

# Változások a nyelvi mérésben Beszámoló az EALTA 2019. évi 16. konferenciájáról

Az EALTA (*European Association for Language Testing and Assessment*) a legjelentősebb nyelvi méréssel és értékeléssel foglalkozó európai szakmai szervezet, amelynek célja a nyelvi mérés elméleti alapvetéseinek és jó gyakorlatainak terjesztése, valamint a nyelvi tesztelés fejlődésének és modernizálásának elősegítése. Ennek legfontosabb terepe a 2004 óta évi rendszerességgel különböző országokban megtartott konferencia, ahol a nyelvi tesztelés első vonalából érkeznek az előadók Európa legkülönbözőbb országaiból, de többször voltak már nagynevű meghívott szakértők más kontinensekről is. Az idei évben a konferencia 2019. május 31. és június 2. között az írországi Dublinban a University College of Dublin szervezésében került megrendezésre.

A konferencia jellemzője, hogy egy tágabb és egy szűkebb értelemben vett programot is nyújt a résztvevőknek. A konkrét konferencia előtt ugyanis mindig szerveznek háromnapos, úgynevezett konferencia előtti workshopokat különböző méréselméleti-tesztelési témákról. Továbbá, mivel az EALTA számos speciális fókuszú csoporttal (SIG, *Special Interest Group*) is rendelkezik, szintén a konferenciát megelőzően tartják ezen csoportok találkozóit párhuzamos szekciókban, ahol több prezentáció is elhangzik. Ezt követi maga a két és fél napos konferencia, valamint ekkor tartják az EALTA évi rendes közgyűlését is.

A konferencia előtti workshopok közül az egyik a Facets program használatának rejtelseibe vezette be a résztvevőket, a másik az R statisztikai szoftver IRT-elemzésekben való használatáról, a harmadik pedig beszédképesség-értékelési skálák fejlesztéséről és alkalmazásáról szólt. A SIG-csoportok közül az osztálytermi értékeléssel, a Közös európai referenciakerettel (KER), a beszédképesség értékelésével, a migrációval és az egyetemi célú nyelvtudás értékelésével foglalkozó csoportok tartottak előadásokat.

A szoros értelemben vett konferencia fő témaként a nyelvi mérés és értékelés elméletében és gyakorlatában végbemenő változásokkal, illetve átalakulásokkal foglalkozott. Három plenáris előadás, három szimpózium, 24 prezentáció, 9 *work-in-progress* (kutatás közbeni) előadás és 13 poszterprezentáció szerepelt a menüben, amelyet kiegészített négy, úgynevezett „ebédidős beszélgetés” (*lunchtime sessions*).

A fő téma, a változások szerepe a nyelvi mérésben, illetve annak kutatásában az összes plenáris előadás összekötő eleme volt. Az első ilyen előadást a University College London szakembere, Talia Isaacs tartotta, akinek prezentációja a nyelvi mérés története során bekövetkezett néhány változást villantott fel, például azt a mozzanatot, amikor a pszichometriai fókusz átváltott a kevert módszerű kutatásokra, illetve kiemelte a kiejtésértékelés mai reneszánszát is, amely Robert Lado 1961-ben írt alapműve óta tetszhalott állapotban volt egészen a mostani időkig. Szintén a változásokat illusztrálандó, az előadás második felében Talia Isaacs részletesen bemutatott egy, a nevéhez köthető interdiszciplináris kutatási projektet, amelynek során nyelvi kisebbségekhez tartozó betegeket vontak be klinikai kísérletekbe, ami egyrészt nagyon innovatív volt a nyelvi mérést illetően, másrészt azonban túlságosan interdiszciplináris jellegű volt, ezért az erről a projektről szóló előadásterveket eleinte nem fogadták pozitívan a tesztelésről szóló konferenciák, mert nem érezték a konferencia profiljába illőnek. Később azonban már befogadták az erről szóló absztraktjait, amivel arra utalt az előadó, hogy leginkább a felhasználón múlik, hogy mi az a téma, amit a nyelvi mérés-értékelés birodalmába tartozónak tekinthetünk, és mi az, amit nem, azaz a változásokra hogyan tekintünk.

A második plenáris előadásban Detmar Meurers, a tübingeni egyetem professzora a számítógépes nyelvészet és az értékelés kapcsolatával foglalkozott, megemlítve, hogy a produktív vizsgateljesítmények (pl. esszék) automatizált értékelése úgy zajlik, hogy az elkészült írásműveket már korábban emberek által értékelt referenciaesszékkel hasonlítja össze a számítógép. Ezt továbbfejlesztve a jövőben a számítógépes nyelvi elemzés érdekes területe lehet a feladatjellemzők és a nyelvtanulói nyelvhasználat jellemzőinek közvetlen összekapcsolása, amelyből mind a nyelvelsajátítás kutatása, mind pedig a nyelvi mérés profitálhat.

A harmadik plenáris előadás során Michael T. Kane, a University of Wisconsin professzora a nyelvi mérés egyik kulcsfogalmának, a validáció értelmezésének változásaival foglalkozott, amely jól kapcsolódott az azt megelőző, validitásról szóló szimpóziumhoz. Mind az előadás mind pedig a szimpózium azt a témát járta körül, hogy a validitás elméleti szakemberei néhány alapelvben ugyan egyetértenek, azonban vita van a validitás hatáskörének kérdésében, illetve abban, hogy mely modellek azok, amelyek segítségével a tesztek validálni lehet. A szimpózium során ezekről fejtette ki véleményét többek között Luke Harding, Constant Leung, Barry O'Sullivan és Claudia Harsch.

A migrációval kapcsolatos kérdésekre a mai Európa különösen nagy figyelmet fordít, ami alól a nyelvtudásmérés világa sem kivétel, amit az is bizonyít, hogy a konferencia második szimpóziumát ennek a kérdésnek szentelték. Ezzel kapcsolatban az egyik kulcskérdés az, hogy a bevándorlók elé állított nyelvi követelmények meglétét miként ellenőrizzék, ugyanis az Európába érkező menekültek egyharmada nem rendelkezik optimális iskolázottsági háttérrel, illetve alacsony szintű az írni-olvasni tudásuk. Ebből következően a tesztelési szakma nincs felkészülve ezen csoport nyelvtudásszintjének

mérésére, ugyanis a nyelvvizsgálók túlnyomó többsége jóval magasabb fokú iskolázottsággal rendelkezik, mint a bevándorlók tömegei. A nyelvtudásmérési szakma ezen felkészületlensége pedig komoly etikai és validitási problémákat vet fel, hiszen, ha hiányosak az ismereteik a vizsgázói populációról, akkor számukra nem lehetséges adekvát tesztek kifejlesztése sem. Az olasz Lorenzo Rocca kimutatta, hogy nincs egységesség az európai országok között abban a kérdésben, hogy az állandó letelepedési engedély vagy az állampolgárság megszerzéséhez milyen KER-szintű nyelvtudás szükséges. A skála A2-től C2-ig nagyon vegyes, bár a legtöbb ország megelégszik a B1 szinttel. A belga Bart Deygers hozzászólásában arra mutatott rá, hogy az iskolázottsági háttér a legmeghatározóbb az állampolgársághoz szükséges nyelvvizsga sikeres abszolválásához. A norvég Cecilie Hamnes Carlsen pedig egyenesen odáig ment érzelmeiktől is átfűtött kissé provokatív prezentációjában, hogy szigorúan szakmai szempontból tekintve a nyelvi tesztek egész egyszerűen nem validak az állandó letelepedési engedély vagy az állampolgárság megszerzéséhez, mert egyrészt ezek a tesztek semmiféle validitásvizsgálaton nem estek át, másrészt pedig egy nyelvi teszt konstruktuma maga a nyelv, nem pedig a beilleszkedési folyamatra való hajlandóság, illetve képesség.

A harmadik szimpózium a C-tesztek változóban lévő felhasználási területeiről szólt. Kínai, török, osztrák és német előadók beszéltek arról, hogy a korábban elsősorban besorolási és diagnosztikus célokra használt C-teszteket ma már kutatási és szűrővizsgálati (*screening*) célokra is használják, utóbbit például bevándorlók nyelvtudásvizsgálatára.

A konferencia előadásai rengeteg témát öleltek fel. Mind a négy alapkészség méréséről hangzott el prezentáció, több előadó foglalkozott a KER-rel, illetve a közoktatásban és a felsőoktatásban alkalmazott nyelvi mérésekkel. Az alábbiakban ezek közül a számomra legérdekesebbeket emelem ki.

A beszédkészség értékelésével kapcsolatban Ching-Ni Hsieh az amerikai ETS-ben a TOEFL számítógépes vizsgájához kifejlesztett SpeechRater rendszer működéséről és validálásáról beszélt, amely a világ egyik legfejlettebb számítógépes eszköze spon-tán szóbeli megnyilatkozások értékeléséhez.

Franz Holzknacht a beszédértésvizsgákon használt hangzó szövegek egyszeri vagy kétszeri meghallgatásának lehetőségét vizsgálta abból a szempontból, hogy milyen módon változik meg így a teszt konstruktumának validitása. Az egyszeri meghallgatás több tesztmegoldó stratégiát váltott ki az alanyokból, míg a kétszeri meghallgatás során változatosabb beszédértési stratégiákat alkalmaztak, tehát nem csupán sikeresebben oldották meg a tesztet, hanem a teszt konstruktum validitása is nőtt ily módon.

Monique Reichert és Charlotte Kramer előadása azt mutatta ki, hogy a serdülő nyelvtanulók hagyományos és digitális környezetben történő olvasott szöveg értésére sincs szignifikáns hatással a digitális szövegek (e-mail, e-könyvek, közösségi média stb.) olvasása. A prezentációt követő vitában az is felmerült, hogy ez nem csoda, hiszen

a digitális olvasás egy teljesen más konstruktum, mint a hagyományos olvasás, tehát nem lehetett azt várni, hogy a digitális szövegek olvasása majd kihat a szövegértésre.

Philip Horne a British Council által kifejlesztett Aptis teszt kapcsán beszélt egy a kiejtés mérésével kapcsolatos esettanulmányról. Vizsgálatai során arra a következtetésre jutott, hogy az Aptis vizsgán használt holisztikus értékelési skála kiejtéssel kapcsolatos deskriptorai a KER kiegészítő kötetében (CEFR Companion Volume 2018) megjelent és alaposan átdolgozott fonológiai tudás skála alapján megújításra szorulnak, így lehet ugyanis továbbfinomítani a holisztikus értékelésen.

Elke Gilin és Lieve de Wachter nagyon érdekes belgiumi helyzetet ismertettek. A flandriai belga egyetemek többségére nem kell felvételizni, aminek az egyik következménye a nagyon magas, közel 60%-os lemorzsolódási arány az első évfolyam végén. Ez a könnyű egyetemre jutás azonban csak a holland anyanyelvű flamand belgákra vonatkozik, ugyanis, ha egy külföldi diák szeretne belga egyetemen tanulni, neki egy B2 szintű holland nyelvvizsgát kell tennie. A kutatók szerint ez a különbségtétel hazai és külföldi hallgatók között nem igazságos, amit azzal támasztottak alá, hogy kísérletképpen 50 holland anyanyelvű flamand belga diákkal is megírátták a holland nyelvvizsgatesztet, akik közül 16 fő nem érte el a sikeres felvételihez szükséges minimális pontszámot.

A poszterszekcióban Santi Lestari azt vizsgálta, hogy különböző értékelési skálák miként próbálják meg operacionalizálni, azaz reprezentálni, magukba foglalni az adott integrált olvasás-írás feladat konstruktumát. Ugyanebben a szekcióban nekem is volt szerencsém előadni, amelynek során a BME Nyelvvizsgaközpont online vizsgáztató-továbbképzésének tapasztalatairól számoltam be az elmúlt két év tükrében. A különböző nyelvek vizsgáztatóinak KER-értelmezését és szóbeli értékelői skála használatát elemeztem, illetve a hagyományos és az online továbbképzés előnyeit és hátrányait, illetve az ebből fakadó teendőket ismertettem.

Az úgynevezett ebédidős beszélgetések (*lunchtime sessions*) négy téma körül forogtak párhuzamos szekciókban. Az egyik a nemzetközi folyóiratokban való publikálás kulisszatitkait hozta közelebb a résztvevőkhöz, a második és a harmadik absztraktok írásához, illetve prezentációkészítéshez nyújtott hasznos tanácsokat, míg a negyedik egy ismerkedési workshop volt az EALTA-konferencián első alkalommal részt vevők számára.

Összességében a konferencia nagyon hasznos volt, számos érdekes témát boncolgatott, és a résztvevők ízelítőt kaphattak azokból a kutatásokból, amelyek az európai nyelvi tesztelés szakembereit jelenleg foglalkoztatják. A konferencia előadásait kötetben sajnos nem szokták megjelentetni, azonban az EALTA honlapján általában közzéteszik az összes előadás és poszter diáit, ami bizonyára az idei évben is így lesz. A következő évi konferenciának pedig Budapest fog helyt adni, tehát érdemes lesz kihasználni a magyar szakembereknek ezt a nagyszerű lehetőséget előadásokkal vagy akár csak részvétellel is.

Kiszely Zoltán