejezetben ezért ezekre a korai mozzanatokra összpontosítok.

A babák életük 10–11. hónapjára elegendő tapasztalatot halmoznak föl ahhoz, hogy megértsék a fejforgatás és a tekintetirány-váltás közti különbségeket (Meltzoff és Brooks 2001, 2007, Meltzoff 2005). Ebben az alig néhány hetes időablakban a babák megértik a szem, a tekintet és a látás kapcsolatait, s elkezdik követni a mama tekintetét. Ezt megelőzően a babák (életük 9. hónapjában) a testtartás és a fejmozgás változásait tekintik önmaguk orientációjának indikátoraként. Ebben a korai időszakban a kommunikatív és egyéb társas cselekvések megértése a célirányos fordulás megjelenésével kapcsolódik össze („mama felém fordul + beszél = hozzám beszél”), majd az összegyűlt tapasztalatok és tanulás hatására a 10. hónapra megjelenik a nézés megértésének képessége.¹² A 12. hónapban történik az az átmenet, mely során a baba rájön arra, hogy a nézésnek nemcsak iránya, hanem célja is van, s e cél egyrészt a társas interakció, a kommunikáció fókusza, másrésztől e célra ráirányítható a kommunikációban résztvevő felek tekintete is (például osztenzív rámutatással) (Gergely & Csibra 2005, Gergely 2007, Csibra & Gergely 2009).

A gépek számára ezen utolsó átmenet, a tekintet céljának megállapítása azért nehéz probléma, mert a gépnek először detektálnia kell a tekintet irányát (analitikus probléma), ezután követnie kell a tekintetet (analízist és szintézist, valamint e kettő koordinációját egyaránt megkövetelő probléma), végül ezekre alapozva be kell határolnia a tekintet tárgyát. Scassellati (2002) rámutat arra, hogy a robotok jelenleg nem is tudják megoldani a tekintet céljának megállapítását, mindössze a Butterworth (1991) által feltételezett ökológiai állapotig képesek eljutni: a tekintet céljának azt a tárgyat tulajdonítják, mely a másik fél tekintetirányának első „útba eső” tárgya.

Az ember–gép interakciók szempontjából a tekintet vizsgálata komoly szerephez juthat és kihívásokkal járhat a közeljövőben. A gépi látás, az analízis és a szintézis párhuzamos működésének fejlesztései lehetővé teszik azt, hogy a gépek is képesek legyenek arra a koordinációra, melyre a babák életük első életévére sajátítanak el. Noha a korai, diadikus, közvetlen prototársalgás jellegzetességei már most jól fölhasználhatók a gépi összetevőt is tartalmazó interakció modellezése során, az egyedfejlődés során később középpontba kerülő triadikus interakciók kiteljesedését jelző mozzanatok még nehézséget okoznak a gépek számára. A tekintet céljának detektálása fontos szerepet játszik a személyközi kommunikációban, s mind fejlődéslélektani, mind komputációs modellezési szempontból további kutatások szükségesek mechanizmusainak megfejtésére.

6 Összegzés: a babák és a gépek analógiájának szerepe a természetes, multimodális ember–gép interakció modelljeiben

Dolgozatom célja az volt, hogy fölvezoljam a korai társas fejlődés és a gépi összetevőt is tartalmazó interakciós lehetőségek közötti analóg viszonyokat. Éppen ezért elsősorban azokra az elméleti és kísérleti megfontolásokra támaszkodtam, melyek hasznos adalékokat szolgáltathatnak az ember–gép interakciós modellek számára. A rendkívül gazdag kutatási hagyományokkal rendelkező szakirodalomban a multimodalitás, az utánzásos tanulás és a tekintet korai szerepe központi jelentőséggel bír, mivel a babák számára ezek képezik a társas megisme-

¹² A nézés megértésének képessége az ún. „csukott szem” paradigmával és ennek változataival vizsgálható.

rés és a személyközi kommunikáció alapját. Ebből fakad az analógia helyénvalósága is a jelen projektben.

A számunkra érdekes és jelentős analógiás párhuzamok tehát az alábbiak:

- Az analízis és a szintézis egyetlen kognitív rendszer sajátja: az egyes cselekvések feldolgozása és produkciója egyetlen, az egyes észlelési modalitásoktól független reprezentációs rendszerben játszódik le (Hunyadi 2011).
- E reprezentációs rendszer magjához tartozik az időbeli szinkronia és átfedés, a tempó, a ritmus és az intenzitás modalitásfüggetlensége.
- Az analízis és a szintézis multimodális koordinációja lehetővé teszi a gesztusok, a mimika, az arckifejezések, valamint a prozódia utánzását.
- Az utánzásos tanulás eredményeként kibontakoznak a prototársalgások fordulóátvételi jellemzői. Mindez kölesönös affektív és interakciós markerek megjelenésével jár.
- A tekintet irányának kifelé irányulása a diadikus interakciókból nehézséget okoz a gépi összetevő számára (s kezdetben a babáknak is), ám a spontán, természetes társalgás során később mégis magától értetődőnek tekintjük a látás és a figyelem eltérő természetét.

E főnti párhuzamok továbbgondolása értékes adalékokat adhat a természetes, spontán ember-gép interakciós modell számára.

Irodalom

- Bahrck, L.E. (1988): Intermodal learning in infancy: learning on the basis of two kinds of invariant relations in audible and visible events. *Child Development* 59, 197-209.
- Bahrck, L.E. (2004): The development of perception in a multimodal environment. In: Bremner, G. & Slater, A. (eds.): *Theories of infant development*. Malden: Blackwell, 90-120.
- Bahrck, L.E. & Lickliter, R. (2004): Infants' perception of rhythm and tempo in unimodal and multimodal stimulation: a developmental test of the intersensory redundancy hypothesis. *Cognitive, Affective, & Behavioral Neuroscience* 4, 137-147.
- Bahrck, L.E. & Lickliter, R. (2002): Intersensory redundancy guides early perceptual and cognitive development. In: Kail, R. (ed.): *Advances in child development and behavior*. New York: Academic Press, 153-187.
- Bruner, J.S. (1983): *Child's talk: Learning to use language*. New York: Norton.
- Butterworth, G. (1991): The ontogeny and phylogeny of joint visual attention. In: Whiten, A. (ed.): *Natural theories of mind: Evolution, development, and simulation of everyday mindreading*. Oxford: Blackwell 223-232.
- Carpenter, M., Nagell, K. & Tomasello, M. (1998): Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development* 63.

- Csibra, G. & Gergely, G. (2009): Natural pedagogy. *Trends in Cognitive Sciences* 13, 148-153.
- Farroni, T., Csibra, G., Simion, F. & Johnson, M. (2002): Eye contact detection at birth. *Proceedings of the National Academy of Science* 99, 9602–9605.
- Fodor, J. (1983): *The modularity of mind*. Cambridge: MIT Press.
- Gergely, G. & Csibra, G. (2005): The social construction of the cultural mind: imitative learning as a mechanism of human pedagogy. *Interaction Studies* 6, 463-481.
- Gergely, G. (2007): The social construction of the subjective self: the role of affect-mirroring, markedness, and ostensive communication in self development. In: Mayes, L., Target, M. & Fonagy, P. (eds.): *Developmental science and psychoanalysis: integration and innovation*. London: Karnac, 45-82.
- Gibson, E.J. & Pick, A.D. (2000): *An ecological approach to perceptual learning and development*. New York: Oxford University Press.
- Hobson, P. (2002): *The cradle of thought*. London: Macmillan.
- Horgan, T. & Tienson, J. (1996): *Connectionism and the philosophy of psychology*. Cambridge: MIT Press.
- Hunyadi, L. (2011, megjelenés alatt): Multimodal human–computer interaction technologies. Theoretical modeling and application in speech processing. *Argumentum*.
- Kimura, D. (1973a): Manual activity during speaking – 1. Right handers. *Neuropsychologia* 11, 45-50.
- Kimura, D. (1973b): Manual activity during speaking – 2. Left handers. *Neuropsychologia* 11, 51-55.
- Kuhl, P.K. & Meltzoff, A.N. (1982): The bimodal perception of speech in infancy. *Science* 218, 1138–1141.
- Kuhl, P.K. & Meltzoff, A.N. (1984): The intermodal representation of speech in infants. *Infant Behavior and Development* 7, 361–381.
- Lewkowicz, D.J. (2000): The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin* 126, 281-308.
- MacKain, K., Studdert-Kennedy, M., Spieker, S. & Stern, D. (1983): Infant intermodal speech perception is a left-hemisphere function. *Science* 219, 1347–1349.
- McClelland, J.L. & Patterson, K. (2002): ‘Words or rules’ cannot exploit the regularity in exceptions. *Trends in Cognitive Sciences* 6, 464-465.
- McClelland, J.L. & Patterson, K. (2002): Rules or connection in past-tense inflections: what does the evidence rule out? *Trends in Cognitive Sciences* 6, 465-472.
- Meltzoff, A.N. (1999): Origins of theory of mind, cognition and communication. *Journal of Communication Disorders* 32, 251–69.

- Meltzoff, A.N. (2002): Imitation as a mechanism of social cognition: origins of empathy, theory of mind, and the representation of action. In: Goswami, U. (ed.): *Blackwell Handbook of Childhood Cognitive Development*. Blackwell: Oxford, 6-25.
- Meltzoff, A.N. (2005): Imitation and other minds: the „Like Me” hypothesis. In: Hilrley, S. & Chater, N. (eds.): *Perspectives on imitation: From neuroscience to social science*. Cambridge: MIT Press, 55-77.
- Meltzoff, A.N. & Prinz, W. (2002): *The imitative mind: development, evolution, and brain bases*. Cambridge: Cambridge University Press.
- Meltzoff, A.N. & Brooks, R. (2001): „Like me” as a building block for understanding other minds: bodily acts, attention, and intention. In: Malle, B.E., Moses, L.J. & Baldwin, R.A. (eds.): *Intentions and intentionality: foundations of social cognition*. Cambridge: MIT Press, 171-191.
- Meltzoff, A.N. & Brooks, R. (2007): Eyes wide shut: the importance of eyes in infant gaze-following and understanding others minds. In: Flom, R., Lee, K. & Muir, D. (eds.): *Gaze following: Its development and significance*. Mahwah: Erlbaum, 217-241.
- Meltzoff, A. & Moore, M.K. (1977): Imitation of facial and manual gestures by human neonates. *Science* 198, 75–78.
- Meltzoff, A.N. & Moore, M.K. (1997): Explaining facial imitation: a theoretical model. *Early Development and Parenting* 6, 179–192.
- Németh T., E. (2010): Results in the communication subproject of the HuComTech IKUT project in its first year and some future tasks. *An overview of the research activities of the HuComTech IKUT project in its first year*. Workshop, Debrecen, 24th of June 2010.
- Piaget, J. (1952): *The origins of intelligence in children*. New York: International Universities Press.
- Piaget, J. (1954): *The construction of reality in the child*. New York: Basic Books.
- Pinker, S. & Ullman, M.T. (2002): Combination and structure, not gradedness, is the issue. *Trends in Cognitive Sciences* 6, 465-472.
- Pinker, S. & Ullman, M.T. (2002): The past and future of the past tense. *Trends in Cognitive Sciences* 6, 456-463.
- Scassellati, B. (2002): Theory of mind for a humanoid robot. *Autonomous Robots* 12, 13-24.
- Seppänen, T. (2009): On research problems, demos and workplan. *HuComTech workshop* 8th October, 2009.
- Slater, A., Quinn, P.C., Brown, E. & Hayes, R. (1999): Intermodal perception at birth: intersensory redundancy guides newborn infants’ learning of arbitrary auditory-visual pairings. *Developmental Science* 2, 333-338.
- Spelke, E.S., Born, W. & Chu, F. (1983): Perception of moving, sounding objects by four-month-old infants. *Perception* 12, 719-732.

- Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. (2005): Understanding and sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences* 28. 675-735.
- Treffner, P.J. & Barrett, R. (2004): Hands-free mobile phone speech while driving degrades coordination and control. *Transportation Research, Part F: Traffic Psychology and Behaviour* 7, 229-246.
- Treffner, P., Peter, M. & Kleidon, M. (2008): Gestures and phases: the dynamics of speech-hand communication. *Ecological Psychology* 20, 32-64.
- Trevarthen, C. (1980): The foundations of intersubjectivity: development of interpersonal and cooperative understanding in infants. In: Olson, D.R. (ed.): *The social foundations of language and thought: Essays in honor of Jerome S. Bruner*. New York: Norton, 316-342.

Bódog Alexa
Debreceni Egyetem
Általános és Alkalmazott Nyelvészeti Tanszék
H-4010 Debrecen
Pf. 24
alexaweirdling@gmail.com