
TARTALOM (TABLE OF CONTENTS)

EINHORN ÁGNES: Mérés és értékelés - múlt, jelen jövő. Előszó
(Measurement and assessment - past, present and future. Introduction).....3

TANULMÁNYOK (ARTICLES)

A nyelvtudásmérés múltja és jövője

BÁRDOS JENŐ: A magyar Állami Nyelvvizsga öröksége: (ne) nézz vissza haraggal?
(The legacy of Hungarian State Language Examinations: (don't) look back in anger?).....5

BENKE ESZTER: Egy siKERTörténet: mit kapott az oktatáspolitikai és a nyelvoktatás
a Közös európai referenciakerettől?
(A success story: how has the Common European Framework of Reference for Languages
informed education policy and language instruction?).....30

LUKÁCSI ZOLTÁN: Új utak a nyelvi mérésben
(New trends in language testing).....45

TANKÓ GYULA: A validálási folyamat érvelésalapú megközelítései áttekintése
(An overview of argument-based assessment validation).....65

A vizsgáztatás problémái

CSIZÉR KATA - ÖVEGES ENIKŐ: Idegennyelv-tanulási motivációs tényezők és a nyelvi vizsgák
Magyarországon: összefüggések vizsgálata egy kérdőíves kutatás segítségével
(The relationship between foreign language learning motivation and language exams:
the results of a questionnaire study).....86

SZABÓ GÁBOR: A szövegnehézség vizsgálata az emelt szintű angol érettségi olvasáskomponensében
(Examining text difficulty in the reading component of the advanced
level school-leaving exam in English).....102

KISZELY ZOLTÁN: Értékelési skálák használata a beszédkészség és íráskészség vizsgákon
(Rating scales in speaking and writing performance assessment in high-stakes examinations).....120

Az osztálytermi értékelés

CSÉPES ILDIKÓ: Hogyan fejleszthető a tanárok nyelvtudásmérési és értékelési műveltsége?
Az Erasmus+ TALE projekt tanulságai
(How to facilitate language teachers' assessment literacy?
Lessons from the Erasmus+ TALE Project).....136

VÍGH TIBOR – KISS-KOVÁCS RENÁTA: Hogyan értelmezik a tanárszakosok a tanulói értékelés fogalmát? (How do teacher trainees interpret student assessment?).....	156
--	-----

MŰHELYTITKOK (WORKSHOP)

Műhelytitkok: a fejlesztő értékelés gyakorlata

EINHORN ÁGNES: Fejlesztő értékelés a felsőoktatásban - módszertani mozaik (Formative assessment in higher education - a methodological review).....	174
--	-----

ERDEINÉ BACHER KATALIN: A fejlesztő értékelés a nyelvórán - példák a szakgimnáziumból (Formative assessment in the language class examples from a vocational secondary school)....	191
---	-----

VARGA ZOLTÁN BALÁZS – SALAMON ANIKÓ: A fejlesztő értékelés gyakorlata a nyelvoktatásban egy szakképzési centrumban (Language teachers' practice of assessment for learning in a VET center).....	202
---	-----

PROJEKTSZEMLE (PROJECT REVIEW)

SZABÓ CSILLA: Az eTransFair projekt portáljai: értékelési technikák és módszertani ötletek fordítást oktató szakemberek számára.....	215
--	-----

HÍREK (NEWS)

KISZELY ZOLTÁN: Változások a nyelvi mérésben. Beszámoló az EALTA 2019. évi 16. konferenciájáról.....	221
--	-----

KISS-KOVÁCS RENÁTA: Beszámoló a XVII. Pedagógiai Értékelési Konferenciáról (PÉK 2019). Szeged, 2019. április 11-13.....	225
---	-----

KIEGÉSZÍTÉS.....	228
------------------	-----

Mérés és értékelés - múlt, jelen jövő

Előszó

A folyóirat e tematikus számában a mérés és értékelés néhány aktuális kérdését járjuk körül, a szerzők a témához különböző szempontokból kapcsolódnak: megjelennek a méréselmélet különböző megközelítésmódjai, a nyelvvizsgáztatás és az érettségi vizsga problémaköre, illetve az osztálytermi mérés szempontjai is. A tanulmányok közül néhány a múlttal, tehát a gyökerekkel foglalkozik, több szöveg az empirikus kutatásokból leszűrhető adatok alapján a jelent elemzi, és természetesen a legtöbb szövegben megjelenik a jövő, a további fejlesztések és innovációk lehetősége is.

Az első írás Bárdos Jenő vállaltan szubjektív esszéje, amelyben a szerző a nyelvtudásmérés magyarországi történetének egy fontos szakaszára emlékezik vissza: a „Rigó utcai” nyelvvizsgára. Reméljük, hogy e személyes hangú visszaemlékezés további szakmai vitát indít el, hiszen a magyarországi nyelvtanítás jelenlegi helyzete szempontjából fontos lehet ennek a kornak az árnyalt recepciója.

A mérés-értékelés szakterület szempontjából meghatározó a KER fejlesztése, ezzel foglalkozik Benke Eszter írása. E dokumentum alapjává vált a nyelvpolitikai intézkedéseknek és a nyelvtanulás szabályozásának, ugyanakkor már megjelenése előtt elkezdődött a szöveg szakmai vitája. Ennek eredménye egy alapos és átgondolt revízió lett, melyről Lukácsi Zoltán számol be. Ez az írás néhány további aktuális kérdéssel is foglalkozik: a globalizáció és a nyelvtudásmérés kapcsolatával, illetve a nyelvtudásmérés technológiai fejlődésének lehetőségeivel. Az elméleti keretekkel foglalkozó tanulmányok sorát Tankó Gyula írása zárja, amelyben a szerző bemutatja az érvelésen alapuló validálási modellt.

Ezt követően a vizsgáztatáshoz kapcsolódó több empirikus kutatás eredményéről is olvashatnak. Csizér Kata és Öveges Enikő egy kérdőíves kutatás keretében vizsgálta a nyelvvizsgákhoz való viszonyulás és a tanulási motiváció kapcsolatát. S bár a kis mintás kutatás eredményei nem általánosíthatók egyértelműen, mindenképpen kérdésessé teszik azt a sokak számára nyilvánvalónak tűnő gondolatot, hogy a nyelvvizsgák egyértelműen erősítik a nyelvtanulási motivációt. Szabó Gábor az érettségi vizsgához felhasznált szövegek elemzését végezte el, tanulságos kísérletet adva arra, hogy objektív mutatók alapján határozzuk meg a szövegértési feladatokhoz felhasznált szövegek

nehézségét. Kiszely Zoltán egy vizsgáztató-továbbképzés adatait felhasználva elemzi, hogy a különböző értékelési skálákat használó vizsgáztatók értékítéleteiben milyen szabályszerűségek figyelhetők meg.

A mérés-értékelés szakterület múltja és jelene után a jövő felé fordulunk. A területen sok fontos innováció, komoly szemléleti változások figyelhetők meg, ám a modernizáció az osztálytermi értékelés gyakorlatában csak komoly megszorítással valósul meg. Ehhez kapcsolódóan két empirikus kutatás eredményei olvashatók: Csépes Ildikó a gyakorló tanárok értékelési műveltségét, Vígh Tibor és Kiss-Kovács Renáta pedig a tanár szakos hallgatók értékelésről alkotott képét vizsgálja.

A Műhelytitkok rovat a fejlesztő értékelés gyakorlatához köthető szövegeket tartalmaz: jó gyakorlatokat olvashatnak a felsőoktatásból és egy szakköznevelésből, végül egy szakképzési centrum továbbképzésének hatásait ismerhetik meg.

A tematikus szám írásait szakmai párbeszéd indításának is szánjuk: a folyóirat későbbi számaiban örömmel látunk olyan szövegeket, amelyek ezekre az írásokra reflektálnak.

Einhorn Ágnes
a tematikus szám szerkesztője

BÁRDOS JENŐ

A magyar Állami Nyelvvizsga öröksége (ne) nézz vissza haraggal?

(Nyelvpedagógiai esszé)

The legacy of Hungarian State Language Examinations: (don't) look back in anger?

This investigation into more than three-decades of HuSLEs is a pioneering study on the recent history of FLT and assessment in Hungary. The text written in the form of an essay pinpoints less obvious facts relating to the only public language examination system of the past century in Hungary which are the mere tip of the iceberg. The author's intention was to demonstrate hidden links between the changes in contemporary LT methods of the time and metamorphoses in the constructs of HuSLEs. The hard work of the members of the responsible institute (the FL Further Development Centre of ELTE University, Budapest) made it possible to continuously develop new variations of HuSLEs, roughly at the end of each decade, labelled tentatively as the 'Ancient' (70s), the 'New' (80s) and the 'Consolidated' (90s). Of the three variants, the HuSLEs of the 80s were examined in the greatest detail: the structure of the exam, examination techniques, error correction, examiner attitudes, etc. The real value of this independently developed FL examination system was its behind-the-scenes human resources: a team of teachers whose strengths were in teaching; designing and renewing exams and teaching materials; and carrying out research in cooperation with one another. It is not examinations but rather people who are so difficult to change.

Keywords: history of Hungarian State Language Examinations (HuSLEs), variants of HuSLEs, glimpses of EFLT methodology in Hungary, methodology of FL examination techniques, quality control and assurance, evaluation and assessment in some aspects of HuSLEs

I. Captatio

A minap hálózati sebességgel terjedt az elektronikus médiában a „Hogyan tudja egyetlen nyelvvizsga romba dönteni az egész nyelvvizsgarendszer hitelességét?” című írás (angol.info). Második bekezdésében egy mondat: „Amíg Magyarországon csak egyetlen nyelvvizsga létezett, nagyon alacsony volt a megbecsültsége – valójában rosszabb volt a híre, mint amit megérdemelt, de ez más kérdés.” Annyi bizonyos, hogy ez a tanulmány tovább lappangott volna a meg nem írtak között, ha nincs az a készlet, amelyet az előző mondatban a „más kérdés” kifejezés váltott ki bennem. Nem más kérdést, hanem egy MÁSIK, sokkal fontosabb kérdést kell feltennünk: hogyan tudta egyetlen nyelvvizsga mégis évtizedeken át megőrizni a hitelességét? Valóban csak azért, mert „egyetlen” volt? Érdemes-e a hőskort felidézni, vagy csak nappal gyűjt lámpást, ki a múltat idézi?

Nincs rá példa a múlt század utolsó évtizedeiben, hogy egy egyetem valamely apró egysége olyan elementáris, országos hatást, hírnevet, sőt hírhedséget váltott ki vizsgáztatási tevékenységével), mint az ELTE Idegennyelvi Továbbképző Központja. Ennek egyik fő oka az a tény, hogy tájékoztatóiban, később vizsgaanyagok publikálásával, az eljárások menete, követelményei, sőt módszerei is nyilvánossá váltak (Bárdos 1982, 1983). Olyan volt ez, mintha valaki kidugná a fejét a lövészárokból, hogy ide lőjtek! Ez később be is következett (lásd Dörnyei–N. Tóth 1987, és rá a felelet: Heltai 1989), bár ugyanezekben az években, a fenti vitákkal egy időben, az Állami Nyelvvizsga ismét újjászületett.

Az elementáris hatás kiváltó oka egyszerű: ez a vizsga *publikus* volt: az egész országban mindenkire egyformán vonatkozott. A bizonyítvány ugyan egyetemi végzettséget nem igazolt, mégis, birtoklása tekintélyt parancsolt. Muszáj ugyanakkor, a sohasem kívánt párhuzam miatt, feltenni azt a kérdést, hogy manapság nyilvános-e, tudja-e pontosan valaki, még ha szakmabeli is, hogy milyen megmérettetéseken át szerez valaki magyar felsőoktatási intézményben diplomát? Sűrű homály fedí emez írásbeli és szóbeli „megmérettetések” mibenlétét, módszereit: ilyen intim kérdéseket nem feszeget senki, aki nem akar vihart aratni. Gyanítható, hogy a pedagógiai mérések világától igen távol eső, ösztönös, szakmailag tudatlan, bár egyáltalán nem szándéktalan megoldások „selyemútján” jut el a jelölt a diplomáig. Ezek szerint a nyelvvizsgák, mint piaci és főként publikus jelenségek, alighanem e kultikus és rituális eljárások „celebjei”, hiszen a nyak.hu honlapján úgy tanulmányozhatók, mint az őshüllők csontjai a British Museumban. (E megkerült gondolatmenet folytatása már csak az lehetne, hogy diplomát, doktorit, sőt egyetemi felvételit is csak nyelvvizsgával szerezhet az ember, hiszen a többi vizsga olyan megbízhatatlan.)

Mielőtt elragadna ez a rémisztő gondolatmenet, a hagyományok jegyében visszatérek a tanulmány megírásának céljához. Azt az időszakot kívánom szakmai és egyéb szempontból elemezni és feleleveníteni, amely a kezdetektől a magyarországi állami nyelvvizsgáztatást jellemzi az ezredfordulóig (1967–2000), merthogy addig létezett. (A nyelvi vállalkozások intézményesülése kényszerítette ki 2000 után azt, hogy az egymással versengő honi és külföldi nyelvvizsgák korszerű szakmai alapokon működjenek: közösen elfogadott standardok alapján. Pár év múlva a Közös európai referenciakeret (CEFR 2001, KER 2002) ismertté válásával a csereszabatosítás olyan sorozatai kezdődtek el, amelyek egyfelől lehűtötték a vizsgaszolgáltatók vérmes kapitalista reményeit, másfelől a beválás alapján szelekcióra voltak képesek.)

Az Állami Nyelvvizsga (és az ELTE Idegennyelvi Továbbképző Központ) több mint három évtizedes történetét egyedi, sajátos, közép-európai, szakmai és emberi törtéetésnek tartom, amelynek valós tényei: a szakmai és emberi erőfeszítések ismerete hiányzik a magyar pedagógiai (mérésmetodikai) köztudatból. Ezért érdemes részletesen

foglalkozni az ÁNYV történetével, vizsgamódszertanával, vizsgafejlesztéseivel, minőségbiztosításával és a korabeli nyelvtanításra gyakorolt hatásával.

A háború utáni első honi nyelvvizsgák a semmiből jöttek: kiváló magyar nyelvtanárok manufaktúrájában váltak kipróbált munkadarabbá. Nem támogatta őket bőséges forrásokkal és nemzetközi gyakorlattal rendelkező, szakértőkkel kibélelt kutatóintézet vagy vizsgaügynökség (mint ahogyan azt az ETS tette saját vizsgáival az Egyesült Államokban, a CITO Hollandiában vagy éppen a UCLES Angliában). Nem volt lehetőség, idő és talán matematikai-statisztikai szaktudás sem (főként nem szoftverek) a teljes körű empirikus vizsgálatokra. Volt viszont többrétű visszacsatolás, amely lehetővé tette a beválást – az egyéni és a közösségi tapasztalatok alapján. Ösztönösség és szándék is hozhat ragyogó eredményeket hozzáértők kezén, feltéve hogy megvan a megfelelő szaktudás (és, esetünkben, a kiemelkedő nyelvtudás mint hozzáadott érték). Ezek a tudós tanárok magas óraszámban tanítottak; vizsgáztattak; nyelvkönyveket és tankönyveket írtak; vizsgát fejlesztettek, és tudományos publikációkat tettek közzé: szakmai kézjegyük mindenütt fellelhető a korabeli idegennyelv-tanítás egészében. Mindezen felül bátrak voltak, mert a lehetetlenre vállalkoztak: a MÁSIK (ember) nyelvtudásának megítélésére.

II. Az ELTE ITK és az Állami Nyelvvizsgák rövid története

A mai olvasónak szól ez a kis bevezető, miszerint az orosz nyelv kötelezővé tétele (1949) után az úgynevezett „nyugati” idegen nyelvek jóformán illegalitásba kényszerültek: iskolai szakkörökbe, szakszervezeti és TIT-tanfolyamokba, egyetemek rejtett tanszékeinek védelmébe; oroszra át nem képzett nyelvtanárok magánórái közé. Létezett ugyan az Idegen Nyelvek Főiskolája (1951–1956), amely leginkább a semlegesnek tűnő franciát preferálta, de a bejutás ide „ellenőrzött és kézben tartott volt”. A forradalom után az intézet megszűnt, de az ELTE-n belül egy Idegen Nyelvtanfolyam és Nyelvvizsga Csoport már rendelkezett nyelvvizsgáztatási jogokkal. Egyetlen évtized alatt jelentősen változott a magyar gazdaság helyzete: a KGST-n kívüli piacok is érdekessé váltak. Az új gazdasági mechanizmust (1968) követően fokozódott az igény (külkereskedelem, ösztöndíjak stb.) az idegen nyelvi diverzitás felélesztésére. A fiatalabb generációk számára világossá vált, hogy az orosz nem lett világméretű közvetítő nyelv, sem a turizmus nyelve; hogy honi elfogadottsága és presztízse csekély; „hogy az orosz nyelv tanításába intézményesen beépített igénytelenség a belenyugvásnak egy olyan társadalmi gyakorlata, amely nem ostromolja a nyelvtudásnak azokat a standardjait, ahol a nyelv mint eszköz egy kultúra közvetítőjévé válhat” (Bárdos 1988: 169).

Szinkronban volt tehát a történésekkel az MM 6/1965-ös rendelete és a rendeletet kiegészítő utasítás (3/1967), amely létrehozta az ELTE felügyelete alá tartozó Idegennyelvi Továbbképző Központot, és megbízta az intézményt a felső- és a középfokú nyelvtudást igazoló vizsgák szervezésével és lebonyolításával. (Egy korábbi rendelet,

a 110/1964 már intézkedett a nyelvtudás anyagi elismeréséről.) A vizsgára iskolai végzettségtől függetlenül bárki jelentkezhett, aki nagykorú. Szakmai nyelvvizsgát csak akkor lehetett tenni, ha a releváns minisztérium gondoskodott megfelelő szaktudású vizsgáztatókról. Gyakran feledésbe merül, hogy az alapítás célja „a dolgozó felnőttek idegen nyelvi oktatásának és továbbképzésének intézményes biztosítása” volt, vagyis nyelvvizsgára előkészítő tanfolyamokat, vállalati céltanfolyamokat kellett szervezni, és mintegy azon túl, a nyelvvizsgáztatást is ellátni. Az ITK belső oktatói és vizsgáztatói gárdája ideálisan 30-40 fő körül mozgott, ám a későbbiekben a külső vizsgáztatókkal az Állami Nyelvvizsga Bizottság több száz főre dagadt, ami súlyos minőségbiztosítási gondokat okozott.

Az intézmény első igazgatója a fiatalon elhunyt Serény Andor volt, akit a francia kultúrán kívül a nyelvtanítás megújítása és az AV-technológiák is érdekelték. Olyan híres tankönyvszerzők vizsgáztattak például angolból, mint Véges István, Tarján Jenő, Korenchy Lajos, Jakabfi László, Székács Györgyné, Seprődi László és mások. Az intézet dolgozói mindnyájan, beosztásuktól függetlenül, keményen húzták az evezőket: feltelítette a napokat a több helyszínű, intenzív nyelvoktatás, a vizsgáztatás, nemkülönben a vizsgák javítása és ellenőrzése. Kétségkívül kutatásra és publikálásra maradt a legkevesebb idő: ez a jelenség a nyelvtanári lét örökzöldjei közé tartozik. A vizsgafejlesztés viszont, mint majd látjuk, a legfontosabb tevékenységgé lépett elő.

Köztudott, hogy a nyelvtanítási módszerek, valamint a társadalmi igények az utóbbi másfél évszázadban világszerte létrehozták a saját nyelvtudás-elképzeléseiket. Ezeket a képzeteket fordították le, ezeket realizálták az iskolai, majd kivetítve, a külső nyelvvizsgákban is. A nyelvtudás, mint a vizsgák célképzete, történeti kategória (vö. Bárdos 2002: 87–104). Távolról sem meglepő tehát, hogy az ELTE ITK is 10-12 évenként megújította vizsgái struktúráját; vizsgatechnikáit; a vizsgáztatókkal és a vizsgafejlesztőkkel szembeni elvárásait, hiszen maga a nyelvtanítás és a nyelvtanulás is, a generációváltások sebességével frissült, élénkült és alakult át. A következő táblázatban az ÁNYV fejlődésének stációit jelenítjük meg, amelyekről a továbbiakban részletesen szólunk.

A bemutatás alapja a továbbiakban nem személyek és évszámok, hanem a tanítás-tanulás, valamint a mérés-értékelés párhuzamos vonulatainak ábrázolása; nyelvtanítás-történeti reprodukció, amely azt kívánja bemutatni, hogy milyen szakmai megfontolások alakították ki az Állami Nyelvvizsgát mint szellemi konstruktumot a korabeli bürokrácia és előítéletesség rideg medrében. Az örökbecsű mondás egyébként ott és akkor is érvényes volt: kellett egy csapat.

Évek	Igazgatók	Vizsgák beállása (a négy korszak)	Jellegzetességek	Vizgszámok (csak megközelítések a növekedés bemutatására)
1951–57 1957–67	ELTE Idegen Nyelvek Főiskolájának megszűnése után	A vizsgáztatási jog az ELTE Idegen Nyelvtanfolyam és Nyelvvizsga Csoport kezébe került		évente 50 körül
1967–1975	Dr. Serény Andor	1969: mindkét irányú fordításokat preferáló, régi típusú nyelvvizsga. Előbb az írásbeli, utána szóbeli. Elégtelen írásbeli után nincs szóbeli.	ELTE ITK: 6 nyelv, 15 főállás. 3/1967 MM utasítás; előtte 6/1965 (VII. 16.) MM rendelet	5591
1975–1987	Dr. Sipőczy Győző	1979: új típusú, beszédcentrikus nyelvvizsga. Magnetofonos hallásértés. Szóbeli és írásbeli jegyek aránya 5:2. Szóbeli és írásbeli nem függ egymástól.	9/1978 (VII. 25.) OM vizsgarend. Kétnyelvű, általános és szakmai nyelvvizsgák. Vidéki egyetemi centrumokban is: Veszprém, Miskolc, Szeged, Pécs, Debrecen.	11 800
1987–1997	Dr. Horváth Iván	1989: modul rendszerű nyelvvizsga. Képleírás, szerepjátszási feladatok. Írásbeli és szóbeli külön-külön is letehető.	ELTE ITK: 40 főállás; vizsga 52 nyelven> összes vizsgáztató: közel hétszáz (!)	29 600 35 000
1997–2015	Dr. Gáborján Lászlóné	1999–2000: kétnyelvű, modul rendszerű, kritériumorientált. A NYAT által elsőként akkreditált, államilag elfogadott nyelvvizsga: az Origó (2000)	Nemzetközi szervezetek vizsgálták az intézményt és vizsgáit. Előtesztelt, statisztikai módszerekkel validált nyelvvizsga bevezetése.	101 796 146 000

1. táblázat. Az ELTE ITK és nyelvvizsgái

III. Az Állami Nyelvvizsgák metamorfózisai

Nincsenek kétségeim, hogy emlékek és történések felelevenítése mind a szakmában, mind a nagyközönség körében vegyes érzelmeket kelt. A nyelvtudás hiányát szokás időről időre felróni genetikánk jellegzetességeként: „más nép, másfajta raj, másként tapad a fején a haj”; meg hogy „nehezebb eltörni a faragatlan fát” és így tovább. Ezek a mélyből fakadó, költői telitalálatok már nem felelnek meg a mai statisztikai mutatóknak, hiszen bizonyítható, hogy Terestyéni (1980) hírhedt felmérése óta sokkal több magyar tanult meg idegen nyelveket (Medgyes 2011 és 2015) valahol, valahogyan. Ugyanakkor a nyelvtudás elavulhat, miközben az egykori bizonyítványok soha nem avulnak el.

Magyarország a második világháború végére egy több évszázada regnáló, birodalmi nyelvet veszített el az iskolákban: a németet. Az 1949-ben kötelező első nyelvként elfogadott orosz szintén birodalmi nyelv lett, ám a politikai ideológia és átnevelés hangsúlyozása elsápasztotta és elapasztotta azt a rendkívül gazdag irodalmi, művészeti és folklórhagyományt, amelynek ez a nyelv is hordozója. Kik tanították ezt az „új” nyelvet? Sem az idősebb, sem a fiatalabb nyelvtanár-generációkban nem voltak képzett szlavisták. Természetesen „átképzett” vagy éppenséggel saját magukat újra kiképző görög–latin–szanszkrit; angol, német, francia, úgynevezett „nyugati” nyelvszakos kollégák, akik főként általános nyelvészeti tudásukra, műveltségükre, pedagógiai tapasztalatukra és intelligenciájukra támaszkodva voltak képesek váltani. Akadt persze néhány, a háború és a megszállás következményeként itt megkötött orosz anyanyelvű feleség is. Szűkebb témánk szempontjából fontosabb, hogy milyen szak módszertani tudás és oktatástechnikai gyakorlat reprodukálható azokban az években, amikor a legelső, írásbeli fordításokon alapuló nyelvvizsgák stabilizálódtak.

Bár a huszadik század elején a magyarok élénken bekapcsolódtak a direkt módszerek (Phillip 1911), vagy éppen Gouin ígeközpontú mondatokkal cselekedtető, memorizáltató módszereinek művelésébe, ezek az akkor modern módszerek hamar feledésbe merültek. A két világháború közötti ismert módszertanok már inkább (és újból) a grammatizáló módszereket emlegetik, akár kontrasztív, akár tudatosító-egybevető, akár memorizáltató és transzformációs változataikban. Felemlítik továbbá az olvastató módszert is. A korabeli bölcsészképzés tudósképzés volt, amelynek középpontjában irodalmi, nyelvészeti és művelődéstörténeti tanulmányok álltak, így nem csoda, ha a közoktatásban is a leíró nyelvtan direkt tanításával, a nyelvtani-fordító módszer különféle változataival igyekeztek a nyelveket megismertetni. Így a tanítás ugyan *a nyelvről* szólt, de a komparatív grammatika égisz alatt eltakarta az erdőt. Bár a nyelvtani-fordító módszer azt ígéri, hogy a fordítás mint technika segítségével az olvasási készségeket fejleszti (eredetileg *auctorok* olvasása), ez a távoli cél a módszer inflálódásával szabályok magolásává, kivételek bevésésévé, deklinációk és konjugációk kántálásává silányult (Bárdos 2005: 45–56). A nyelvről szóló nyelvtani tudások (és tiltások) pókhálója rátelepedett a tanulók agyára: a homo grammaticusok

diktatúrája ez, amely túl sokáig tartott. Az volt az elvárás, hogy a tanuló a megtanult nyelvi szabályok alkalmazásával helyes mondatokat gyártson, esetleg fordításokat készítsen. Ezeknek a készségeknek a fejlesztése gyakran mondatgyakorlók segítségével történt, amelyek előszeretettel fókuszáltak a két nyelv eltéréseire, illetve nyelvtani alapú trükkjeire. Olvasás esetén a szöveg nyelvtani elemzése is elvárás volt, gyakran olyan szinten és olyan terminológiával, amelyre a tanuló anyanyelvén is képtelen lett volna. Az igen intenzív szótanuláson túlmenően (szódolgozatok) a rendszer mai időnkben is érvényes, sőt kívánatos egyetlen előnye volt az eredeti szövegek (dalok, versek stb.) memoriterként kezelése és számonkérése. A félelem a hibás megszólalástól, valamint az unalom a véget nem érő nyelvtani gyakorlatoktól (ragozások, kiegészítések, transzformációk, előljárószók behelyettesítése stb.) passzivitásba taszította és depriálta a tanulókat.

Nagyjából hasonló módszertani alapokon tanították az orosz nyelvet is, amelynek tanulását az iskolai nehézségeken túl az ellenszenv is hátráltatta, az új generációk számára főként a bezártság miatt. Természetesen mindig voltak és lesznek olyan sugárzó tanáregyéniségek, akik a legunalmasabb nyelvtani témát is (pl. rejtelmek az orosz igeszemléletben) egy aggteleki cseppkőbarlangban tett látogatás színességéhez tették hasonlóvá; meg olyanok is, akik mit sem törődve a kötelező tantervekkel, orosz irodalmat, kultúrát és művelődést hoztak be a nyelvórákra a szépre éheseknek, akár egy lexikai egység tanításának ürügyén is.

Nehéz elképzelni ma már azt a középkori állapotot, amikor a tanár egyes-egyedüli megtestesítője az idegen nyelvi szemléltetésnek mind hanggi, mind fizikai értelemben. Az áttörés, hogy anyanyelvű beszélőket is halljunk, már igen közel volt. A magnetonok elterjedése kiszorította a nehézkes lemezjátszókat: 1963-ban jelent meg Magyarországon az első (Cedamel típusú francia) nyelvi laboratórium. Ezek az eszközök kiváltották, beindították, mintegy hozták magukkal a nyelvtanítás első tömeges módszerét: az audiolingvális módszert, amely a legjellegzetesebb és a leggyakoribb mondatminták (*pattern drills*) begyakoroltatását ígerte a strukturalista nyelvsemlelet és a behaviorista tanuláselmélet összefonódásának következményeképpen. Ugyanekkor vált hozzáférhetővé az audiovizuális (strukturális-globális) módszer, amely az intonáció és az idegen nyelvi kiejtés tökéletes utánzásával jókora lehetőséget biztosított a kevésbé kognitív beállítottságú, ám kiváló hangutánzó képességű tanulók számára (Bárdos 2005: 95–110).

IV. A „régii” nyelvvizsga (1969–1979)

A régi típusú nyelvvizsgában ez az áttörés még nem jelentkezett. Az írásbeli (másfél-két oldalnyi szöveg magyarra és célnyelvre történő fordítása) megelőzte a szóbelit, illetve kizárta, ha az anyanyelvre vagy az idegen nyelvre történő fordítások közül bármelyik sikertelen volt: szóbelire már nem hívták be a jelöltet. Szinte minden ember

nyelvi készség-repertoárjában ezek a leggyengébben fejlett készségek; részben mert tanult készségek, amelyeket csak temérdek munkával lehet elfogadhatóan cizellált állapotba hozni. Ebben az értelemben a vizsga, kimondatlanul bár, de elitista és féloldalas volt: a beszéd fontosságának, primátusának ignorálása mellett csak igen keveseket engedett be a „kiválasztottak” közé. Akik mégis eljutottak a szóbeliig, azokat gyakran szószintű kérdés várta (pl. soroljon fel háziállatokat; időjárás jelenségeket; virágokat; járműveket stb.); vagy nyelvtani tudatosságról tanúskodó mondatszerkezetek bemutatása.

A direkt módszer óta jól ismert örökzöld társalgási témák szerény mértékben kerültek terítékre: mit mond barátjának a kórházban (*speedy recovery*); születésnapkor (*many happy returns*) stb. A társalgás az akkor divatos Angol társalgási zsebkönyv témáiban, inkább számonkérően, mintsem interaktívan folyt. Véges István (1972) számtalan kiadást megért munkájában kísérletet tesz a fonetikai átírás és a magyar helyesírás összebékítésére: *Áj lájk máj wők (I like my work)*. Ehhez hasonló, emlékezetes próbálkozás csak egyszer történt, a Basic English (Ogden 1930) első magyar kiadásában: mindent a magyar kiejtés szerint tüntettek fel. Tehát a 'szoksz' a zokni többes száma. A 'bút, trí, sört, csin, kap, boksz, csíz, mít, kik' stb. szavak kitalálását az olvasóra bízuk.

Sok korabeli tanár az akkor már nálunk is kapható hírhedt Essential English (lemezzen is hallható) szövegei mellett (Eckersley 1938) többnyire a nagy példányszámban többször is kiadott Báti–Véges, vagy a Tarján–Korenchy nyelvtanát gyakoroltatta. Jelen sorok szerzője az Essential English nagyméretű faliképeit (Gatenby–Eckersley 1955) más nyelvek tanításában is használta – a képekhez igazított direkt módszerekkel. Belátom, hogy veszélyes vizekre tévedtem azzal, hogy nyelvtanári „anyagismeretet” eleveníték fel, de fontos, hogy a mai olvasó érzékelje, hogy a hatvanas években már áthatolt az autentikus (már amennyire egy tankönyv autentikus) brit nyelvkönyvek garmadája a vasfüggönyön. Ráadásul Lado (1964) vagy Mackey (1965) alapvető elméleti munkáinak ismerete nélkül zuhant rá a magyar nyelvtanárookra a rengeteg „csempészárú”: az audiolingvális tananyagok teljes spektruma. A szövegeket kísérő hangfelvételek és laborgyakorlatok már szakszerű és korszerű, „elrontani- sem-lehet” (*foolproof*) kézikönyvekkel segítették a vakpróbálkozást. Röviden: a hetvenes években párhuzamosan futottak a régi (nyelvtani-fordító, a direkt és utóhatásai, az olvasatú) módszerek; csúcson volt az audiolingvális és az audiovizuális, miközben a láthatár peremén megjelentek a humanisztikus módszerek (pl. Lozanov [1979] szugesztópédiája). Nem kétséges, hogy a hetvenes évek végére a nyelvtanárok túlnyomó része metodikai szempontból felkészült volt arra a nyelvpedagógiai fordulatra, amit a kommunikatív nyelvtanítás áttörése jelentett. Dúskáltunk a javakban, jobbnál jobb tananyagokban, és nem volt zavaró, hogy Allen (1954) legalább egy évtizeddel megelőzte az audiovizuális módszert a Living English Speechcsel, vagy hogy French (1947) jól használható mondat-táblázatai megelőlegezik az audiolingvális laborgyakorlatokat

(miközben Erasmus Colloquiáira kacsintanak). Virágkorát élte az idiómák tanítása (Kundt 1938, McMordie 1909, 1971, Seidl 1974, Worrall 1964). (Csak évekkel később derült ki számomra, hogy milyen elavult szólások voltak ezek, amikor Canterburyben a szállásadóm tinédzser fia egyetlen délután, nyerítő vihogások közepette, hogy ti. melyek azok, amelyeket ő soha nem mondana, „kipreparálta” Seidl remekművét: alig maradt benne valami tanítanivaló.)

A mérésmetodikában kitört a tesztláz (bár sokak szerint az úgynevezett objektív feleletválasztós teszt (*multiple choice*) csak a nyelvtanuló idegen nyelvi bizonytalanságának mértékét képes megállapítani, s nem lehet mércéje egészes nyelvtudásának). A réges-régi szerzők tesztgyűjteményeit (Mills 1961, Stone 1970, Fowler–Coe 1976 stb.) felváltották a korabeli brit tesztgyűjtemények és vizsgaanyagok példatárai; s bár a sziget azóta is süllyed évente néhány millimétert, az angol nyelv, valamint vizsgáinak terjedése globális hatást ért el. A kommunikatív nyelvtanítás – a beszédfunkciókra fókuszálás és az éppen szükséges funkcionális nyelvtan arányainak megtalálásával – ekkor hatja át a magyarországi nyelvtanítást is. A globális nyelvtanítás és nyelvvizsgáztatás beszüremkedése megmagyarázhatja, hogy miért került a következő alcímbe az „új” szó mégis idézőjelbe.

V. Az „új” nyelvvizsga (1979–89)

Madarak röptét figyelve nehéz eldönteni, hogy mi van előbb: a fej vagy a szárnyak? A nyelvtanítás mindennapi gyakorlata (mint szárnyak) számtalan esetben megelőzte a később elméletileg is igazolt tételeket. Például a funkcionális szemléletű, a kommunikáció lehetőségét felvállaló tananyagoknak is voltak korai előfutárai: András – Murvai 1961; Binham 1968 stb. A hetvenes években jelentősen bővült a modern szemléletű, kompakt (vagyis hanganyaggal és tanári segédkönyvekkel ellátott) modern nyelvkönyveket létrehozó, kreatív magyar szerzők tevékenysége. Volt, aki az Iskolarádióban (pl. Bodóczky–Horlai–Lukács 1979 stb.); mások a Tankönyvkiadó standard tankönyveiben (pl. Czobor–Horlai 1980, Bodóczky–Csonka 1979, Csonka 1990 stb.) mutatták be a legfrissebb módszereiket, magyarokra szabott nyelvtanítási elképzeléseiket. Mondanom sem kell, hogy a Rigó utca tanárai derekasan kivették a részüket ebből. Kétségtelen tény viszont, hogy a brit angoltanítási nagyipar a rákövetkező évtizedben, különlegesen szép kivitelű, óriási példányszámú nyelvkönyveivel lekörözte a magyar könyvkiadást, ami bizonyos értelemben csökkentette a publikálásra szánt magyar nyelvtanítási vágyakat, illetve a tantermek zárt ajtajai mögé lökte a honi ötleteket. Mégis, a hazai metodika színpadán már minden készen állt a kommunikatív nyelvtanítás befogadására. Miért jelentett mégis óriási előrelépést az 1979-es (és az 1980-as években már jól dokumentált) vizsgareform?

A vizsga szerkezete

Az „új” vizsga szerkezeti felépítésében, valamint az osztályzatok megoszlásának számában *biztosította a beszéd primátusát*. Másfelől, a *magnetofonos hallásértés* bevezetésével kiküszöbölte azokat az egyenetlenségeket, amelyeket az egyetlen, a jelöltet tanító nyelvtanár; vagy az akár csak akkor és ott, a jelölttel szemben ülő vizsgáztató egyéni nyelvmegjelenítései példázhattak. A magnetofonok megjelenésének országos hatása felmérhetetlen volt: megszűnt a középkori „egyke” kiejtés-állapot: sokféle anyanyelvű beszélőt kellett hallgatni, és ez a „fültiprás” jelentősen kitágította a befogadhatóság kapuit (pl. amerikai angol). Az új vizsga bevezetésének feltétele volt az a magas technikai színvonal, amely a Rigó utca testületét és infrastruktúráját jellemezte: a heti 12 órás kurzusokon az ITK tanárai rendszeresen használták a nyelvi laboratóriumokat (1971–1985: Tandberg IS 6; 1985–92: Tandberg IS 9; 1992–2016: Tandberg IS 10) és tanóráikon is használtak magnetofonokat (Tandberg AT771). (A nyomda, az AV-eszköztár fejlesztése, a tananyagok tárolása és előkészítése elképzelhetetlen lett volna a tanárokkal és anyanyelvű beszélőkkel teamekben dolgozó hangmérnök: Kovács István nélkül.) A vizsgákon használt autentikus hanganyagok felvételei szintúgy a Rigó utca „süketszobájában” készültek. A hallásértés feladatot nyelvi laborokban folytatták le, amely megkötés a nyolcvanas években az egyetemi városokkal bővülő nyelvvizsgákon is kritérium volt. Kezdetben ezek a helyszínek: Veszprém, Pécs, Debrecen, Miskolc, Szeged, Sopron még biztosították ezt a feltételt. A vizsgahelyek rohamos számú növekedésével ez a fontos minőségbiztosítási szempont megszűnt (tisztelet a kivételnek).

Ez a vizsga vezette be a *rövid szövegek olvasását* (gyakorlatilag blattolását), *az újságnyelv* és általában a modern médiák nyelvezetét a szóbeli vizsga részeként. Ez a vizsgapont késhegyre menő metodikai viták forrása lett, hiszen a legtöbb ember az anyanyelvén is nehezen olvas első látásra (arról nem is beszélve, hogy értse is, amit mond). A hangos olvasásra azért volt szükség, hogy rövid idő alatt még több mintát lehessen kiváltani a jelöltből a kiejtés jobb megítélése céljából. (A blattolás természetesen kiküszöbölhető, ha hagyjuk, hogy a jelölt magában egyszer elolvassa a szöveget.) Abban az időben nem voltak elérhetőek az angol újságok: itthon az AP, UPI, Reuters szövegekből összeollózott Daily News volt a legkönnyebben hozzáférhető anyag. A szöveg érthetőségét veszélyeztető „ollózás” oka vagy politikai volt, vagy egyszerűen az alapvető pragmatikai követelmények ismeretének hiánya. Az újságnyelv tanulmányozását nyelvkönyvek is segítették (Land 1975, 1983, Véges 1983 stb.) és ez az átállás egy modern médianyelvezetre ugyanolyan nagy hatással volt a hazai lexika-oktatásra, mint a hallásértésre a magnetofonok használata. (Jelen sorok szerzője a nyolcvanas évek elején rendszeres telexolvasást tartott, hiszen azok a szövegek még nem voltak megvágva.)

A vizsgához az ITK oktatói *tesztbankokat* állítottak fel megegyezéses nyelvtani minimum alapján, de ez csak akkor vált nyilvánossá, amikor tesztgyűjtemények formájában,

több nyelven megjelentek a Tankönyvkiadó gondozásában (pl. Bárdos – Sarbu 1984). Maga a tesztbank sokkal több, kártyákon rögzített és rendezett, nyelvtani szempontból tematizált, de nem statisztikai alapú, szakértői item-elemzésen alapuló rendszert alkotott. Az intézet rendszeresen kiadta a lefutott vizsgaszövegeket, gyakran magyarázatokkal, gyakorlás céljaira. A több nyelven kiadott tesztgyűjtemények mondatgyakorlókat is tartalmaztak. A tematikát „Az Állami Nyelvvizsgák tematikája: közép- és felsőfok (1985)” című kiadvány pontosította. (Az alábbi ábrán csak az általános középfokot mutatjuk be; praktikus okokból az alap- és felsőfokkal, valamint a szakmai nyelvvizsgákkal ebben a tanulmányban nem foglalkozunk.)

	SZÓBELI	ÍRÁSBELI
1.	BESZÉLGETÉS (Ismerkedés, „small talk”, párbeszédés „mezőnyjáték”, témaválasztás, téma kifejtése (Szótár nem használható) Idő: 15 perc a célképzet	NYELVI KÉSZSÉGEK ELLENŐRZÉSE (Nyelvtani és szókinccstesztek; példamondatok fordítása stb.) (Szótár nem használható) Idő: 90 perc
2.	HALLÁSÉRTÉS Célnyelvű szöveg tartalma magyarul (terjedelme másfél-két perc: összefoglalás, értelmezés magyarul) (Szótár nem használható) Idő: 5-10 perc	FORDÍTÁS célnyelvről magyarra: 1000–1100 n terjedelmű általános szöveg (Nyomatott, egy- vagy kétnyelvű szótár használható) Idő: 90 perc
3.	OLVASÁSÉRTÉS (500-600 n terjedelmű, általános idegen nyelvű szöveg fordítása, értelmezése magyarul) (Szótár nem használható) Idő: 5-10 perc	–
Együtt	Minimum 20, maximum 40 perc	Három óra (180 perc)
Jegyek	Beszédértés Beszédkészség (lexika) Nyelvhelyesség Kiejtés Olvasott szöveg értése	Nyelvi készségek Fordítás

2. ábra. Az általános (nem szaknyelvi) középfokú nyelvvizsga szerkezete a nyolcvanas években

A mindenkori vizsga szerkezete az aktuális tájékoztatókból követhető (Sipőczy 1981, Horváth–Gáborján 1990, Katona 2000 – az utóbbi már az Origó), de a módszertana, amely a szikár struktúrába életet lehel, kezdetben csak kevés kiadványban lelhető fel (Bárdos 1982, 1986, Heltai 1990, Szentiványi 1990 stb.), a laikus visszacsatolás pedig

ritka (Kubinszky 1985). A *Vizsgamódszertan az ÁNYB tagjai számára* (Bárdos 1982) a vizsgáztatókkal szemben támasztott követelményeket, betartatásuk módját írja elő. A vizsgáztató vizsgatechnikája, nyelvi felkészültsége, műszaki-technikai felkészültsége és magatartása című fejezetekben.

A *vizsgatechnika* négy témát tárgyal: a kérdezés technikáit, az apropókészség jelentőségét, a határozott vizsgáztatói magatartás kritériumait, valamint az egyes feladatok lebonyolításának lényegi, módszertani fogásait. A nyelvi felkészültség evidenciának tűnik: minden tekintetben anyanyelvi szintű teljesítményre kell törekedni (mindamellett a szerző elismeri (Bárdos 1982: 16), hogy vannak olyan vizsgáztatók, akik inkább alkalmasak középfok, mintsem felsőfok lebonyolítására). Az alelnökök számtalan esetben váltak meg olyan minisztériumok által delegált szakmai vizsgáztatóktól, akik nem feleltek meg nyelvileg az elvárt színvonalnak (ez utóbbi kérdés már inkább a minőségbiztosítás területe).

A *műszaki-technikai felkészültség* nem csak a *magnetofon* vizsgán történő helyes használatára vonatkozott, hanem a lebonyolítás mikéntjére; a szelíd, állásfoglalás-mentes feltárás módszereire (arra kérdezzünk rá, amit a vizsgázó már kimondott). Nem minősítünk a jelölt előtt (olyasmit persze szabad mondani, hogy nem egészen erről volt szó). A maximális hallhatóság a cél, ehhez igazodnak a körülmények (nincs papírzörgés, székhúzogató, vizsgáztatók közti beszélgetés stb.). (A [kazettás] magnókat nem kellett állítgatni, mert egy oldalon 12–14 hanganyag szerepelt egymás után a vizsgán előírt gyakorisággal, így minden új vizsgázónál ott lehetett folytatni, ahol korábban megálltak.) Az olvasott cikkek kiválasztása a vizsgáztatók (időigényes) feladata volt: az elvárásnak megfelelően a köznyelvi társalgás legegyszerűbb pl. igéi esetén, az újságnyelvre jellemző, kicsit elvontabbak is szerepelhetnek (elismer, fenntart, hangsúlyoz, vádol, figyelmeztet stb.).

A technikai felkészültség körébe sorolhatjuk az írásbeli javítását is (Heltai 1989). Nem a nyelvtani tesztek (többnyire rozettás) javítása a probléma (kivéve a mondatfordításokat), hanem a teljes szövegeken a megfelelő hibajelek gondos és alapos feltűntetése (ez különösen jelentőssé vált attól az időponttól, amikortól a jelöltek utólag megtekinthették írásbeli munkáikat). A magyarra vagy célnyelvre fordítás mikéntje gyakran váltott ki vitát az anyanyelvű lektorok és a nem anyanyelvű oktatók között, sőt még tanszékek között is. Az elbírálás egységesítése céljából ez a vizsgamódszertan (Bárdos 1982) már *teljes hibatipológiával* szolgált. Súlyos az a hiba, amely torzítja a jelentést, vagy értelemzavaróvá válik (pl. jelentéssel ellentétes, vagy a közvetítendő információ teljes hiánya); nem súlyos az olyan szűk vagy tág fordítás, amely nem a leginkább odaillő szót használja, illetve súlyos következménnyel nem járó alaki vagy nyelvtani hiba. Az alakítól az információs hibáig tehát szerepel:

- alaki hiba (helyesírás, központosítás, elírás stb.);
- nyelvtani hiba (nyelvtani szerkezetek tévesztése);
- lexikai hiba (téves szóhasználat, vagy válogatás a lexikában);
- stilisztikai hiba (nyelvhasználati funkciók, regiszterek tévesztése);
- fordításképtelenség (értelmetlen, zagyva szöveg; ellentétes jelentés; szöveg kihagyása).

A felsorolás úgy is szemléltethető, hogy a legutolsó a súlyos információs, vagyis „minőségi” jellegű hiba, míg a többi négy inkább „mennyiségi” mutató, amelyeknek nem feltétlenül vannak minőségi következményei.

A vizgáztató magatartása című rövid fejezet arra hívja fel a figyelmet, hogy a nyelvvizsga kulturálisan érzékeny jelenség, így a vizgáztató magatartását a célnyelvben uralkodó normák szabják meg. A pontosság és következetesség mellett az eleve empátiás, odaadó jelöltre figyelés a legfőbb munkaeszköz: eldöntheti a társalgás irányát, a segítségnyújtások mikéntjét. Az ideál a nyugodt vizgáztató (számomra Lukács Krisztina vagy Radványi Tamás), mert sem modorunk, sem mozdulataink, sem mosolyunk nem lehet harsány vagy tolakodó. „A vizgáztató ne legyen fölényes; viselkedése legyen szolid, nyugodt, összefogott, kiegyensúlyozott, kellemes és derűs” (Bárdos 1982: 17).

Az első vizsgamódszertan legnagyobb erénye mégis az, hogy az élő nyelvvizsga mindennapi gyakorlatából származó *értékelési skálákat* közöl minden bírálendő területre (beszédkészség, beszédértés, kiejtés, nyelvhelyesség, olvasott szöveg értése, olvasott szöveg célnyelven (szóbeli), fordítás magyarrá, fordítás célnyelvre. Több mint érdekes most elolvasni a KER „mire képes” („*can do*”) skáláit *húsz évvel (!) megelőző* értékelési javaslatokat. Az ÁNYV kategóriái az egészet, az egészszletes teljesítményt tartják szem előtt, a KER skálái azokat az apró kapaszkodókat, amelyekkel az egymásra hasonlító kategóriák mégis elválaszthatók: az előbbi integrált, az utóbbi analitikus módszerekkel összeszerkesztett. Vessd össze: „The best descriptors – those that were sorted consistently to the correct pile and that were marked as clear and useful – were then used to create a series of overlapping 50-item checklists of descriptors for different levels” (North 2017). Az utóbbi már a „big data” világában lebeg, jobban bízik a statisztikában, mint a saját látásában: a kettő persze nem zárja ki egymást. A vizsgakkreditációk „benchmarking” eljárásaiban, a kritikus pontokon, mindig gyakorlott szakértők segítségét veszik igénybe, nem pedig egy virtuális világ statisztikai látszatait. Szemléltetésként bemutatjuk a beszédkészség értékskáláját, ahol a közép- és felsőfok még egy skálán szerepel (és rejtetten a beszéd folyamatossága kritériumot is figyelembe veszi).

<p>1 – elégtelen: a jelölt nagyon habozva, alig egy-két mondatot vagy mondattöredéket mond, az anyanyelvi beszélő számára elviselhetetlenül hosszú szünetekkel. Amit mond, többnyire érthetetlen, kivehetetlen.</p>
<p>2 – elégséges: nagyon lassan, de éppen elviselhető hosszúságú szünetekkel kifejezi a legegyszerűbb nyelvi szükségleteit, igen sok nyelvtani hibával.</p>
<p>3 – közepes: nem veszíti el a természetes társalgás fonalát, határozottan képes kifejezni igényeit, de pontatlanul beszél, elég sok szünettel.</p>
<p>4 – jó: (Középfok) – határozottan, folyamatosan fejezi ki magát, nem kell szavakat keresgélnie, de még vannak beszédében pontatlanságok. (Felsőfok) – könnyedén boldogul a mindennapi köznyelvvél, folyamatosan, majdnem teljesen hibák nélkül beszél. Összetettebb, idiomatikus kifejezésekbe még bele-belebonyolódik, kerüli az elvontabb lexikai síkokat.</p>
<p>5 – jeles: (Középfok) – könnyedén boldogul a mindennapi köznyelvvél, folyamatosan, hibák nélkül beszél. (Felsőfok) – folyamatosan, hibák nélkül beszél. A célnyelvre jellemző bonyolultabb, idiomatikus kifejezéseket természetesen, anyanyelvi tempóban használja. Elvont témákról is folyamatosan – ha nem is szakszerűen – beszél.</p>

3. ábra. A beszédkészséget értékelő skála (Bárdos 1982: 8)

A vizsga minőségbiztosítása

Kezdetben még minden résztvevő („stakeholders”: vizsgaszolgáltató, vizsgázó, munkaadó) elégedett volt a vizsga bevalásával, és így inkább a nagyközönség igényeit igyekezett az intézet teljesíteni: ne csak Budapesten legyen vizsga; hamarabb tudják meg az eredményeket; legyenek az írásbelik megtekinthetőek, stb. A legutóbbival túl is lépett egy határt: egyetlen ismert, nemzetközi nyelvvizsga sem csinált ilyet. A vizsgák minőségéért az adott nyelv alelnöke vagy alelnökei (később fővizgáztatók) feleltek. A vizsgaszövegeket, tesztek, példamondatokat tanszéki szinten válogatták és itt döntötték el, hogy melyik legyen (gyakran rejtett kipróbálás után, hiszen az ITK oktató intézmény is volt, illetve oktatói más intézetekben is dolgoztak). A szóbeli vizsgákat rendszeresen látogatták az alelnökök, akik nem a jelöltekkel, hanem a vizgáztatókkal, a vizgáztatók szaktudásának minősítésével foglalkoztak. Az írásbeli vizsgákat véletlenszerű szűrőpróbákkal, de közel fele-fele arányban szintén az alelnökök ellenőrizték. A vizgáztatók a legismertebb, legjobb nyelvtanárok közül kerültek ki, az ÁNYV Bizottság tagjainak javaslatára. Felkészítésük vizsgalátogatásokkal kezdődött, majd a módszertani tájékoztatók ismeretének birtokában próbavizgáztatásba foghattak, párban egy régi vizgáztatóval. Akit nem találtak alkalmasnak, azt eltanácsolták (a kiváló nyelvtanár nem biztos, hogy jó vizgáztató, nem neki kell elmondani a feladat

megoldását (!). A vizsgáztató párok állandó forgása dacára kialakultak olyan duók, akik nagyon szerettek és jól tudtak együtt vizsgáztatni. Kérésük, hogy együtt maradhassanak, jelentősen nehezítette a Nyelvvizsga Csoport tagjainak munkáját (akkor Rózsa Ágnes vezetésével), akiknek rendkívüli adminisztratív pontossága és odafigyelése nélkül aligha szaladt volna a szekér – a folyamatosan növekvő vizsgaszámok mellett. Az együttműködésnek ez az ideális „csúcsformája” hosszú éveken át megmaradt: ez biztosította az intézet és egyben a nyelvvizsgák stabilitását, elfogadottságát. Miként lehetséges, hogy egy ilyen, kiemelkedő kreativitással és összhangban szerkesztett nyelvvizsgarendszer nemzetközi szempontból tökéletesen ismeretlen maradt?

Kitekintés

A fenti kérdésre több, meglehetősen triviális válasz adható. Egyfelől, kinek kellene olyan angolszász (vagy más célnyelvű) nyelvvizsga, amelyet (még ha anyanyelvi kontrollal is, de) Magyarországon fejlesztettek; miközben a brit és amerikai nyelvvizsgák globális összecsapásokban igyekeztek minden elképzelhető piacot megszerezni. Arról, hogy egy ország milyen nyelvvizsgát, nyelvvizsgákat fogad el, kizárólag saját maga dönt, és nálunk, akkor, a magyar Állami Nyelvvizsga volt elfogadott. Bezártságunk miatt még a kétely sem merült fel, hogy most jobbat, vagy rosszabbat kapott-e az ország, mint amit megérdemelt: véleményem szerint mindvégig a világszínvonalnak megfelelőt, azzal a többlettel, hogy a tanítási kontextusoktól nem távolodott el. Meglehetősen titokban ugyanekkor lehetett a követségeken is külföldi nyelvvizsgákat letenni, természetesen a kinti áron. (A bátrakról, akik be mertek menni egy „nyugati” követségre, fénykép készült.) 1989-ben az ELTE ITK szerződést kötött a cambridge-i nyelvvizsgák itthoni forgalmazására. Már ez a tény is bizonyítja, hogy az ELTE ITK nem törekedett arra, hogy „monopóliummá” váljon (később ugyanis ez lett a vád), brit szakmai kapcsolatai pedig kifejezetten erősek voltak, ahogy azt doktori védések is bizonyították. Arra, hogy a külföldi nyelvvizsgák változatos csapatával itthon „megmérkőzzék”, még jó tíz évet kellett várni. Nem véletlen, hogy a népesebb mezőnyben az első, új kritériumok fényében elfogadott nyelvvizsga mégis az Origo lett (Katona 2000).

Amennyiben olvasóink – mint szakemberek – figyelemmel követték a brit nyelvi mérés és értékelés fejlődését (Davies 1978, Skehan 1988, Bachman 2000, Alderson – Banerjee 2001, 2002 stb.), akkor azt is érzékelhették, hogy micsoda hatalmas szakmai potenciál rejlik a kommerciális érdekek által is támogatott nyelvvizsgák tökéletesítésében. Nem kétséges, hogy az idegen nyelvi mérés és értékelés a pedagógiai mérés és értékelés legdinamikusabban fejlődő ágazata: élcsapata volt a század utolsó két évtizedében. Akadt ugyan egy-két olyan vizsgatechnikai fogás, amely az ELTE ITK következő vizsgáiban mintegy visszaköszön (pl. a képleírás), de a brit nyelvvizsgák többsége, éppen a gazdasági siker miatt, nem nagyon törekedett

a legmodernebb, ám költségesebb teszt-technikák bevezetésére (tisztelet a kivételnek: Cambridge). Továbbra is világméretben dúlt a harc az inkább megítélésalapú ('judgement') brit, és a sokkal inkább empirikus, észak-amerikai nyelvvizsgák között. A megfogalmazás afféle minőség–mennyiség szembenállást sugall, erről azonban szó sincs. A brit vizsgák – csakúgy, mint az Állami Nyelvvizsga – a legjobb szakértők segítségével hoztak létre vizsganyagokat, amelyeket azután soha többé nem használtak. Bizonyosságul mellékelünk egy elemzést Aldersontól (1987: 3): „Due to the constant need to produce new examinations and the lack of emphasis by exam boards on the need for empirical rather than judgemental validation, these examinations are rarely, if ever, tried out on pupils or subjected to the statistical analyses of typical test production procedures. Examination boards do not see the need to pretest and validate their instruments, nor conduct post-hoc analyses of their tests' performance.” A brit vizsgafejlesztő cégeknél is a fővizsgáztató (alelnök) feladata volt az új tesztek kifejlesztése, amely munkálatokban szakértői véleményekre támaszkodott. A kijelölt szakértői bizottság hozta létre, ellenőrizte és véglegesítette a teszteket („moderation”), és ennek jegyében tájékoztatták, illetve képezték tovább a vizsgáztatókat. Ebben az értelemben a nyolcvanas évek magyar állami nyelvvizsgái összhangban voltak más (külföldi) vizsgaszolgáltatók eljárásaival. Még azt az előnyt is mellé tehetjük, hogy a magyar Állami Nyelvvizsga fejlesztői ugyanakkor az adott nyelv avatott oktatói is voltak, így a tanítás és a mérés-értékelés tevékenységkörei a kívánatos irányban: párhuzamosan mozogtak (vesd össze: Báthory Zoltán gondolatai a könyvborítón, Bárdos [2002] mérésmetodikai monográfiáján). Mindeközben mind a tanításban, mind a nyelvi értékelésben, a kommunikatív kompetencia mint célképzet igen tág, általános rendező elvvé vált, amelynek eredményeként a nyelvtanításban használt modern technológiák (akkor a videózás és a televízió is: képi többletek), valamint a funkcionális beszédszemlélet és a kulturális másságot is tudatosító pragmatikus elképzelések győzedelmeskedtek. Újfént nagyjából tíz év kellett ahhoz, hogy a kerék megint forduljon egyet, és létrehozzon egy újabb nyelvvizsgát (Bárdos 1987). Az ELTE ITK vizsgáztatói jelentős publikációkkal, folyóiratok indításával, kitűnő nyelvkönyvekkel és öntökéletésítéssel (!) a modern tesztfejlesztés útjára léptek. A sajnálatos nem az, hogy erről a megújulásról külföldön nem szereztek tudomást, hanem inkább az, hogy a nyelvészet vagy a neveléstudomány hazai elefántcsonttornaiba ez a jelentős alkalmazott nyelvészeti, nyelvpedagógiai és mérésmetodikai fejlődés alig-alig volt képes beszivárogni.

IV. Az ÁNYV megszilárdulása és átalakulása (1989-1999)

A kilencvenes években a nyelvtanítás és a nyelvtudás célképzete, konstruktuma messze jutott a huszadik század elejének szedercsíra-állapotától: szinte egy kifejlett embrió képét mutatja. A magyarázat egyszerű: a mai nyelvtanítás nem kizárólag az iskolai nyelvtanítás lombikjában fogant, vagy ha mégis, akkor számtalan egyéb forrás is táplálja. A tanár már nem ura a bemenő, bejutó tananyagoknak: hogyan is korlátozhatná a filmek,

az elektronikus kommunikáció, a keresők és műsorkínáló szoftverek beözönlését – egyáltalán: a hálózati világ szüntelen be- és átáramlását agysejtjeinkben. A digitális „big brother” vette hátára a nebulókat, rafináltan kontrollálja és tartja fogságban gondolataikat. Ez az uralom egyszerre nyelvi és képi, kultúrája globalista, tagadhatatlanul az angolszász nyelvek hegemoniájával. A modern nyelvtanár maga is „áramlatívá” válik, és megpróbálja az aranymosást újra megtanítani a világháló sekélyes vizeiben. Ugye milyen könnyű átcsúszni a mai időkbe? NEM, a kilencvenes években *még nem ez volt*, ami most: akkor cseréltük a Robotron írógépeket személyi számítógépekre; még használtuk a nyolcvanas évek videóit (Készüljünk a nyelvvizsgára!); és futottak a nyelvi laborok is. A bőség kétségkívül jelen volt: egyfelől az ELTE ITK szerzői gárdája egyre több és jobb tankönyvvel állt elő, és egyre többet lehetett importálni a célnyelvek anyaországaiból is. A legfontosabb változás azonban az volt, hogy megnőtt a nyelvvizsgáért dolgozó szakemberek meta-tudása: már képesek voltak saját magukat kívülről szemlélni, méghozzá a korszerű mérésmetodikai szakma szemüvegén át. Ennek jele a szakfolyóiratok megjelentetése és benne a célképzetek és módszerek pontos megfogalmazása, nemkülönben kutatásokkal történő alátámasztása (Testing Matters 1996–97, Nyelvvizsga Fórum 1998–2000, Nyelvi Mérce 2001-től – Fekete Hajnal, Heltai Pál, Horváth Iván, Staub Valéria, Tamássyné Bíró Magda és kollégáinak érdeme). Ezek a folyóiratok azonban nem rendelkeztek ISSN-számmal, kis példányszámuk és terjesztési problémák miatt nem ragadták meg a témához valóban értő (bár maroknyi) kutató figyelmét. A szerzők inkább alkalmazott nyelvészeti konferenciákon szerepeltek, de a nagyobb mérésmetodikai tapasztalatokat felmutató neveléstudományi rendezvényektől (pl. ONK, PÉK) rendre távol maradtak; országos és nemzetközi folyóiratokban keveset publikáltak. Ugyanakkor tovább utaztak az ÁNYV „szekerén”, kényszeresen is felvállalva azokat a szakmai kihívásokat, a nem mindennapi presszúrákat és dresszúrákat, amelyeket a kilencvenes évek hoztak magukkal. A rendszerváltás után minden „állami” eleve privatizációs rohamban részesült, és olyan könnyű volt a világszerte elfogadott szakmai erők összpontosítását (amilyen egy jól fejlesztett vizsga mint márka) politikai színezetű monopóliumnak lefesteni. Gyakran kicsinyes, provinciális üzleti érdekek húzódtak meg a végzálás mögött (jobb lenne, ha ezt a vizsgát mi csinálnánk...). Az átalakítások nagyformátumú személyisége dr. Manherz Károly professzor és államtitkár volt, aki úgy tudott érdekeket képviselni különleges menedzseri adottságaival, hogy ahhoz pénz is szerzett, ugyanakkor a szakma is jól járt: a nyelvvizsgák modern akkreditációja és minőségbiztosítása döntően „földről” indult el, háttere a minisztérium és az asszisztáló OIT (Országos Idegennyelvi Tanács). Így került sor arra, hogy az ELTE ITK Állami Nyelvvizsgálója az első külföldi „akkreditációval” is rendelkező szellemi termék lett: voltaképpen a minisztérium felkérésére (a Világbank támogatásával) a Goethe Intézet szakértői, illetve a British Council által felkért holland és angol szakértő (Crighton–West 1993) többféle igényt kielégítő elemzést készítettek. A British Council képviselői az intézet múltjára is kíváncsiak voltak, alapos munkát végeztek; ugyanakkor a német jelentésben számos alá nem támasztott állítás, mendemonda

szerepelt: áttételesen nyilvánvaló volt az üzleti érdek, az „egyeduralom” megtörése, a „nemzetköziesítés” előtérbe helyezése, (az akkor még inkább a csak várhatóan EU-országok) gyarmatosítása – globális érdekeknek álcázva. Ma már igazoltak tekinthetjük Heltai Pál és Fekete Hajnal (1996) kasszandrai tanulmányait a két akkreditációs dokumentumról. A fő paradoxon voltaképpen az volt, hogy mindkét jelentésben szerepeltek olyan elvárások és követelményszintek, amelyeket a bírálók a saját vizsgarendszereikben sem tudtak megvalósítani.

A vizsga szerkezete

Apróbb változtatások mellett, közel egy évtizedig, az alábbi szerkezetben működött az általános középfok (4. ábra).

Tétel	SZÓBELI (A TÍPUS)	ÍRÁSBELI (B TÍPUS)
1.	IRÁNYÍTOTT BESZÉLGETÉS (Ismerkedés, rövid párbeszéd megadott témakörökből, lásd vizsgatematikák; inkább interjú) (Szótár nem használható) Idő: 5–8 perc	NYELVI KÉSZSÉGEK ELLENÖRZÉSE Nyelvtani és szókincesztek (45+5): 50 feladat (tesztlépés); Idő: 30 perc (Szótár nem használható)
2.	ÖNÁLLÓ TÉMAKIFEJTÉS KÉPEK ALAPJÁN (szintén a megadott témakörökkel összefüggésben) (Szótár nem használható) Idő: 5–8 perc	MAGYARRÓL CÉLNYELVRE FORDÍTÁS 500 n terjedelmű általános, egybefüggő szöveg (Szótár használható) Idő: 30 perc (?)
3.	SZITUÁCIÓS PÁRBESZÉD (Magyar nyelvű szituáció alapján nyelvileg megfelelően kell reagálnia: partnere a vizsgáztató) (Szótár nem használható) Idő: 5 perc	IRÁNYÍTOTT FOGALMAZÁS Több téma közül választva: levél írása, vagy egy téma kifejtése a megadott utasítások alapján, szintén a megadott témakörök alapján (15–20 sor). Idő: 60 perc (?) (Szótár használható) (A fordítás és az irányított fogalmazás feladatok sorrendje felcserélhető!)
4.	HALLÁSÉRTÉS (Beszédértés laboratóriumban) Célnyelvű beszélt nyelvi szöveg (1500 n), normál tempóban: háromszori meghallgatás. A „ráhangolódás” után feladatlapokat kell megoldani a hallottak alapján: magyar nyelvű kérdésekre magyar nyelvű válasz (Szótár nem használható) Idő: 20–25 perc	CÉLNYELVRŐL MAGYARRA FORDÍTÁS 1000–1200 n (később 3000–3200 n) terjedelmű, célnyelvű általános szöveg <i>fordítása</i> magyarra (Szótár használható)
5.	----- (csak felsőfokon: tolmácsolás)	OLVASÁSÉRTÉS 1500–1700 n terjedelmű, célnyelvű szöveg(ek) alapján <i>feladatlapok kitöltése</i> (kérdésekre adott válaszok) Idő: 90 perc (?) (Szótár használható) (Csak felsőfokon tömörítés célnyelvre)
Összes idő és értékelés	Kb. 35–40 perc Kommunikatív érték, szókincs, nyelvhelyesség, kiejtés, beszédértés	210 perc (három és fél óra)

4. ábra. Az Állami Nyelvvizsga a kilencvenes években: általános (nem szaknyelvi) középfok

pontosabban megítélhető nyelvi megnyilvánulásokat váltott ki a jelöltekből. A feladatok megváltoztatása nemcsak a külső elvárásoknak szóló megfelelést, hanem a modern nyelvtanításban teret nyert újabb és tömegesen elterjedt módszertani megoldásokat is tükrözi. A kommunikatív nyelvtanítás túlsúlya az írásbeliséget vagy a nyelvhelyességet ugyanolyan helyzetbe hozta vagy inkább hagyta, mint régen: továbbra is a leggyengébb (Sipőczy 1988). Megszüntetett néhány kevésbé kedvelt tételt (pl. újságszöveg blattolása), valamint az élő vizsgán rögtöni elmeséléshez kapcsolt (prompt) hallásértést. Újdonság a képleírás (lehetőleg tematikus elvárásokkal). Ez az önálló, viszonylag folyamatos szóbeli leírást elváró feladat még csak ment valahogy, viszont az interaktivitás távol állt az (egykori) iskolai reflexektől. Így nagyon nehéz volt interjú helyett társalogni, nem is beszélve a szituációs kártyákról (amely eleve megszünteti a valós kommunikáció egyik éltető elemét: a véletlenszerűséget). Eleinte a jelöltek meg sem értették, hogy mi az elvárás: hogy el kell játszani, elő kell adni a szituációt, ráadásul a vizsgáztatóval mint egyenértékű partnerrel.

Új feladat volt az irányított fogalmazás is. A fantázia fejlesztése, a kreatív írás képességének szisztematikus felépítése az anyanyelven nem tartozott a kiművelt, kimódolt és főként begyakorolt területekhez. A testes fordítás és olvasásértés egyaránt hozzájárult az írásbeli időtartamának növekedéséhez (210 perc), a Rigó utca mégis hosszú ideig kiállt és kitartott a kétnyelvű vizsgák mellett. (A kétnyelvű vizsga csak annyiban más, mint az egynyelvű, hogy az utóbbi nem vallja be, nem ismeri el, hogy egy kétnyelvűben mindenképpen van oszcillálás a két nyelv között, még ha ez nem is mindig tudatosul.) Kétségtelen, hogy a vizsga egészében a kommunikációra való képesség könnyebben exponálódott mind szóban, mind írásban: pragmatikai okokból értelmetlenné vált a tematikus magolás. A képekben ábrázolt, valamint elképzelt és végül az éles vizsgán kiváltott szókinccs nem mindig harmonizált a leggyakoribb szavak anyanyelvűeknél is megtapasztalható ökonómiájával: magyarul, a ritka szavak kizárása nehezebb volt a vizsgafejlesztőknek, mint az újságszövegek esetében korábban. Összességében az új vizsgarendszer sokrétűségével, változatosságával szervesen illeszkedett az állami nyelvvizsgák fejlődésének ívébe, és nemzetközileg is érvényesen képviselte a lehetséges mérészetodikai világok legjobbját.

Támadások és fejlesztések az ANY utolsó éveiben

A kilencvenes évekre kedvelt ujjgyakorlattá vált újságírói körökben az Állami Nyelvvizsga pellengérré állítása, megbélyegzése. Varga (1989: Mégse fütyülnek a Rigónak?) például a következő allegóriával indít. „Nézem a hölgyet, aki bizonyára a Rákóczi tér közelsége miatt foglalt puccot magának a Rigó utcában. Nézem a hölgyet, amint hanyagul nekidől az Idegennyelvi Továbbképző Központ falának, és az jut eszembe, hogy végső soron felfedezhető valami hasonlóság közöttük: neki is, az intézménynek is rossz híre van.” Kétségtelenül rendkívül szuggesztív indítás ez, már ha valaki ilyen sokáig

nézi, akarom mondani olvassa ezt a cikket... Szerencsére, az épületben dolgozó nyelvtanárokat és egyben nyelvvizsgáztatókat nem zavarták meg az ilyen krokik (vagy kokik?) (Nógrádi 1989), hanem jobbításokra törekedtek.

A vizsgáztatók kiképzésére videotréning készült (Fekete–Dalglish–Hughes 1993). Korábban már említett folyóirataikban rendszeresen publikáltak vizsgamódszertani, vizsgatechnikai eljárásokat, fejlesztéseket, amelyeknek valóban fogantja lett a vizsgák tényleges lebonyolításában. Ugyanezekben a folyóiratokban egyre gyakrabban jelentek meg „külső” szakemberek írásai (Csépes Ildikó, Dróth Júlia, Einhorn Ágnes, Hock Ildikó, Károly Krisztina, Kontra Edit, Kormos Judit, Major Éva, Öveges Enikő, Szabó Gábor és mások). Bizonyos értelemben a kilencvenes évektől érvényes, hogy a szakma valamiképpen mégis a Rigó utca köpönyegéből bújt elő: olyannyira előbújtak, hogy egyszerre teremtették meg az eredet (!) konkurenciáját, azzal, hogy más nyelvvizsgák fejlesztésébe is bekapcsolódtak, ugyanakkor segítettek is e tudományág fejlődését, később intézményesülését is (pl. Nyelvvizsgáztatók és Nyelvtudásmérési Szakértők Országos Egyesülete, Dávid Gergely). Többen szerepeltek a Nyelvvizsgát Akkreditáló Testület pályázatott tagjaiként (NYAT). Az utóbbi testület a Nyelvvizsgát Akkreditáló Központ (NYAK) hathatós támogatásával végezte el a honi és külföldi nyelvvizsgák heroikus csereszabatosítását a Közös európai referenciakerettel az új évezred első évtizedében – ez azonban már egy másik történet. Az ELTE ITK minőségbiztosítási eljárásai, maga a koncepció (a vizsgáztatók rendszeres, az írásbelik és szóbelik arányos és véletlenszerű ellenőrzése) nem változott jelentősen a korábbiakhoz képest a század utolsó évtizedében, ellenben a statisztikai és technikai háttér fejlődése jelentősen emelte az objektivitást. Mind a külső laikusok, mind a belső, felelős vizsgafejlesztők megtapasztalhatták az évek során, hogy tökéletes nyelvvizsga nem létezik, csak újabb, divatos megalkuvások és kétségbeesett erőfeszítések sorozata következhet a nyelvtudás-fogalom bekerítésére. Számos ilyen „talányt” ismerünk: nincsenek végleges tudások, csak cáfolható és cáfolandó approximációk. Ugyanakkor ez az utolsó vizsga – az Origo kifejlesztésével – nagyot lépett előre a végtelenség tűnő úton.

Epilógus

A huszadik század utóbbi néhány évtizedéről szóltunk egy intézmény szellemi-szakmai tükrében: nem a teljes feltárás, hanem inkább az érdeklődés felkeltése céljából. A szelektív emlékezet és számos szubjektív szuperplán alkotja az alábbi „tabula gratulatoria” kistotálját. A szereplők mindegyike, így vagy úgy, hozzájárult és formálta a „Rigó utca” és az Állami Nyelvvizsga imázsát. Ennek elméleti vonulata a vizsga- és tananyagfejlesztés; gyakorlati megvalósulása pedig a vizsgáztatás és a nyelvtanítás, mint több évtizedre kiterjedő, a nemzetközi trendeket időnként megelőző, majd követő, intenzív és következetes szakmai tevékenység. Ahogy hullnak és múlnak munkák és napok, egyre inkább látom, hogy milyen szerencsés voltam, hogy ilyen közösségben dolgozhattam.

AZ ELTE ITK LEGJELENTŐSEBB FŐÁLLÁSÚ TANÁRAI	(A SZELEKTÍV EMLÉKEZETBEN)
<p>ANGOL TANSZÉK</p> <p>Bartáné Aranyi Edina Bárdos Jenő Bodóczky Caroline Czobor Zsuzsa Cseppkövi Ilona Csonka Margit Fekete Hajnal Fonyódi Jenő Fonyódi Jenőné Heller Anikó Heltai Pál Horlai György Katona Lucia Kontráné Hegybíró Edit Kutor Susanne Lukács Krisztina Marosán Lajos Németh Pálma Patthy Miklósné Rawlinson Zsuzsa Sarbu Éva Szentgyörgyvári Artúr Szentiványi Ágnes Székács Györgyné Walkó Zsuzsa</p>	<p>NÉMET TANSZÉK</p> <p>Antal Mária Borbély Mária Csizmadia Miklós Deákné Heidrun Emericzzy Tibor Gáborján Katalin Hahn György Molnár Judit Szabó László Szinna Undine Szitnyainé Gottlieb Éva Tamássyné Bíró Magda</p> <p>OROSZ TANSZÉK</p> <p>Galló András Gecső Sándor Gergely Katalin Halász Lászlóné Horváth Iván Juhász Gyula Sipőczy Győző</p> <p>FRANCIA TANSZÉK</p> <p>Dési Ágnes Gergelyi Mihály Kaposi Tamásné Kiss György Staub Valéria</p> <p>SPANYOL TANSZÉK</p> <p>Hübner Katalin</p>

5. ábra. Ötven emlékezetes tanáregyéniség az ELTE ITK-ból

Gondolom, hogy többek számára ez a befejezés világossá tette, hogy e sorok szerzője *megbékéléssel tekint vissza az* emberi és szakmai küzdelmeknek erre a három évtizedére: a mai magyar idegen nyelvi értékelés és nyelvvizsgáztatás gyakorlatát meghatározó tudós tanárok életművére. Farewell!

IRODALOM

- Alderson, J. Ch. – Krahnke, K. J. – Stansfield, Ch. W. (1987, eds.): *Reviews of English Language Proficiency Tests*. Washington, D.C.: TESOL.
- Alderson, J. Ch. – Banerjee, J. (2001): Language Testing and Assessment. I. *Language Teaching* 34, 213–236. Cambridge: Cambridge University Press.
- Alderson, J. Ch. – Banerjee, J. (2002): Language Testing and Assessment. II. *Language Teaching* 35, 79–113. Cambridge: Cambridge University Press.
- Allen, W. S. (1954): *Living English Speech. Stress and Intonation Practice for the Foreign Language Student*. London: Pearson ELT.
- András László, T. – Vándorné Murvai Márta (1961-től folyamatosan): *How to Say it in Hungarian?* Conversational Phrase Book. Budapest: Tankönyvkiadó.
- angol.info/2019/04/28/hogyan-tudja-egyetlen-nyelvvizsga-romba-donteni-az-egesz-nyelvvizsgarendszer-hitelesseget (Máthé Elek) *Hogyan tudja egyetlen nyelvvizsga romba dönteni az egész nyelvvizsgarendszer hitelességét?*
- Állami Nyelvvizsga Bizottság (1985): *Az Állami Nyelvvizsgák tematikája: közép- és felsőfok*. Budapest: ELTE ITK.
- Bachman, L. F. (2000): Modern Language Testin at the Turn of the Century: Assuring that What we Count Counts. *Language Testing* 17 (1) 1–42.
- Bárdos Jenő (1982): *Vizsgamódszertan az Állami Nyelvvizsga Bizottság tagjai számára*. Budapest: ELTE ITK.
- Bárdos Jenő (1983): A nyelvvizsgamódszertan néhány kérdéséről. *Modern Nyelvtanítás. A felnőttoktatás tájékoztatója. 1963–1983*, 247–259. Budapest: TIT Idegennyelv-oktatási Központ.
- Bárdos Jenő (1986): *State Language Examinations in Hungary. A Survey on Levels, Requirements and Methods*. Bulletin, National Institute of Education. Budapest, Országos Pedagógiai Intézet (OPI).
- Bárdos Jenő (1987): *Programtervezet az Állami Nyelvvizsga rendszerének továbbfejlesztésére*. Budapest, ELTE ITK 1–46.
- Bárdos Jenő (1988): *Nyelvtanítás: múlt és jelen*. Gyorsuló Idő. Budapest: Magvető Kiadó.
- Bárdos Jenő (2002): *Az idegen nyelvi mérés és értékelés elmélete és gyakorlata*. Budapest: Nemzeti Tankönyvkiadó.
- Bárdos Jenő (2005): *Élő nyelvtanítás-történet*. Budapest, Nemzeti Tankönyvkiadó.
- Bárdos Jenő – Sarbu Aladárné (1984): *Angol feladatgyűjtemény a középfokú nyelvvizsgálóhoz*. Budapest: Tankönyvkiadó.
- Báti László – Véges István (1965): *Angol nyelvkönyv kezdőknek*. Tanuljunk nyelveket! Negyedik kiadás. Budapest: Tankönyvkiadó.
- Binham, Ph. (1968): *How to Say it?* London: Longman Group Limited.
- Bodóczy Caroline – Csonka Margit (1979): *Do you Speak English?* I–IV. (Államvizsgára előkészítő társalgási gyakorlatok. Budapest: Kereskedelmi Továbbképző és Módszertani Központ.
- Bodóczy Caroline – Horlai György – Lukács Krisztina (1979): *Down to Earth*. Angol nyelvvizsga előkészítő. Budapest: Iskolarádió.
- Common European Framework of Reference for Languages*. Learning, Teaching, Assessment. (2001): Council of Europe. Modern Languages Division, Strasbourg. Cambridge: Cambridge University Press.

- Crichton, J – West, R. (1993): *Az idegennyelvi vizsgák fejlesztésével kapcsolatos magyarországi konzultánsi munka. Jelentés és javaslatok*. Budapest – Manchester: The British Council.
- Csonka Margit (1990): *Let's Speak English!* (Angol nyelv középfokon, társalgási gyakorlatok). Angol nyelvkönyv III–IV. Budapest: Tankönyvkiadó.
- Csépes Ildikó (1999): Issues in Testing Spoken Language in Pairs. *Nyelvvizsga Fórum. Nyelvvizsgáztatók és Nyelvtanárok Lapja*. II/1.14–16. Budapest: ELTE ITK.
- Dávid Gergely (1992): Objektivitás, értékelés, standardizáció. *Iskolakultúra*. 2/1, 26–29.
- Davies, A. (1978): Language Testing (a survey article) in Kinsella, V. (ed.) *Language Teaching and Linguistics: Abstracts*. Vol. II. 3, 4, 145–159; 215–231. Cambridge: Cambridge University Press.
- Dörnyei Zoltán – N. Tóth Zsuzsa (1987): Megméretett – A magyar állami nyelvvizsgarendszer áttekintése. *Pedagógiai Szemle*. XXXVII. évfolyam, 11. szám, 1119–1132.
- Dróth Júlia (2001): A fordítások nyelvi megvalósításának értékelése a fordításoktatás folyamatában. *Nyelvi MÉRCE*. I/1–2, 30–36.
- Eckersley, C. E. (1938–): *Essential English for Foreign Students*. (Illustrations by Salibury, Ch. and Sharrocks, B.) Books I–IV. London: Longmans (Green and Co Ltd).
- Einhorn Ágnes (1999): Az idegen nyelvi érettségi reformjának eddigi eredményei. *Nyelvvizsga Fórum. Nyelvvizsgáztatók és Nyelvtanárok Lapja*. II.2. 51–52. Budapest: ELTE ITK.
- Fekete, H. – DalGLISH, R. – Hughes, M. (1993): Good exams, bad exams. *Testing Matters*. Budapest: State Language Examination Board. 97/II/1, 10–11.
- Fekete Hajnal (1996): Az akkreditáció kérdéséhez. *Testing Matters*. Budapest: State Language Examination Board, 96/I/1, 9–13.
- Fowler, W. S. – Coe, N. (1976): *English Language Tests*. Sunbury-on-Thames: Thomas Nelson and Sons Ltd.
- French, F. G. (1968): *English in Tables*. Oxford: OUP.
- Gatenby, E. V. – Eckersley, C. E. (1955): *General Service English Wall Pictures*. Teacher's Handbook with Exercises. London: Longmans.
- Gáborján Lászlóné – Horváth Iván (1990): *Tájékoztató az állami nyelvvizsgákról*. Budapest: Téka Könyvkiadó.
- Heltai Pál (1989): Elméretett... Válaszok Dörnyei Zoltán – N. Tóth Zsuzsa: Megméretett... c. tanulmányára. *Pedagógiai Szemle*, XXXIX évfolyam, 1. szám, 63–76.
- Heltai Pál (1989): Lehet-e fellebbezni? *Nyelvvizsga Híradó*/2, 5-8. Budapest: ELTE ITK.
- Heltai Pál (1990): Nyelvvizsga tévhitek. *Nyelvvizsga Híradó*.4. Budapest: ELTE ITK, 9–16.
- Heltai Pál (1996): A Világbank / British Council és a Goethe Intézet jelentéséről. *Testing Matters*, Budapest: State Language Examination Board, 96/I/1 1–8.
- Hock Ildikó (2000): On Threats to the Reliability of Oral Examinations. *Nyelvvizsga Fórum. Nyelvvizsgáztatók és Nyelvtanárok Lapja*. III/1. 45–46. Budapest: ELTE ITK.
- Katona Lucia (2000, szerk.): *Tájékoztató az ITK 'Origó' nyelvvizsgarendszeréről*. Budapest: ITK.
- Károly Krisztina (et al.) (2000): Diplomáciai szakmai nyelvhasználat. Nyelvhasználati igényfelmérés a diplomáciai szakmai nyelvvizsga korszerűsítéséhez. *Nyelvvizsga Fórum. Nyelvvizsgáztatók és Nyelvtanárok Lapja*. III/1. 27–36. Budapest: ELTE ITK.
- Király Zsolt (1997): Áldás, vagy átok? Az angol középfokú nyelvvizsga a tanár szemszögéből. *Testing Matters*. II/3, 27–32. Budapest: ELTE ITK.

- Kontráné Hegybiró Edit (2002): Szövetségest keres a nyelvtanár. *Nyelvi MÉRCE*. II/2. 98–100. Budapest: ELTE ITK.
- Kormos Judit (1997): A pragmatikai, szociolinvisztikai és szociokulturális kompetencia tesztelése a szerepjátékok segítségével. (Eredetileg: Play your cards right, *Novelty*, 1996) *Testing Matters*. II/2. 15–18. Budapest: ELTE ITK.
- Közös Európai Referenciakeret (KER) (2002): *nyelvtanulás, nyelvtanítás, értékelés*. Európa Tanács Közoktatási Bizottság. Élő Nyelvek Osztálya, Strasbourg. Budapest: OM Pedagógus-továbbképzési Módszertani és Információs Központ Kht.
- Kubinszky Mihály (1985): Az állami nyelvvizsgák tapasztalatai. *Felsőoktatási Szemle* XXXIV. évf. 3. szám. Budapest: az MM folyóirata.
- Kundt Ernő (1938-től folyamatosan): *Anglicizmusok*. Budapest: Aetheneum, Vajna György és Társa.
- Lado, R. (1964): *Language Teaching: A Scientific Approach*. New York: McGraw-Hill.
- Land, G. (1975, 1983): *What's in the News?* (A Selection of Newspaper Extracts with Exercises). London: Longman.
- Lozanov, G. (1979): *Suggestology and the Outlines of Suggestopedya*. New York: Gordon and Breach Science Publishers.
- Lukács Krisztina (2002): Nyelvoktatás Magyarországon – európai mércével. *Nyelvi MÉRCE*. II/2 5–15. Budapest: ELTE ITK.
- Mackey, W. F. (1965): *Language Teaching Analysis*. London: Longman.
- Major Éva (1999) A közép- és keleteurópai országok idegen nyelvi vizsgáztatási modelljei és az európai harmonizáció. *Nyelvvizsga Fórum. Nyelvvizsgáztatók és Nyelvtanárok Lapja*. II.2. 3–14. Budapest: ELTE ITK.
- McMordie, W. (1909): *English Idioms (and how to use them)*. (Revised by Goffin, R. C. in 1971, 15th impression) Ely House, London: OUP.
- Medgyes Péter (2011): *Aranykor. Nyelvoktatásunk két évtizede: 1989–2009*. Budapest: Nemzeti Tankönyvkiadó.
- Medgyes Péter (2015): *Töprengések a nyelvtanításról*. Budapest: ELTE Eötvös Kiadó.
- Mills, A. E. (1961-től folyamatosan): *Intermediate Test Papers in English*. Grammar and Vocabulary. London: Longman Group Limited.
- Nógrádi Gábor (1989): Nyelvvizsgáztatás alapfokon. *Élet és Irodalom* 1989. április 7.
- North, B. (2017): *The Common European Framework of Reference: Development, Theoretical and Practical Issues*. Lausanne, Switzerland: Eurocentres Foundation.
- Ogden, C. K. (1930): *Basic English: A General Introduction with Rules and Grammar*. First Edition, London: Kegan Paul, Trench, Trubner and Co.
- Öveges Enikő (1999): Helyzetkép az új angol érettségi fejlesztési munkálatairól. *Nyelvvizsga Fórum. Nyelvvizsgáztatók és Nyelvtanárok Lapja*. II.2. 53–55. Budapest: ELTE ITK.
- Phillip Kálmán (1911): *A direkt módszer. Tanulmányok*. Budapest: Franklin
- Radányi Tamás – Székács Györgyné (1981): *Angol II. Tanuljunk nyelveket!* Budapest: Tankönyvkiadó.
- Seidl, J. (1988): *English Idioms* (5th Edition) Oxford: OUP.
- Skehan, P. (1989): Language Testing: Survey Article. Part I. *Language Teaching Abstracts*. 21/2. 211–221. Cambridge: Cambridge University Press.

- Skehan, P. (1988): Language Testing: Survey Article. Part II. *Language Teaching Abstracts*. 22/1. 1–13. Cambridge: Cambridge University Press.
- Sipóczy Győző (1981, szerk.): *Tájékoztató az Állami Nyelvvizsgáról a 3/1980 (X. 25.) MM sz. rendelet alapján*. Budapest: TIT.
- Sipóczy Győző (1988): Egy kis statisztika. *Nyelvvizsga Híradó*. 88/1, 32–33. Budapest. ELTE ITK.
- Stone, L. (1970): *Tests in English Structure and Usage*. London, Macmillan.
- Szabó Gábor (1997): Item Response Theory in Language Testing. (Később megvédett PhD-értekezés.) *Testing Matters*. II/2, 9–11. Budapest: ELTE ITK.
- Szentiványi Ágnes (1990): *A nyelvvizsga, a vizsgázó, a vizsgáztató, a vizsga*. Budapest: Közgazdasági és Jogi Könyvkiadó.
- Tamássyné Bíró Magda (1998): Kell-e a középfokú nyelvvizsgán a gátakról is beszélgetni? *Nyelvvizsga Fórum. Nyelvvizsgáztatók és Nyelvtanárok Lapja*. II/1. 36–39. Budapest: ELTE ITK.
- Tarján Jenő – Korenchy Lajos (1965): *Angol nyelvkönyv haladók számára*. Tanuljunk nyelveket! Budapest: Tankönyvkiadó.
- Terestyéni Tamás (1980): *Idegennyelv-tudás a magyar lakosság körében*. Budapest, Tömegkommunikációs Kutatóközpont.
- Varga Zsolt (1989): Mégse fűtyülnek a Rigónak? *Mai Nap*. március 28.
- Véges István (1972): *Angol társalgási zsebkönyv*. Budapest: Tankönyvkiadó.
- Véges István (1983): *What about the News?* Budapest: Tankönyvkiadó.
- Worrall, A. J. (1964): *More English Idioms*. London: Longmans and Green Co. Limited.

BENKE ESZTER

Egy siKERtörténet: mit kapott az oktatáspolitikai és a nyelvoktatás a Közös európai referenciakerettől?

A success story: how has the Common European Framework of Reference for Languages informed education policy and language instruction?

With the appearance of the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR), a comprehensive and highly influential language education policy document has emerged which appears to have gained widespread recognition well beyond the borders of Europe. The original purpose of the CEFR was to support education policy reforms and thus contribute to the improvement of the quality of language instruction as well as to develop and strengthen professional dialogue between stakeholders in various areas of language teaching. Its recommendations have become a crucial element of language education, educational policy and have contributed to the emergence of a metalanguage that makes language learning, teaching and assessment coherent and transparent. The purpose of this paper is to provide a brief overview of some of the CEFR-related projects and toolkits that stemmed from the CEFR. The practical results of these projects seem to suggest that the CEFR is likely to inspire continued improvement in language teaching and learning and frame educational policy for years to come.

Keywords: CEFR, language education, metalanguage, coherence, transparency

Bevezetés

A Közös európai referenciakeret: nyelvtanulás, nyelvtanítás, értékelés (2002) angol (2001a) és francia (2001b) nyelvű változatának egyidejű megjelenése a Nyelvek Európai Évében hosszú, széles körű empirikus kutatáson alapuló munkát zárt le, és egyúttal a 21. század nyelvpedagógiai metanyelvének, gondolkodásának és eseményeinek is egyik legmeghatározóbb műve lett. Elgondolkodtató lehet a címben sikertörténetként definiált, nem szakmai diskurzusba illeszkedő kifejezés, ugyanakkor vitathatatlan, hogy az eredetileg európai kontextusra szánt keretrendszer nem ok nélkül vált népszerűvé Európán kívül is (Byram–Parmenter 2012) és olvasható több mint 40 nyelven. A Közös európai referenciakeret (KER) létrejöttével olyan meghatározó oktatáspolitikai dokumentum került az idegennyelv-pedagógiai színtérre, amelynek népszerűsége és gyakorlati haszna egyebek mellett a mű alapelveként kitűzött rugalmas, kontextusfüggő alkalmazásának, iránymutató jellegének, valamint a nyelvtudás dinamikus, cselekvésközpontú, „can do”, pozitív, tudásorientált megközelítésének köszönhető (Figueras 2012, North 2014). E „can do”, vagyis a „tud” állítások alkalmazásának felhasználásával

történő nyelvtudás-megközelítés eredetéről a KER új, kiegészítő kötete további részletekkel is szolgál (Council of Europe 2018: 32), gyakorlati alkalmazásáról és hasznáról pedig jó néhány szerző beszámol (lásd pl. North 2014: 109). A KER elméleti alapvetésére és gyakorlati útmutatására épülő dokumentumok, források és eszközök jelentős része az Európa Tanács Nyelvpolitikai Osztályának honlapján (<https://www.coe.int/en/web/common-european-framework-reference-languages/home>), részben pedig a grazi Európai Modern Nyelvek Központja KER tematikus oldalán (<https://www.ecml.at/Thematicareas/CEFRandELP/tabid/1935/language/en-GB/Default.aspx>) érhető el.

Az állítás, hogy a KER a nyelvoktatás, a nyelvtanulás, a nyelvpedagógiai diskurzus és gyakorlat meghatározó elemévé vált, előre jelzi azt is, hogy a mű nyelvtanulásra, nyelvoktatásra és oktatáspolitikára gyakorolt hatását reménytelen vállalkozás egy rövid tanulmányban szisztematikusan és átfogóan összefoglalni. Írásunkban mozaik jellegű, általános áttekintést kívánunk adni arról, hogy a KER-ben szereplő nyelvoktatás- és nyelvtudás-koncepció milyen elméleti és gyakorlati eredményei érzékelhetők és hasznosíthatók a nyelvoktatás-politikában.

A Közös európai referenciakeret: nyelvtanulás, nyelvtanítás, értékelés

Az 1991-ben „Átláthatóság és kohézió a nyelvtanulásban Európában” címmel a svájci Rüşchlikonban megrendezett kormányközi szimpóziumon vetődött fel olyan alapelvek kidolgozásának szükségessége, amelyek a nyelvoktatás minden érintettje számára érthetővé és átláthatóvá teszik a nyelvtudás szintjeit. Célként fogalmazódott meg az az elvárás, hogy váljon jól megfogalmazhatóvá, érthetővé és átláthatóvá a nyelvtanuló számára, hogy a nyelvelsajátítási folyamat során honnan hová tud vagy hová szükséges eljutnia, a nyelvoktató számára, hogy milyen lépcsőfokok, mennyiségi és minőségi tényezők azok, amelyek tevékenysége során nyelvoktatási célokként jelentkeznek, valamint a döntéshozó számára, hogy egy adott szinten milyen nyelvtudást várhat el.

A KER létrejöttének körülményei és célja nem egyszerűsíthető le csupán arra a tényre, hogy a nyelvoktatási színtéren is jelentkező globalizáció szükségessé tette egy új, átfogó, nyelvsemleges nyelvtudásprofil leíró keretrendszer kidolgozását. A mű összegezte az Európa Tanács nyelvpolitikai tevékenységének mintegy 30 évet átívelő eredményeit, valamint a nyelvelsajátítás és nyelvoktatás területén végzett kutatások hatását (Green 2012). Alapvető célkitűzésként jelentkezett az igény a közös hivatkozási pont kidolgozására tananyagírók, tantervkészítők és vizsgáztatók számára. További cél volt globális és lokális

oktatáspolitikai reformok ösztönzése és támogatása, valamint az abban érintettek párbeszédének elősegítése is (North 2014). A KER felhasználásának fókuszja a megjelenése óta eltelt közel két évtized alatt kétségtelenül az értékelés irányába tolódott el, ahogyan ezt az erre vonatkozó felmérések is igazolják (North 2014), de fontos hangsúlyozni, hogy a referenciakeret eredeti és alapvető célját a szerzők a nyelvtanulás és nyelvtanítás támogatásában látták egy átfogó, átlátható és következetes keretrendszer létrehozásával (North 2000).

Az 1975-ben elsőként kidolgozott Küszöbszint (Ek 1975), amely idővel felfelé és lefelé is további szintekkel bővült, egy hatfokú szintrendszer kiindulásaként szolgált. A KER leggyakrabban idézett 53 szintleíró skálája mellett nem kevésbé érdekesek és hasznosak a teljes mű kilenc fejezetének és négy függelékének további tartalmi elemei sem. Az első fejezet meghatározza a Közös európai referenciakeret funkcióját, és definiálja szerepét az Európa Tanács nyelvpolitikai célkitűzéseivel összhangban. A következő fejezet bemutatja azt a cselekvésközpontú szemléletet, amelyet a nyelvvoktatás és a nyelvtanulás leírása során a KER alkalmaz. Ez a fejezet részletes leírását és magyarázatát adja azoknak a nyelvhasználati kategóriáknak, amelyeket a KER referenciaként alkalmaz. A harmadik fejezet tartalmazza a közös referenciaszintek részletes bemutatását, a negyedik és ötödik fejezet pedig a KER általános elveként alkalmazott cselekvésközpontú megközelítéssel, szemléltető skálák segítségével illusztrálja a nyelvhasználói kompetenciákat, nyelvi tevékenységeket, mikrokészségeket. A hatodik fejezet nyelvtanulási és nyelvvoktatási kérdéseket tárgyal, de kitekintést nyújt általános, nem nyelvspecifikus kompetenciák tanítására is. A korábbiakban alkalmazott megközelítéshez hasonlóan, egy-egy téma megtárgyalása után a leírtak tükrében a KER az olvasót saját gyakorlatának átgondolására, önreflexióra ösztönzi a felvetett kérdésekkel kapcsolatosan. A hetedik fejezet a nyelvvoktatás során használt feladatokra fókuszál: ezek típusait, célját, valamint a nyelvtanulói stratégiákat és a feladatmegoldás körülményeit és kritériumait vázolja fel. A nyolcadik fejezet középpontjában a tanterv, a kilencedikben pedig a nyelvtudásmérés áll. A négy függelék közül az első kettő a szintleíró skálák kidolgozását és jellemzőit mutatja be, az utolsó kettő pedig további szintleíró skálákat tartalmaz: a DIALANG és az ALTE skálákat.

A 2001-ben végleges formát öltött mű érezhetően és minden valószínűség szerint szándékoltan hagyott teret további bővítésnek. A Közös európai referenciakeret közel húszéves használata során szerzett tapasztalatok, illetve a nyelvhasználati igények a KER-skálák módosítását, kibővítését, valamint újabb nyelvi tevékenységek leírását tették szükségessé. Az öt szakaszból álló, 2014 és 2017 között megvalósult projekt eredetileg a KER-ben negyedik nyelvi tevékenységként szereplő, de leíró skálákkal nem rendelkező közvetítés kidolgozását tűzte ki célként. A három nyelvi alaptévékenység, a produkció, recepció és interakció mellett a közvetítés komplex készségként jóval összetettebb megközelítést igényel, mint az eredeti műben kidolgozott nyelvi

tevékenységek (North–Piccardo 2016, Piccardo–North 2019). Az új és a cikk írásakor mindössze angol nyelven elérhető kiegészítő kötetben, mely a „Common European Framework of Reference for Languages, Companion volume with new descriptors (Council of Europe 2018)” címet viseli, a közvetítési tevékenységeket jellemző szintleíró skálák mellett további deskriptorok is helyet kaptak. Ezek az Európai Modern Nyelvek Központja ProSign projekt keretében született jelnyelvi deskriptorok, továbbá a 21. századi digitális környezetben népszerűvé és elterjedté vált digitális kommunikációra, valamint az együttműködésen alapú tevékenységek végzésére vonatkozó skálák. A 2018-ban megjelent kötet a teljességgel új és a csekély mértékben módosított deskriptorokon kívül újabb szinteket is felvonultat: ilyen például a pre-A1 vagy A1-et megelőző szint, illetve az újabb plusz jelölésű szintek. Újdonság még a kötetben nem szereplő, de a honlapon elérhető deskriptorgyűjtemény, amely a projekt keretein belül a fiatal nyelvtanulók 7–10 és 11–15 éves korcsoportjai számára készült. Különösen érdekesek és értékesek azok az új elemek, amelyek a közvetítési tevékenységeken keresztül, valamint a többnyelvű és plurikulturális kompetencia leírása segítségével tükrözik a KER nyelvi és kulturális sokszínűségén is alapuló és azt hirdető értékrendjét. Ezek a skálák a multikulturális környezetben történő hatékony kommunikációhoz szükséges eszköztár szintjeit mutatják be. Az új szintleíró skálák kidolgozásában jóval több szakértő vett részt, mint az eredeti projektben. Fontos megjegyezni, hogy több magyarországi intézmény nyelvoktatói és nyelvtudásmérői is részt vettek az új deskriptorokat kidolgozó szakértői munkában.

A Közös európai referenciakeretre épülő dokumentumok és segédanyagok

Szintharmonizációs és a szintillesztési kézikönyv

A globális célként meghatározott referenciakeret szerepével összhangban a KER egyik legmarkánsabb hatása a hivatkozási rendszer vagy viszonyítási szisztéma funkciójához kapcsolódik. A KER, és azon belül is a szintleíró skálák megjelenése az oktatáspolitikai és a nyelvoktatáshoz kötődő tevékenységek számos területén törekvéseket indított meg összehasonlításra, harmonizálásra, összekötésre, illesztésre, viszonyításra, megfeleltetésre, egyenértékűsítésre, amely a mai napig, változó intenzitással ugyan, de rendkívül széleskörűen és változatos fókusszal folytatódik. A kevésbé szisztematikus módon meghatározott szintmegjelölés vagy szintbesorolás a nyelvoktatásban használt tankönyvek, tananyagok és kurzusok esetében rendkívül gyorsan elterjedt és igen jól érzékelhető, ugyanakkor a szintillesztés empirikus formában történő megvalósítása elsősorban a nyelvvizsgák szintjeihez köthető. A számos szinonima nem véletlen, hiszen a megfeleltetés kifejezést még akkor is körültekintéssel kell kezelni, ha az illesztési tevékenység a KER kiegészítő dokumentumaként született szintillesztési kézikönyv

(Council of Europe 2003, 2009a, 2009b) előírásait követve történik. Mind a KER-szintek és a nyelvvizsgaszintek, mind pedig a KER és az egyéb létező szintleíró skálák közötti harmonizáció és megfeleltetés iránt folyamatos a szakmai érdeklődés és az igény Európában és Európán kívül is. A szintillesztési kézikönyv próbaváltozatának megjelenésével együtt kezdődtek meg a szintillesztési munkálatok (Martyniuk 2010), melyek eredményeinek bemutatása folyamatos (Green 2018). A legnagyobb vizsgaszolgáltatók, mint például az IELTS és a TOEFL, az elsők között kezdték meg a szintmegfeleltetésre irányuló kutató-fejlesztő tevékenységeiket (Tannenbaum–Wylie 2008, Taylor 2004, Taylor–Jones 2006), de a jóval kisebb vizsgapopulációval rendelkező, nemzetközi vagy akár csak nemzeti vizsgázói bázissal rendelkező vizsgaközpontok is hamar követték a legnagyobbak példáját. A szintillesztés, ahogyan az erre vonatkozó szakirodalom sokasága ezt igazolja, nem korlátozódik egy vizsgatípusra: a mérés céljától, formájától és tétjétől függetlenül tapasztalható a törekvés arra, hogy a teszt tulajdonságait és eredményét a KER szintbesorolásaira hivatkozva és azokkal összhangban lehessen közzétenni és értelmezni.

Magyarországon a nyelvvizsgák akkreditációját 1998-ban törvénybe iktató 71/1998. (IV. 8.) kormányrendelet bevezetése után a nyelvvizsgarendszereiket elsőként akkreditáltató intézmények követik az „Államilag elismert nyelvvizsgák akkreditációjának kézikönyvét” (Lengyel 1999), amely ugyan lehetőséget ad arra, hogy az adott nyelvvizsgarendszer önállóan határozza meg a KER-hez viszonyítva nyelvvizsgái szintjeit, ugyanakkor definiálja is „az államilag elismert nyelvvizsgákon” használatos szintrendszert. Ennek értelmében a középfok lefedi a B2 tartományt, és átnyúlik a C1 alsó tartományába, a felsőfok pedig a C1 felső tartományát és a C2 szintet foglalja magában. Fontosnak tartjuk megjegyezni, hogy míg az Akkreditációs Kézikönyv első kiadása 1999-ben került kiadásra, a Közös európai referenciakeret hivatalos, nyomtatott formában csak 2001-ben vált mindenki számára elérhetővé, vagyis a mű végleges változata akkor jelent meg, mikor Magyarországon jó néhány nyelvvizsgarendszer akkreditációja már megtörtént. Meg kell ugyanakkor azt is jegyezni, hogy a KER 1996-tól elérhető munkaanyagát nagyon sok szervezet és projekt felhasználta.

Magyarországon a felsőoktatási törvényben 2005-ben jelent meg (2005. évi CXXXIX. törvény a felsőoktatásról [XII. 14.]) az az elvárás, hogy a nyelvvizsgaszinteket a Közös európai referenciakeret szintjeinek kell megfeleltetni, illetve a hivatalos dokumentumokban a korábbi alap-, közép- és felsőfok szintmegnevezéseket B1, B2 és C1 szinttel kell helyettesíteni. Ez a lista 2014-ben bővült az opcionálisan akkreditálható A2 szinttel. A 2004 decemberében napvilágot látott új Akkreditációs Kézikönyv (AKK, 2004) ugyanakkor már egy az egyben megfelelteti a magyarországi nyelvvizsgaszinteket az Európa Tanács-skála három szintjének. A nyelvvizsgák minőségbiztosításával kapcsolatos részletes elvárásokat tartalmazó Akkreditációs Kézikönyv ezt a kívánalmat a következőképpen fogalmazta meg: „Az államilag elismert nyelvvizsgák rendszere

összhangot kíván teremteni a nyelvvizsgáztatás hazai és nemzetközi (európai) elmélete és gyakorlata között. Ez a törekvés abban is kifejezésre jut, hogy a meghatározott kedvezményeket és jogosultságokat biztosító háromszintű hazai rendszer és az ET munkabizottságának a nyelvtudás színtezésére vonatkozó, részletesebb tagolást javasoló, korszerű szemléletmódja közötti átjárás biztosított (ez a folyamat nem tekinthető lezártnak)” (AKK 8. old, 25. old.).

Az akkreditált nyelvvizsgák színtillesztésére irányuló kutató-fejlesztő tevékenységek Magyarországon is már a 2000-es évek közepén megkezdődtek, mivel az Európa Tanács Nyelvpolitikai Osztálya által irányított programhoz, a színtillesztési kézikönyv kipróbálásához több magyarországi nyelvvizsgaközpont is csatlakozott (Dávid 2012, Szabó 2011). Fontos megjegyeznünk azt is, hogy a kézikönyv kidolgozását koordináló tudományos operatív bizottságnak is voltak magyar résztvevői, valamint azt, hogy a NYESZE, a British Council és a Nyelvvizsgáztatók és Nyelvtudásmérési Szakértők Országos Egyesülete szervezésével a tapasztalatok megosztására egy budapesti konferencián is lehetőség nyílt 2006-ban.

Összhangban a nemzetközi nyelvtudásmérési szintéren folyó tevékenységekkel, a KER-szintek megjelenítése a nyelvvizsgáztatásban, illetve a nyelvvizsgák színtillesztése olyan mértékben prioritássá vált a magyarországi nyelvvizsgáztatásban is, hogy kormányrendelet (137/2008. [V. 16.]) is kimondta/szabályozta a színtillesztés szükségességét. A KER-szintek frissen tartását és a harmonizáció folyamatosságát a tematikus akkreditációs felülvizsgálatok és a két évente esedékes akkreditáció megújítása során a nyelvvizsgáztatást felügyelő kormányzati szerv, az Oktatási Hivatal az érvényben lévő Akkreditációs Kézikönyv elvárásainak megfelelően számon kéri.

A színtillesztés egyes fázisainak megkönnyítését hivatott segíteni számos olyan nemzetközi projekt, amelynek eredményei a legtöbb esetben még ma is használhatók. Ilyen volt a több ország nyelvtudásmérési szakértőinek részvételével megvalósult Dutch Grid projekt, amely a színtillesztés úgynevezett specifikációs fázisában nyújt segítséget az olvasott szöveg értése és a hallott szöveg értése tesztek specifikációinak elemzéséhez, illetve KER-szintekhez történő illesztéséhez. Az online felületen a felhasználó által feltöltött tesztek szövegeit és itemeit lehet a specifikációkban meghatározott tulajdonságok mentén elemezni a KER-rel összhangban lévő kategóriák alapján. Ugyan a Színtillesztési Kézikönyv is tartalmaz táblázatokat, amelyek a teszt-elemzést megkönnyítik, a Dutch Grid részletesebb tesztleírást tesz lehetővé az előre meghatározott, a specifikációkban is használt tulajdonságok és kategóriák segítségével.

Ahogy a Dutch Grid is hangsúlyozza, a KER-harmonizáció abban az esetben lehet hatékony és megbízható, ha az azt végző szakértők megfelelő ismeretekkel rendelkeznek a KER-szintekről. A familiarizációs gyakorlatok célja és részletes leírása a Színtillesztési Kézikönyvben (2003) olvasható, és a segédanyag első kiadásának

megjelenése óta számos változata is született az ott leírt eljárásoknak. Ilyen például a szintekhez illesztett nyelvhasználói teljesítményeket is bemutató CEFTrain (<http://www.helsinki.fi/project/ceftrain/index.html>) online felület is. A minden érdeklődő felhasználó számára szabadon elérhető magyar KER familiarizációs gyakorlatokat is tartalmazó KEROnline alkalmazás (<http://www.keronline.hu/>) az OH NYAK kezdeményezésére készült azzal a céllal, hogy a magyarországi nyelvvizsgáztatók számára is könnyen elérhető, felhasználóbarát formában rendelkezésre álljon egy KER-ismerteket frissen tartó gyakorlási lehetőség. A KER-skálákkal végzett familiarizáción túl angol, francia és német nyelvű, kipróbált és megfelelő beválási mutatókkal rendelkező feladatsorokon és vizsgáztatói teljesítményeken nyílik lehetőség a szintezés gyakorlására.

A szintillesztés, a KER-harmonizáció eljárásai és eredményei számos kérdést vetnek fel. A felvetett kritikák elemzésekor érdemes Green (2018) álláspontját szem előtt tartani, mely szerint a KER csak támpontként szolgálhat az eltérő tartalmú tesztek összehasonlítása esetében. Egy eljárás és eredményének értelmezése során a felelősség megoszlik (North 2014). A szintmeghatározási módszerek sokasága (Cizek 2001, Cizek–Bunch 2007, Kaftandjieva 2004) és az ezek megvalósítása során alkalmazott széles körű információn alapuló, ámde mégis szubjektív emberi értékítélet megbízhatóságának kérdése (Harsch–Hartig 2015, Jaeger 1992, Kahneman és mtsai 1982, Papa-georgiou 2010) még hosszú ideig biztosít megválaszolendő kérdéseket és megoldandó problémákat elméleti és gyakorlati szakemberek számára.

DIALANG: diagnosztikus önértékelés

A Közös európai referenciakeret „C” függelékében található DIALANG-skálák (KER 286–299) önértékelés segítségével adnak visszajelzést a nyelvtanuló nyelvtudásszintjéről. A 25 oktatási intézmény részvételével és európai uniós támogatással készült online DIALANG-platform (<https://dialangweb.lancaster.ac.uk/>) 14 nyelven (angol, dán, finn, francia, görög, holland, ír, izlandi, német, norvég, olasz, portugál, spanyol, svéd) és 18 különböző nyelvű instrukcióval kínál készségenkénti diagnosztikus nyelvtudásmérést. A felhasználó által kiválasztott készség mérését szókinccsteszt és önértékelés előzi meg, melyek eredménye meghatározza a komputer-adaptív teszt kezdőszintjét. A nyelvtudásprofil az alkalmazás a következő területeken tudja meghatározni: hallott szöveg értése, írás, olvasott szöveg értése, nyelvhelyesség, szókinccs. Lehetőség nyílik itemenkénti, azonnali visszajelzés kérésére, illetve a teszt végén az összesített eredmény megtekintésére. Ez utóbbi esetben is áttekinthetők a válaszadó megoldásai, és összevethetők a helyes megoldásokkal. A diagnosztikus mérés céljának megfelelően a program visszajelzést ad a nyelvtudásszintről, a fejlesztendő területekről, valamint lehetőséget kínál arra, hogy az önértékelés és a valós teljesítmény alapján megállapított nyelvtudásszint közötti különbség okait a nyelvtanuló megvizsgálja.

A DIALANG önértékelési skálák alapjául a KER-deszkriptorok szolgálták, amelyeket a diagnosztikus és önértékelési célnak megfelelően módosított a projekten dolgozó szakértői csoport (Alderson 2005, Alderson–Huhta 2004, Huhta és mtsai 2002, KER 2002, Luoma 2004). E diagnosztikus teszt előnyeiről és gyengeségeiről számos tanulmány született az elmúlt mintegy másfél évtizedben (Chapelle 2005, Harding és mtsai 2015). A készségek alkésszükszintű mérhetőségéről megfogalmazott, eltérő eredményeket hozó empirikus kutatásokon alapuló álláspontok, vagy a gépi, mechanikus és uniformizált visszajelzés érvényessége és hatékonysága az online diagnosztikus tesztelés esetében jogosan vet fel megválaszolendő kérdéseket. Ugyanakkor az ingyenesen, immár webes platformon elérhető, KER-alapú diagnosztikus mérőeszköz méltán kerülhet fel a többi, hasonló funkciójú diagnosztikus teszt listájára és a nyelvoktatók eszköztárába. A webalapú, komputeradaptív DIALANG-teszt kiindulási alap és elmozdulási kísérlet lehet a dinamikus tesztelés (Feuerstein és mtsai 2002, Lantolf–Poehner 2004, Poehner–Lantolf 2005) irányába.

English Profile Network

A KER a létrehozók szándéka szerint nyelvsemleges, vagyis konkrét nyelvtől függetlenül is felhasználható számos nyelvoktatási kontextusban, ugyanakkor az eredeti műből kiindulva több olyan dokumentum és online alkalmazás kidolgozásának is tanúi lehetnek a felhasználók, amelyek egy adott nyelvhez vagy speciális kontextusokhoz alkalmazkodnak. Példaként említhetjük a német, osztrák és svájci együttműködés eredményeként létrejött Profile Deutschot (Glaboniat és mtsai 2002, 2005), amely a KER-skálák alapján és felhasználásával hozta létre a német nyelv közös referenciakeretét, és amely a szintleíró skálákon túl nyelvhelyességi leírásokat is rendel az egyes szintekhez (Víg 2005). De a németen kívül léteznek további, konkrét nyelvekhez kapcsolódó szintleírások RLD – Reference Level Descriptions elnevezéssel, amelyek a közös útmutató (Council of Europe 2005) alapján készültek. A horvát, cseh, francia, grúz, olasz, litván, portugál, spanyol és török nyelvből elkészült anyagoknál az egyes szintek vagy az összes szint elérhető, de további nyelvekből is készülnek a nyelvspecifikus szintleírások. Ezek a leírások a KER-skálákból kiindulva nyelvi szerkezetekkel, szókinccsel és nyelvhasználati funkciókkal is jellemzik a szinteket.

Hasonló elvek alapján folyik a munka az English Profile Network (Green 2010, Harrison–Barker 2015, Saville 2010, Saville–Hawkey 2010) fejlesztése körül, amely több írott és beszélt nyelvi korpusz (pl. Cambridge International Corpus, Cambridge Learner Corpus) elemzésével és felhasználásával rendel szókinccset, nyelvi szerkezeteket és nyelvhasználati funkciókat az egyes KER-szintekhez. A jelenleg mindenki számára ingyenesen elérhető English Vocabulary Profile brit és amerikai adatbázisa szavak szintbesorolását adja meg a felhasznált korpuszokból származó példákkal együtt. Van lehetőség szavakra keresni, amelyeket szótári és nyelvtanulói példák illusztrálnak. Az English Grammar Profile a szókinccskorpuszhoz hasonló módon a

KER-szintekhez kapcsol olyan nyelvtani szerkezeteket, amelyekről az empirikus adatok alapján elmondható, hogy az adott szint nyelvhasználói alkalmazzák (McCarthy 2016). Az alkalmazás példaként a „may” használatán keresztül mutatja be, hogy a Vocabulary Profile-hoz hasonló módon egyes szerkezetek jelentéstől függően több KER-szinten is megjelenhetnek.

- A2 Weak possibility: „The weather may be hot.”
- B1 Formal permission: „May I borrow your bike?”
- C1 'May well': „You may well find that this is not the case.”
- C2 'May as well': „We may as well go home.”

Forrás: <https://www.englishprofile.org/english-grammar-profile>

A funkcionális nyelvhasználat tárháza (English Functions Profile) jelenleg még fejlesztés alatt áll; a tervek szerint a Vocabulary és a Grammar Profile-hoz hasonló módon nyelvhasználói példák fogják illusztrálni az egyes KER-szintekre jellemző funkcionális nyelvhasználatot, kiegészítve a módosított és átdolgozott „can do” állításokkal (Green 2012).

Az English Profile hangsúlyozza leíró jellegét, nevezetesen azt, hogy adott szinten mit tudnak a nyelvtanulók, szemben azzal, hogy mit kell tudniuk. Ezzel összhangban határozzák meg az alkotók az adatbázis célját is, amely empirikus kutatáson alapuló háttér-információ biztosítása többek között tananyagírás, tantervkidolgozás és további nyelvoktatási tevékenységek támogatására (McCarthy 2016).

Európai nyelvi portfólió és a Europass

A nyelvi útlevelekből, dossziékból és életrajzból álló Európai nyelvi portfólió 2001-es, a KER-rel közel egyidejű megjelenését hatalmas érdeklődés övezte. Sorra jöttek létre a különböző nyelvoktatási szintekre és különböző nyelveken kidolgozott, validált, akkreditált és regisztrált portfóliók, amelyek népszerűsége és használata mára némileg alábbhagyott. Továbbra is népszerű ugyanakkor az egyik fő alkotóelem, az önértékelési rendszer, amelyek a portfólió újabb változataiban is az önálló tanulást, az önreflexiót és a tanulás menedzselésének fontosságát népszerűsíti. A portfólió felhasználását, alkalmazásának előnyeit és eredményeit bemutató kutatások (Little 2009a, 2012, 2016) gyakran hivatkoznak az autonóm tanulás első és a mai napig az egyik leggyakrabban idézett kutatójára és népszerűsítőjére (Holec 2001), akinek az önálló tanulóval kapcsolatos megállapításai igen nagyfokú hasonlóságot mutatnak az Európai nyelvi portfólió autonóm tanulást népszerűsíteni hivatott pedagógiai célkitűzéseivel (Council of Europe 2011: 5, 12, Little 2009b). Fontos célja még a portfóliónak az önértékelésen kívül az interkulturális érzékenység, tudatosság és az ehhez szükséges kompetenciák fejlesztésének elősegítése. Több tanulmány igazolja, hogy a portfólió és

annak elektronikus változata (Kühn 2016) a felsőoktatásban bizonyos képzési területeken a mai napig népszerű (Argondizzo–Sasso 2016, Carson 2016, Dalziel és mtsai 2016).

Az Európai nyelvi portfólióhoz kapcsolódik a 29 nyelven rendelkezésre álló Europass dokumentumgyűjtemény, amelyet az Európai Parlament és Tanács a képesítések és a szakmai alkalmasság átláthatósága és elismertetése céljából hozott létre (2004). A dokumentumgyűjtemény részei a következők: önéletrajz, valamint az Európai készségi levél, amelynek részét képezi a nyelvi útlevél, a mobilitási igazolvány, az oklevélmelléklet és a bizonyítványkiegészítő. Az önéletrajz idegennyelv-tudásra vonatkozó részét az önértékelési skálák segítségével lehet kitölteni. Az önéletrajz mellett a Nyelvi útlevél részletes nyelvtudásprofil ad a felhasználó nyelvismeretéről. Az Europass alkalmazás fejlesztése folyamatos: a mindenki számára elérhető, online formában tárolható és módosítható dokumentumgyűjtemény sablonjai a képesítések és kompetenciák változásával összhangban bővülnek. A KER „can do” elve alapján készült a például az Europass részeként felhasználható, digitális kompetenciákat bemutató önértékelő keretrendszer is (European Union 2015). Érdemes megemlíteni az Europass alkalmazás levélszerkesztő funkcióját is, amely beépített és kiválasztható kifejezésgyűjteménnyel nyújt segítséget az önéletrajzot kísérő motivációs levél megírásában a nyelvtanulónak/munkavállalónak. Az Europass népszerűségét igazolja, hogy a honlap statisztikái szerint 2005 óta 128 483 624 online dokumentum készült az alkalmazás segítségével (Cedefop 2019).

Az European Profiling Grid és a nyelvtanári kompetenciák

A KER népszerűvé és elfogadottá vált szerkezete és felépítése jelentős mértékben hozzájárulhatott más, hasonló kompetenciaszint-leíró keretrendszerek létrejöttéhez is. Ezek közül nyelvoktatók számára az egyik legérdekesebb az EAQUALS minőségbiztosítási keretét alapul felhasználó (North 2009), az Európai Bizottság támogatásával, valamint a Centre international d'études pédagogiques (CIEP) irányításával kidolgozott *European Profiling Grid (EPG)*. A jelenleg 13 nyelven elérhető keretrendszer a nyelvoktatói kompetenciák leírását adja a KER keretrendszerének alapelveire építve. A nyelvtanári kompetenciákat bemutató deskriptorok ennél az eszköznél is a „tud” állítások, vagyis a „can do” alapelv szerint épülnek fel, és a három fő „fejlődési szakasz” alkategóriái és szintjei mennyiségi és minőségi növekedést írnak le. A négy tanári kompetencia-profil (a képzés és képesítés/végzettség, a tanári kulskompetenciák, alap/támogató-kompetenciák és szakmai felkészültség) további alkompetenciákon keresztül mutatja be azokat a fázisokat, amelyeken keresztül az oktató a szakmai tapasztalatai és továbbképzés eredményeként előre tud lépni tanári kompetenciái fejlesztésében (Rossner 2018). Az eszköz mind letölthető, táblázatos, mind pedig online formában elérhető, és lehetővé teszi a tanári önreflexiót és önértékelést (European Profiling

Grid 2011). A tanári kompetenciaprofil nem feltétlenül és szükségszerűen egységes minden kategória szempontjából. A felhasználó elhelyezheti saját kompetenciáit a keretrendszerben, és útmutatást kap arra vonatkozóan, hogy mely irányban érdemes kompetenciáit fejleszteni.

Bár az EPG alkalmazásának gyakorisága egyelőre még távolról sem közelíti meg a mintegy tíz évvel korábban született KER népszerűségét, számos tanulmány erősíti meg pozitív hatását a tanári önértékelésre, önreflexióra és szakmai fejlődésre (lásd pl. Linaker 2018, Rossner 2017). Az oktatói tevékenységek a kontextustól függően szer-teágazóak, így a tanári tevékenységek palettája is igen sokszínű. Az EPG természetesen nem az egyetlen olyan skála, amely részletes leírását adja a tanári kompetenciáknak. Az Európai Modern Nyelvek Központja „Közös Európai Referenciakeret tanárok számára” munkacsoport weboldalán (ECML, n. d.) mintegy 50 különböző tanári kompetencia-keretrendszerbe kínál betekintést, amelyek ugyan némileg eltérő fókusszal, de azonos céllal, az EPG-hez hasonlóan a tanárok folyamatos szakmai fejlődésének motorjaként/mozgatórúgójaként működhetnek.

Összefoglalás

Mit kapott az oktatáspolitikai és a nyelvoktatás a Közös európai referenciakerettől? Szinte észrevétlenül kialakult és természetessé vált az a közös metanyelv, melynek a KER szolgált forrásul és alapjául. A hosszú időkre visszanyúló vita kapcsán, hogy mennyiben alakítja a nyelv a gondolkodást, a KER hatására kialakult metanyelv esetében meglehetősen nagy biztonsággal állíthatjuk, hogy igen nagy mértékben meghatározza nyelvoktatással kapcsolatos gondolkodásunkat, nyelvpolitikai tevékenységeinket. A KER javaslatainak tudatosságnövelő szerepe a nyelvoktatás számos területén megkérdőjelezhetetlen. A KER felhasználásának lehetőségeit, illetve gyakorlati alkalmazásának eredményeit vizsgáló és értékelő szakmai közösségek párbeszéde megtermékenyítőleg hat a nem kutató felhasználók gondolkodására is.

Fontos ugyanakkor szem előtt tartani a referenciakeret szerzői által gyakran hangoztatott alapelvet és célkitűzést: a KER javaslatokat ad, nem előír. Eredeti célja oktatáspolitikai reformok támogatása volt, amely hozzájárulhat a nyelvoktatás minőségének fejlesztéséhez és a szakmai párbeszéd kialakulásához és erősödéséhez a nyelvoktatás különböző színtereinek érintettjei között. A felmerülő nehézségek, problémák és gyengeségek mellett North (2014: 231) olyan közös referenciakeret létrejöttét is felveti, amelynek alkalmazhatósága az Európán kívüli kontinenseket is célzottan lefedi.

Nem könnyű megjósolni, mit hoz a jövő, de továbblépésre mindig számos lehetőség kínálkozik. Az elmúlt húsz év meghatározó szerepet jelölt ki a KER számára a nyelvoktatás minőségbiztosításának alapvető pilléréként a plurilingvális beszélő közösségek és a multikulturális világban történő boldogulás érdekében. A KER igen nagy valószínűséggel a jövőben is fontos szerepet fog betölteni a nyelvoktatás területén történő fejlődés

mozgatórugójaként, katalizátoraként, útmutatójaként. A felhasználó és a döntéshozó felelőssége, hogy az abban foglalt javaslatokat hogyan ülteti át a gyakorlatba, és milyen célokra használja az ezekre épülő döntéseket.

IRODALOM

- Alderson, J. C. (2005): *Diagnosing foreign language proficiency: the interface between learning and assessment*. London; New York, NY: Continuum.
- Alderson, J. C. – Huhta, A. (2005): The development of a suite of computer-based diagnostic tests based on the Common European Framework. *Language Testing*, 22/4, 301–320.
- Argondizzo, C. – Sasso, M. I. (2016): The ELP through time: Background motivation, growing experience, current beliefs. *Language Learning in Higher Education*, 6/2, 377–395.
- Byram, M. – Parmenter, L. (2012): *The Common European Framework of Reference: the globalisation of language education policy* (Languages, Intercultural Communication and Education). Clevedon: Multilingual Matters.
- Carson, L. (2016): Fostering engagement with the Common European Framework of Reference for Languages and the European Language Portfolio: Learning from good practice in university language centres. *Language Learning in Higher Education*, 6/2, 275–282.
- Chapelle, C. A. (2006): Test review. *Language Testing*, 23/4, 544–550.
- Cedefop, 2019 <https://europass.cedefop.europa.eu/resources/statistics/custom-reports#/generated> CEFTrain <http://www.helsinki.fi/project/ceftrain/index.html>
- Cizek, G. J. (2001, ed.): *Setting performance standards: Concepts, methods and perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Cizek, G. J. – Bunch, M. (2007): *Standard-setting. A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage.
- Council of Europe (2001a): *Common European framework of reference for languages: learning, teaching assessment*. Cambridge: Cambridge University Press.
- Council of Europe (2001b): *Cadre européen commun de référence pour les langues: Apprendre, enseigner, évaluer*. Paris: Didier.
- Council of Europe (2003): *Relating language examinations to the common European framework of reference for languages: Preliminary pilot manual*. Strasbourg, France: Language Policy Division.
- Council of Europe (2005): *Reference level descriptions for national and regional languages (rld) draft guide for the production of RLD*. Version 2. Strasbourg, France: Language Policy Division.
- Council of Europe (2009a): *Reference supplement to the manual for relating language examinations to the CEFR*. Strasbourg, France: Language Policy Division.
- Council of Europe (2009b): *Relating language examinations to the common European framework of reference for languages: learning, teaching, assessment (CEFR)*. A manual. Strasbourg, France: Language Policy Division.
- Council of Europe (2011): *European Language Portfolio (ELP): Principles and Guidelines*. With added explanatory notes. Strasbourg: Council of Europe. Version 2. <https://rm.coe.int/16804586ba>

- Council of Europe (2018): Common European framework of reference for languages, Companion volume with new descriptors, Strasbourg: Council of Europe. <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>
- Dávid Gergely (2012): A szintleírások nyelvének szerepe a Közös Európai Referenciakeret magyar, angol és német nyelvű kiadásában. *Magyar Pedagógia*, 112/1, 19–39.
- Dalziel, F. – Davies, G. – Han, A. (2016): Using the ELP as a basis for self- and peer assessment when selecting “best” work in modern-language degree programmes. *Language Learning in Higher Education*, 6/2, 397–417.
- Európai Parlament és Tanács (2004): Europass, Elérhető: <https://europass.cedefop.europa.eu/sites/default/files/europassdecision-hu.pdf>
- European Profilig Grid: <http://www.epg-project.eu/grid/>
- European Union (2015): Europass, <https://europass.cedefop.europa.eu/sites/default/files/dc-en.pdf>
- Feuerstein, R., Feuerstein, S., Falik, L – Rand, Y. (2002): *Dynamic assessments of cognitive modifiability*. Jerusalem, Israel: ICELP Press.
- Figueras, N. (2012): The impact of the CEFR. *ELT Journal*, 66/4, 477–485.
- Glaboniat, M. és mtsai (2002): *Profile deutsch. Gemeinsamer europäischer Referenzrahmen*. Langenscheidt, Berlin.
- Glaboniat, M. és mtsai (2005): *Profile Deutsch A1-C2 (Version 2.0)* Langenscheidt, Berlin.
- Green, A. (2010): Requirements for Reference Level Descriptions for English. *English Profile Journal*, 1.
- Green, A. (2012): *Language Functions Revisited: Theoretical and Empirical Bases For Language Construct Definition Across the Ability Range*. Cambridge: UCLES/Cambridge University Press.
- Green, A. (2018): Linking Tests of English for Academic Purposes to the CEFR: The Score User’s Perspective. *Language Assessment Quarterly*, 15/1, 59–74.
- Harding, L. – Alderson, J. C. – Brunfaut, T. (2015): Diagnostic assessment of reading and listening in a second or foreign language: Elaborating on diagnostic principles. *Language Testing*, 32/3, 317–336.
- Harrison, J. – Barker, F. (eds.) (2015): *English Profile in practice*. English Profile Studies. Cambridge, United Kingdom: Cambridge University Press.
- Harsch, C. – Hartig, J. (2015): What Are We Aligning Tests to When We Report Test Alignment to the CEFR?. *Language Assessment Quarterly*, 12/4, 333–362.
- Holec, H. (1981): *Autonomy and foreign language learning*. Oxford: Pergamon.
- Huhta, A. – Luoma, S. – Oscarson, M. – Sajavaara, K. – Takala, S. – Teasdale, A. (2002): DI-ALANG: a diagnostic language assessment system for learners. In: Alderson, J. C., ed., *Common European Framework of Reference for Languages: learning, teaching, assessment. Case studies*. Strasbourg: Council of Europe, 130–145.
- Jaeger, R. M. (1991): Selection of judges for standard-setting. *Educational Measurement: Issues and Practice*, 10/2, 3–6.
- Kaftandjjeva, F. (2004): *Council of Europe Reference Supplement to the Preliminary Pilot Version of the Manual for Relating Language Examinations to the CEFR. Section B: Standard setting*. Language Policy Division.

- Kahneman, D. – Slovic, P. – Tversky, A. (1982): *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.
- KER = Közös Európai Referenciakeret. *Nyelvtanulás, nyelvtanítás, értékelés*. 2002. Budapest: Pedagógustovábbképzési Módszertani és Információs Központ Kht.
- KEROnline <http://www.keronline.hu/>
- Kühn, B. (2016): EPOS – the European e-portfolio of languages. *Language Learning in Higher Education*, 6/2, 335–354.
- 71/1998. (IV. 8.) Korm. rendelet az idegennyelv-tudást igazoló államilag elismert nyelvvizsgáztatás rendjéről és a nyelvvizsga-bizonyítványokról
- Lantolf, J. P. – Poehner, M. E. (2004): Dynamic assessment of L2 development: Bringing the past into the future. *Journal of Applied Linguistics*, 1/1, 49–72.
- Lengyel Zsolt (1999, szerk.): *Államilag elismert nyelvvizsgák akkreditációjának kézikönyve*. Budapest: Professzorok Háza, Nyelvvizsgáztatási Akkreditációs Központ.
- Little, D. (2009a): *The European Language Portfolio: Where pedagogy and assessment meet*. Strasbourg: Council of Europe. Elérhető: www.coe.int/portfolio
- Little, D. (2009b): Language learner autonomy and the European Language Portfolio: Two L2 English examples. *Language Teaching*, 42/2, 222–233.
- Little, D. (2012): The European language portfolio: History, key concerns, future prospects. In: Bärbel Kühn – María Luisa Pérez Cavana (eds.): *Perspectives from the European language portfolio: Learner autonomy and self-assessment*, 8–21. London: Routledge.
- Little, D. (2016): The common European framework of reference for languages, the European language portfolio, and language teaching/learning at university: An argument and some proposals. *Language Learning in Higher Education*, 6/2, 283–296.
- Linaker, M. H. (2018): A catalyst for growth? How does the European profiling grid act as an instrument for change in language teachers' development?, *Educational Action Research*, 26/2, 301–313, DOI: 10.1080/09650792.2017.1326962
- Luoma, S. (2004): Self-assessment in DIALANG: an account of test development. In: Milanovic, M. – Weir, C. (eds.): *European language testing in a global context. Studies in Language Testing 18*. Cambridge: Cambridge University Press, 143–156.
- Martyniuk, W. (2010): *Aligning tests with the CEFR: Reflections on using the Council of Europe's draft manual*. Cambridge, UK: Cambridge University Press.
- McCarthy, M. (2016): Putting the CEFR to good use: Designing grammars based on learner-corpus evidence. *Language Teaching*, 49/1, 99–115.
- NYAK (2004): Akkreditációs Kézikönyv.
- North, B. (2000): The development of a common framework scale of language proficiency. New York: Lang, Peter, Publishing Inc.
- North, B. (2009): A profiling grid for language teachers. Paper presented at the International Meeting on Training, Quality and Certification in Foreign Language Teaching, Siena.
- North, B. (2009): The Educational and Social Impact of the CEFR. In: L. Taylor – C. Weir (eds.): *Language Testing Matters: Investigating the wider social and educational impact of assessment – Proceedings of the ALTE Cambridge Conference, April 2008, Cambridge. Studies in Language Testing series 31*, Cambridge, UK: Cambridge University Press, 357–377.

- North, B. (2014): *The CEFR in practice*. Cambridge: Cambridge University Press.
- North, B. – Piccardo, E. (2016): Developing illustrative descriptors of aspects of mediation for the Common European Framework of Reference (CEFR). *Language Teaching*, 49/3, 455–459.
- Papageorgiou, S. (2010): Investigating the decision-making process of standard setting participants. *Language Testing*, 27/2, 261–282.
- Piccardo, E. – North, B. (2019): *The action-oriented approach*. Clevedon: Multilingual Matters.
- Poehner, M. E. – Lantolf, J. P. (2005): Dynamic assessment in the language classroom. *Language Teaching Research*, 9/3, 233–265.
- PTMIK (2002): *Közös európai referenciakeret. Nyelvtanulás, nyelvtanítás, értékelés*. Pedagógustovábbképzési Módszertani és Információs Központ Kht., Budapest.
- Saville, N. (2010): The English Profile Programme: Background, current issues and future prospects. *Language Teaching*, 43/2, 238–244.
<https://www.ecml.at/Portals/1/5MTP/Bleichenbacher/CEFRLT%20list%20of%20instruments.pdf>
- Saville, N. – Hawkey, R. (2010): The English Profile Programme – the first three years. *English Profile Journal*, 1.
- Rossner, R. (2017): *Language teaching competences*. Oxford: Oxford University Press.
- Szabó, G. (2011): Relating language examinations to the CEFR: ECL as a case study, In: Walde-
mar Martyniuk (szerk.) *Aligning Tests with the CEFR Reflections on Using the Council of
Europe's Draft Manual*, Cambridge: Cambridge University Press, 133–144.
- Tannenbaum, R. J. – Wylie, E. C. (2008): *Linking English language test scores onto the Com-
mon European Framework of Reference: An application of standard-setting methodology*
(TOEFL iBT Research Report RR-08-34). Princeton, NJ: Educational
Testing Service, <http://dx.doi.org/10.1002/j.2333-8504.2008.tb02120.x>
- Taylor, L. (2004): Issues of test comparability, *Research Notes*, 15, 2–5. [https://www.cam-
bridgeenglish.org/Images/23131-research-notes-15.pdf](https://www.cambridgeenglish.org/Images/23131-research-notes-15.pdf)
- Taylor, L. – Jones, N. (2006): *Research Notes*, 24, 2–6. [https://www.cambridgeenglish.org/Ima-
ges/22627-research-notes-24.pdf](https://www.cambridgeenglish.org/Images/22627-research-notes-24.pdf)
- Van Ek, J. A. (1975): *A Threshold level in a European Unit/Credit system for modern language
learning by adults*. Strasbourg: Council of Europe.
- Vígh Tibor (2005): A kommunikatív tesztelés elméleti alapjai. *Magyar Pedagógia*, 105/4, 381–407.
- Zhang, S. – Thompson, N. (2004): DIALANG: A Diagnostic Language Assessment System
(review). *The Canadian Modern Language Review / La revue canadienne des langues vi-
vantes*, 61/2, 290–293.
2005. évi CXXXIX. törvény a felsőoktatásról (XII. 14.)

LUKÁCSI ZOLTÁN

Új utak a nyelvi mérésben

New trends in language testing

The Common European framework of reference (CEFR) has seen major revisions in the recent past. The Companion Volume to the CEFR (Council of Europe 2018b) presents the altered construct of mediation along with phonology together with their scales, and it also gives a detailed description of plurilingualism and pluriculturalism. The present study is an overview of research into these new areas by European applied linguists and measurement professionals. Given the prevalence of and advances in technology, the overview discusses some of the novel aspects of IT in language testing. In the closing section, a brief comparison shows what effect the new European trends have had on the Hungarian practice of foreign language testing and assessment.

Keywords: Council of Europe, CEFR, mediation, phonology, migration

Bevezetés

Jelen tanulmány a 21. századi európai idegen nyelvi mérés és értékelés új és fokozott érdeklődéssel figyelt területeiről ad átfogó képet. A válogatás alapját az EALTA, a nyelvi teszteléssel és nyelvtudásméréssel foglalkozó legnagyobb európai szakmai szervezet tevékenységi köre, a szervezet éves konferenciájának témája, illetve kutatócsoportjainak munkája képezi. Ennek megfelelően a tanulmány különálló egységekbe rendezve tárgyalja: (a) a KER nyelvi mérésben beöltött szerepét; (b) a KER kiegészítő kötetét; (c) a korai idegennyelv-tanulást és idegen nyelvi mérést; (d) a közvetítés kiegészített konstruktumát; (e) a fonológia átdogozott skálarendszerét; (f) a plurilingvális és plurikulturális kompetenciát; (g) a nyelvtudás szerepét a társadalmi mobilitásban és integrációban; (h) a technológia nyelvi mérésben betöltött szerepét; valamint (i) a magyar gyakorlatot.

A Közös európai referenciakeret (KER, Council of Europe 2001, Európa Tanács 2002) az európai nyelvi mérés elméleti kereteit szabályozó irányelveket tartalmazza, ezért a nyelvtudásmérési kutatások és fejlesztési projektek témaválasztásában is alapvető szerepet játszik. Az elmúlt évtizedben ilyen vizsgálatok tárgyát képezte a közvetítés és a fonológia fogalma, a korai idegennyelv-tanulás és nyelvi mérés, a jelnyelvi mérés és a többnyelvűség és plurikulturalizmus szerepe a társadalmi integrációban. A tanulmány eseteken és beszélőkon keresztül ad áttekintést az érdeklődő olvasónak a kurrens trendekről.

John Trim, a KER egyik megalkotója egy interjúban (Cambridge Assessment 2011) azt jósolta, hogy a szakma figyelme a következő évtizedben a nyelv grammatikai szerveződésére fog irányulni, és az alkalmazott nyelvészek az európai nyelvek grammatikai fejlődési szintjeinek leírásával foglalkoznak majd. Áttekintésemben megvizsgálom, hogy a fenti állítás mennyiben bizonyult pontosnak.

A KER és a nyelvi mérés

A KER meghatározó jelentőségű az európai nyelvtanulás, nyelvtanítás és nyelvi mérés fő irányvonalainak a megrajzolásában (Figueras–North–Takal –Verhelst–van Avermaet 2005). Ennek megfelelően a KER a hazai nyelvpolitikában is alapvető fontossággal bír: az akkreditált nyelvvizsgáztatást szabályozó 137 / 2008 kormányrendelet 2. § (6) bekezdése a nyelvvizsgákat a KER négy szintjéhez (A2, B1, B2, és C1) társítja. Szakemberek és politikusok számos, egymástól sok esetben gyökeresen eltérő célt társítanak a KER-hez. A dokumentum saját céljaként az élő nyelvek oktatásával foglalkozók közötti kommunikációs korlátok áthidalását jelöli ki, tehát egy metanyelv, amely áttekinthetővé teszi a különböző országok nyelvtanulási céljait és eredményeit (Európa Tanács 2002: 1). A Miniszteri Bizottság kiemelte a páneurópai többnyelvűség előmozdításának a fontosságát (Európa Tanács 2002: 2), ez Figueras és társai (2005) szerint a sokszínűség leírására alkalmas eszköz megalkotását inkább jelenti, mint egy uniformizálásra törekvő rendszert. North és Panthier (2016: 18) mégis elsősorban politikai eszközként tekint a KER-re, amely a művelt többnyelvű európai népesség megvalósítását segítheti azzal, hogy oktatáspolitikai reformokat, alaptantervi változásokat mozdít elő, és hozzájárul a plurikulturális Európa értékeit közvetítő nyelvtanárok képzéséhez. Az értékalapú politikai szándék érhető tetten az Európa Tanács felnőtt bevándorlók integrációját célzó programján és a demokratikus kultúrák kompetenciáit soroló kiadványban is (Council of Europe 2016).

A KER első mondatának önmeghatározása szerint: „A Közös európai referenciakeret közös alapot teremt Európa-szerte a nyelvi tantervek, tantervkészítési irányelvek, vizsgák, tankönyvek stb. kidolgozásához” (Európa Tanács 2002: 1). Paradox módon az európai referenciakeretet számos ázsiai ország is a nyelvtanítás és nyelvtudásmérés alapdokumentumának tekinti (Alderson 2017, Pham 2017). A Japán Nyelvtanítási Egyesület 2019 májusában jelentette a *CEFR Journal: Research and Practice* online folyóiratának első számát. A belső ellentmondást a névleges ellentétén túl tovább feszíti, hogy a jelenkor idegen nyelvi méréseinek egyik irányelve a lokális meghatározottság és a helyi kontextus fontossága (O’Sullivan 2018), miközben a KER kifejezetten a specifikumoktól mentes, átfogó kép megalkotásának segédeszköze.

O’Sullivan (2018) a lokalizáció hat szintjét különbözteti meg nyelvvizsgák helyi megrendelésre történő fejlesztésében. Az egyik szélsőérték, a 0. szint az, amikor a szolgáltató a polcról levéve, változtatás nélkül adja át mérőeszközét a megrendelőnek helyi

használatra. Ennek az ellentételezése az a szélsőséges lokalizáció, amikor a vizsgaközpont teljesen új vizsgát fejleszt, ami már nem tekinthető a korábbi vizsgarendszer változatának. A közbülső szintek a lokalizáció lépcsőfokai: (1) válogatás a vizsgarendszer már kifejlesztett és működő vizsgamátrixából, például vizsgarészek vagy vizsgaformátum szerint; (2) kontextualizálás szókészleti vagy tematikus módosításokkal; (3) átszerkesztés az itemek vagy a mérhető szintek megváltoztatásával; és (4) részleges konstruktum újradefiniálás esetlegesen új feladattípus bevezetésével. O’Sullivan (2018) szerint a lokalizáció nyilvánvaló oka, hogy nem lehetséges univerzálisan elfogadott és felhasználható vizsgát fejleszteni egy kontextuálisan behatárolt környezetben. A nyelvvizsgák oldaláról vizsgálva a KER maga egy összehasonlító eszköz a nyelvvizsgáztatási gyakorlat értékelésére és kontextusba helyezésére (Európa Tanács 2002: 1).

A KER az európai országok helyi nyelvtudásmérési sajátosságaitól mentes, ebben az értelemben kontextusfüggetlen keretrendszer az Európa Tanács jellegzetesen európai nyelvpolitikai céljaival és feladataival (Európa Tanács 2002: 2–5). A KER-t mint rendszert O’Sullivan lokalizáció lépcsőfokain keresztül szemlélve felmerül a kérdés, hogy a kontextualizálás európai gyakorlatban bevett szintjeinek egyike elégséges lehet-e arra, hogy ezt az európai dokumentumot a kontinens kultúrkörain kívül is hatékonyan használni lehessen. Másként megfogalmazva lehetséges-e egy az európai tagállamok számára megfogalmazott együttműködési javaslatrendszert markánsan más kultúrkörben sikeresen alkalmazni anélkül, hogy az eredeti kulturális és nyelvpolitikai célok alapjaikban sérüljenek?

Az európai nyelvi mérésben a KER meghatározó jelenlétéről tanúskodik az EALTA KER kutatócsoportjának megalakulása 2015-ben. A kutatócsoport az éves konferenciát megelőző tanácskozáson túl évközi találkozókat is szervez. Dublinban 2019-ben a nyolcadik KER-kutatócsoport-értekezletre került sor. A sokat kutatott területek között találjuk a produktív nyelvi tevékenységeket (Folny 2018), a szintmeghatározást (Kanistra 2018) és az angol nyelvű oktatást (Moe–Wiger 2019). A KER adta a témáját az EALTA 14. konferenciájának: *A KER és a nyelvi mérés és értékelés*. Budapest ad otthont 2020-ban az EALTA 17. konferenciájának *Változatok a közös standardokban* címmel. Az egyesület korábbi elnöke, Sauli Takala életművét elismerő emlékkötet számos tanulmánya foglalkozik a KER-hez történő szintillesztéssel (Benigno–de Jong 2019, Erickson 2019, Verhelst–Figueras–Prokhorova–Takala–Timofeeva 2019).

A KER kiegészítő kötete

A KER gyakorlati jelentőségét a szemléltető skálák adják (Alderson 2007). Ugyanakkor az első kiadásban a négy nyelvi tevékenység és stratégia közül a közvetítés csak az említés szintjén van jelen, így tehát a közvetítés tanítását vagy mérését megcélzó nyelvtanítással vagy nyelvvizsgáztatással foglalkozó szolgáltatókat a dokumentum magára hagyja. Egyebek között ezt az ellentmondást hivatott feloldani a KER kiegészítő

kötete (Council of Europe 2018b). A tanulmány következő része KER kiegészítő kötetének újdonságait mutatja be. Kitér az apróbb, pontosító változtatásokra, a fiatalkori nyelvtanulásra, a hallássérültek kommunikációjának szintezésére, és a két nagyobb projektre, jelesül a közvetítés és a fonológia skálarendszerére.

A KER kiegészítő kötetének apróbb változtatásai a szövegszerkesztési hibák javításától a skálák átszövegezéséig, illetve kiegészítéséig terjednek. Külön függelék tartalmazza az első kiadás szóhasználatát érintő változtatásokat (Appendix 7, Council of Europe 2018b: 223). Itt tételes lista található arról, hogyan gyomlálták ki a szerzők az anyanyelvi beszélőre történő utalásokat a szövegből. Az „anyanyelvi beszélők közötti társalgás megértése” skála a kiegészítő kötetben „más beszélők közötti társalgás megértése” kissé tautologikus módon. A fejezetek elején megjelenő folyamatábrák is újak, hatásukra a szemléltető skálák rendszere a korábbinál lényegesen átláthatóbb. Új vonás az online interakció nyelvi tevékenységeinek kiemelése, melyek így a szóbeli és az írásbeli interakcióval azonos szintre kerültek.

Számos korábban hangoztatott észrevétel azonban válasz nélkül maradt. Így nem kap iránymutatást az olvasó egy sor pontatlan, meghatározatlan minősítő kifejezés tekintetében, amilyen például az „összetett”, „nagymértékben köznyelvi” vagy a „megjósolható” (Alderson 2007). Szintén változatlanul maradtak az átdolgozott szövegben a nyelvhasználó képességeire vonatkozó korábban feltárt következtetések, mint „*Megértem az ismerős neveket*” (A1, olvasás, Európa Tanács 2002: 34), szemben a „*felismer ismerős neveket*” leírással (A1, olvasás tájékozódás céljából, Európa Tanács 2002: 86 – *kiemelés a szerzőtől*). Végül a szövegekkel kapcsolatban felvetett aggodalmak is tovább élnek. Így a kiegészítő kötet szerint is nehezebb a hosszabb szöveg. Mindezek mellett sok esetben továbbra is rejtve marad, mit is kell a nyelvhasználónak egy adott szöveggel tennie (Alderson 2012).

A kiegészítő kötet újdonsága, hogy szerepelteti az A1 szintet megelőző nyelvi szintet (pre-A1). A meghatározás szerint ezen a szinten a nyelvhasználó félúton van az A1 szint felé, egy szűk szókészletre és nyelvi sablonokra épít, ennek megfelelően a leírások turisztikai jellegű feladatokra vonatkoznak. Szintén önálló kutatásra építő változás, hogy a skálák a ProSign logót viselik, amennyiben létezik elérhető jelnyelvi adaptációjuk. A jelnyelv a KER szerint nem csupán gesztikuláció, és nem szűkíthető a beszélt nyelv sajátos médiumára. A jelnyelv a siketek első nyelvává válhat, és mint minden emberi nyelv, rendelkezik a saját szabályrendszerével és korlátaival egyaránt. A hangzó nyelvhez hasonlóan jellemzik a nyelvészajátítás, -feldolgozás és nyelvvesztés pszicholingvisztikai folyamatai. A kiegészítő kötet tartalmaz hét olyan skálát, amelyet a svájci ProSign Projekt kifejezetten jelnyelvi kompetenciákra dolgozott ki (Council of Europe 2018b: 145–156).

A korai idegennyelv-tanulás

A korai idegennyelv-tanulás vonzerejének markáns növekedése indokolja, hogy a KER egyes eredetileg fiatal felnőttek és felnőttek nyelvhasználatát leíró deskriptorait kiskolásokra adaptálták. Mivel a szerkesztők el akarták kerülni a fiatalkori nyelvtanulás skáláinak esetlegesen zavaró párhuzamos rendszerét, a KER kiegészítő kötete nem tartalmazza ezeket a sávleírásokat. Az életkor alapján két csoportra osztott fiatal nyelvtanulók deskriptorai önálló kiadványként érhetőek el (Council of Europe 2018a). A kétkötetes gyűjtemény afféle válogatás és adaptáció eredménye, így segíti a KER-t egyébként már ismerő nyelvtanárokat reális oktatási célok kitűzésében. A 7–10 évesekkel foglalkozó 1. kötet jelentősen leszűkíti a releváns szinteket, így a C2 szint meg sem jelenik, a C1 szint nyelvi tevékenységeinek a köre pedig egy beszélt és egy írott receptív tevékenységre szorítkozik. A 11–15 évesekkel foglalkozó 2. kötetben már megjelenik a C2 szint, de a mesterfokú nyelvhasználó deskriptorai kevés kivétel mellett csak részleges relevanciával bírnak, jellemzően életkor szerint és kontextuálisan szűkítettek.

Ellentét feszül a fiatal nyelvtanulók nyelvi mérése iránti piaci igény és a kutatási eredményeket bemutató konferenciajelenlét között. Miközben a nemzetközi szolgáltatók sorra fejlesztik a gyerekek és kiskamaszok számára a nyelvvizsgákat, amilyen a Cambridge Assessment „For schools” terméksora, az ETS TOEFL „Young students” vizsgái, a Pearson „Young learners” szintjei, vagy a Goethe Intézet és az Osztrák Intézet által közösen fejlesztett Jugendliche vizsga, az EALTA konferenciákon az elvégzett kutatásokhoz képest arányaiban is kevés előadás hangzik el a témában (Hasselgreen–Torsheim 2016, Hsieh–Wang 2016, Liontou–Tsagari 2017). Az egyesület a hiátust tudatosan kezelve Bochumban szimpóziumot hívott össze a kutatási eredmények megvitatására (Barras–Brunfaut–Caudwell–Hasselgreen–Karges–Lenz–Papp 2018).

A közvetítés kiegészített konstruktuma

Évtizedekkel első megjelenését követően a KER még mindig innovatívnak tekinthető vonása, hogy szakít a Lado (1961) által felállított négy alapkészség rendszerével és helyébe a kommunikatív nyelvi tevékenységek és stratégiák osztályozását állítja (North–Panthier 2016: 18). Talán kevésbé ismert tény, hogy North (1992) eredeti javaslatában a produkció, recepció és interakció mellett a feldolgozás szerepelt negyedik nyelvi tevékenységként, amit Coste javaslatára közösen dolgoztak át közvetítéssé (North–Panthier 2016: 18). Idővel a közvetítés fogalma is változott, módosult. Az első kiadás szűk értelmezése szerint a közvetítés során egymás közvetlen megértésében akadályozott partnerek közötti kommunikáció lehetővé tétele a cél. Ez jellemzően azonos nyelvet nem beszélők között fordul elő, így a közvetítés legtipikusabb esete a fordítás, tolmácsolás vagy az átfogalmazás (Európa Tanács 2002: 103). A KER kiegészítő kötete (Council of Europe 2018b) ennél lényegesen tágabban értelmezi a közvetítést

Coste és Cavalli (2015) definíciója alapján, miszerint a közvetítés célja két távoli vagy ellentétes pólus közelítése. A polaritást a nyelvi akadályokon túl okozhatja kulturális, szemantikai, de akár technikai korlát is.

A közvetítés bővített skálarendszerének kidolgozása 2014. január és 2016. április között zajlott sok száz nemzetközi műhelyfoglalkozás keretében az Európa Tanács Nyelvpolitikai Programja által indított projektben. A deskriptorok fejlesztése három szakaszból állt. Az intuitív szakaszban a kutatók a szakirodalom áttekintésén túl használatban lévő szintleírásokat szedték csokorba, és újakat írtak, majd a deskriptorokat ismét csoportokba rendezték, szövegezéseiket szükség szerint átszerkesztették. A kvalitatív szakaszban 137 műhely során 990 nyelvtanár értékelt a sávleírásokat, illetve tematikusan csoportokba rendezték őket. A kvantitatív szakaszban a deskriptorok Rasch-módszeren szerinti kalibrálása történt, melynek eredményeként 24 kategóriában 427 szintminősített sávleírás született (North–Docherty 2016).

A projekt során a vezetőknek két problémával kellett megküzdeniük. Egyrészt nagyszámú érdeklődés mutatkozott a kutatás iránt, mint remélték. Az eredetileg 200 főre tervezett vizsgálat szerkezeti átalakítást igényelt annak érdekében, hogy a számításokat kapcsolt elrendezésben el lehessen végezni (North–Docherty 2016: 26). A gyakorlati problémánál nehezebben orvosolhatónak bizonyult a fogalmi azonosság hiánya a résztvevők között. A plurilingvális és plurikulturális kompetenciák esetében például az alacsony szintűnek tervezett sávleírásokat a tanárok jellemzően B2+ szintűnek ítélték, sok esetben pedig hosszas mellékmondat láncolatokkal javasolták kiegészíteni a 20–25 szavasra szövegezett leírásokat.

A 24 kategóriát tartalmazó komplex közvetítés skálarendszer megtartotta a fordítást és a tolmácsolást is, azonban a KER kiegészítő kötete (Council of Europe 2018b) számos olyan nyelvi tevékenységet is idesorol, amelyet hagyományosan másik osztályba tartozónak vélnénk. Régi adóssággal számol el a dokumentum, amikor a szépirodalmi művek személyes, de akár kritikai interpretációját a közvetítés szövegfeldolgozás ágára helyezi. Egyes nyelvvizsgákban választható írásfeladat előre megadott irodalmi művek elemzése. Az évente frissített irodalomjegyzékkel a vizsgázó idejekorán megismerkedhet, azonban szakmai aggály merülhet fel azzal kapcsolatban, mennyiben tekinthető tisztán írásprodukciónak egy olyan feladat, ami sok száz oldal értő olvasását és feldolgozását igényli. Ezt a feladattípust a KER kiegészítő kötete szerint (Council of Europe 2018b: 116–117) közvetítésként kell értelmezni.

Sajátos feszültség érhető tetten az új csoportos együttműködés és az átdolgozott célirányos együttműködés skálák között. Míg az előbbi a közvetítés, az utóbbi az interakció részét képezi, a fogalmi átfedés pedig olyan mértékű, hogy a két tartalom nehezen elkülöníthető. Egyes esetekben a fogalmi letisztultság hiányán felül szinteken keresztüli következetlenség is megfigyelhető (*1. táblázat*).

Meg tud kérni másokat, hogy fejtsék ki véleményüket.	B1	csoportos együttműködés	119. o.
Elő tudja segíteni a munka menetét azzal, hogy megkér másokat, hogy csatlakozzanak, kikéri véleményüket, stb.	B2	célirányos együttműködés	88. o.
Meg tud kérni másokat, hogy osszák meg vele a tennivalókkal kapcsolatos nézeteiket.	B1	célirányos együttműködés	88. o.

1. táblázat. Nehezen elkülöníthető sávleírások a KER kiegészítő kötetében

A plurikulturális tér kialakítása skála már olyan elemeket is tartalmaz, amelyek a nyelvhasználaton kívüli cselekvésekből is mélyen merítenek. Így a B2 szint egyik ismérve, hogy idegen kulturális háttérű emberekkel történő együttműködés esetén a jelölt képes legyen módosítani munkamenetét annak érdekében, hogy közös gyakorlat alakulhasson ki (Council of Europe 2018b: 123).

Az EALTA KER kutatócsoportja megalakulásától kezdve aktívan vizsgálta a közvetítés konstrukcióját. Lukácsi (2016) a közvetítés nyelvi tevékenységeinek fordításra és tolmácsolásra történő leszűkítését tárgyalja a magyar akkreditált nyelvvizsgáztatás kontextusában. Lankina (2018) a szóbeli közvetítés globális és analitikus értékelésének kihívásait vizsgálja kis csoportos kísérletében. Rodriguez (2018) a spanyol akkreditációt érintő jogszabályi változásokhoz kapcsolódó kutatásában arra a kérdésre keresi a választ, mennyire volt a szóbeli nyelvvizsga korábban beszédkézségként értelmezett része a ma közvetítésként kezelt mérési cél. A tudományos angol egyeduralma és a kulturális sokszínűség közötti feszültségre hívja fel a figyelmet Boyd (2019) kutatása, amelyben azt vizsgálja, alkalmas lehet-e a közvetítés kibővített konstrukciója a globális szakmai norma és a helyi hagyományok közötti ellentét csillapítására.

A fonológia átdolgozott skálarendszere

A közvetítés konstrukciójának átdolgozása számos szürke területre irányította a kutatók figyelmét. Természetes igény mutatkozott arra, hogy a kiegészítő kötet választ adjon a pedagógia fejlődésével és a társadalmi változásokkal együtt járó elvárásokra. Piccardo (2016: 6) szerint a fonológia az egyik ilyen, az alkalmazott nyelvészetben mostohagyerekként kezelt és következként kevésbé kutatott és feltárt terület. Az egyik cél a KER „fonológiai tudás” skálájának az átdolgozásakor az volt, hogy a nyelvtanítás és a nyelvi mérés számára realiztikus szintleírásokat adjon.

A kutatás a közvetítés skáláinál leírtakhoz hasonló terv alapján zajlott. A KER kritikái olvasata feltárta, hogy a 2001-ben kiadott skála nem tükrözi világosan az akcentust és a kiejtés fejlődését, illetve sávleírásai következtelenek egymástól jelentősen eltérő építőelemek felsorakoztatásában. Piccardo (2016: 9) hibaként emeli ki, hogy a korábbi skála „nem teljes”. Ezt érdemes összevetni azzal a megjegyzéssel, miszerint

máig „egyetlen nyelvről sem készült még teljes, kimerítő leírás” (Európa Tanács 2002: 130), és már az olvasóhoz intézett bevezető is világossá teszi, a szöveg gondolatébresztő jellegű, nem törekszik a teljességre (Európa Tanács 2002: xiii).

A fonológiai tudás átdolgozása a következő kulcsfogalmakat kívánta operacionálizálni: (a) kiejtés; (b) hanglejtés; (c) akcentus; és (d) érthetőség. A gyakorlati megvalósításban azonban a lista az általános fonológiai tudás, a hangejtés és a prozódia három komponensére szűkült. Az érthetőség a szintek elkülönítésében bizonyult meghatározó jelentőségűnek, az akcentus pedig a fonológiai tudatosság mutatójaként került a leírásokba. Az eredeti fonológiai tudás helyét tehát a KER kiegészítő kötetében (2018b: 136) három skála foglalja el A1–C2 szintig 18 deskriptorral.

A fonológia átdolgozott skáláit tanulmányozza Kallio (2017) egy a finn diákok svéd nyelvtanulását vizsgáló kismintás próbában. Kallio szerint a hangejtési sebesség, illetve a bizonyos kiejtésfolytonossági törések mutatják legbiztosabban az elért nyelvtudásszintet. Az IELTS vizsgarendszer munkahelyi interakciót vizsgáló kutatásában Galaczi és Taylor (2018) kitér a fonológia szerepére is, azonban éles vizsgák felvételeit elemezve arra a következtetésre jut, hogy a konstruktum az adott nyelvhasználati tartományban az érthetőségre szűkítendő. Ennek a megállapításnak kritikai vonatkozása van a KER-fonológia fogalomértelmezésével kapcsolatban, és nyilvánvaló implikációkat tartalmaz a nyelvvizsgát levezető vizsgáztatók képzésének tekintetében is.

A fonológia, de különösen a közvetítés skáláinak átdolgozása egyben lehetőség is volt arra, hogy a KER eredeti célja és üzenete ismételt megerősítést nyerjen. A közvetítésben a nyelvhasználó társadalmi szereplő, aki a nyelvi jelentés közösségi kialakításának folyamatában vesz részt nyelvi és egyéb, például plurilingvális és plurikulturális kompetenciákat mozgósítva. A KER közvetítéssel kiegészített szintleíró rendszere nyelvoktatási paradigmaváltást sürget, melyben a nyelv cselekvés és együttműködés, a nyelvhasználat közösségi jelentésalkotás, és az alapkészségek helyét integrált kommunikációs módok veszik át.

A plurilingvális és plurikulturális kompetencia

A multilingvalizmus, azaz soknyelvűség a KER (Európa Tanács 2002: 5) szerint „bizonyos számú nyelv ismerete, vagy több nyelv együttes jelenléte egy adott társadalomban”. Ezzel szemben a plurilingvalizmus, a többnyelvűség az egyén élményeinek és tapasztalatainak folyamatos bővülését is jelenti a nyelvek használatával és az ettől elkülöníthetetlen kulturális ismeretekkel együtt. A többnyelvűség a nyelvtanulást nem szűkíti egy jelrendszer elsajátítására, egy mentális tréningre. Az idegen nyelvek a nyelvtanulás különböző fejlettségi fokain egymás mellett párhuzamosan léteznek a nyelvtanuló fejében, aki ezeket kommunikációs céllal – akár keverten – felhasználja, és a visszajelzéseknek megfelelően finomítja őket. A nyelvhasználat során kulturális ismeretekre is szert tesz, azok tárházát bővíti, az elvárásokhoz igazodik témaválasztásában,

szóhasználatában, hanghordozásában stb. A nyelvtanulás célja ennek megfelelően nem az anyanyelvi beszélő szintjének elérése, hanem a beszélgetőpartnerrel folytatott sikeres kommunikáció.

Hasonlóképpen, míg a multikulturalitás kultúrák egymás mellett élését jelenti, a plurikulturalitás alapvetése, hogy minden kulturális csoport heterogén és eltérő értékrendű egyénekből épül fel, akik egyszerre több ilyen csoport tagjai. Így tartozhatunk egyszerre – a teljesség igénye nélkül – nemzetekhez, faji és etnikai csoportokhoz, vallási felekezetekhez, településekhez, szakszervezetekhez, szakmai egyesületekhez, bizonyos szexuális orientációjú közösségekhez, generációkhoz és családokhoz (Council of Europe 2016: 19). A korábban említett sikeres kommunikáció feltétele nem csupán a kommunikációs kompetenciák rugalmas használata, hanem a plurikulturális kompetenciáké is: a kulturális diverzitásban a kölcsönös megértés csak akkor lehetséges, ha a nyelvhasználó egyszerre fejleszti nyelvi és kulturális eszköztárát (Council of Europe 2018b: 157).

A plurilingvális és plurikulturális kompetencia skáláinak fejlesztése a közvetítés skálarendszerének átdolgozásával egyszerre történt. A plurikulturális tér kialakítása skála a közvetítésbe integrálva található, ezzel is hangsúlyozva a nyelvhasználó hathatós közreműködését kulturális mediátorként. A kommunikációs kompetenciától elkülönített három kompetencia: (a) a plurikulturális készlet használata; (b) a plurilingvális megértés; és (c) a plurilingvális készlet használata. A középszintű nyelvhasználó egyebek mellett képes a kulturális sajátosságokat megfelelően értelmezni, észreveszi a kultúrközt viselkedésformákat, és ismeri ezek jelentőségét, és képes a plurilingvális eszköztárában lévő nyelvek között váltani, ezzel is elősegítve a beszélgetőtársak megnyilvánulását.

A plurikulturalitás fogalma és új skálái esetenként szétfeszítik a nyelvtanulás, nyelvtanítás és nyelvi mérés hagyományos kereteit. A B2+ szintű nyelvhasználó az öt jellemző egyik leírás szerint képes a saját és idegen társadalmi csoportok álláspontjának és gyakorlatának értékelésére, tudatában van az ítéletek és előítéletek alapját képező kulturális értékeknek (Council of Europe 2018b: 159). Dunlea (2017) ennek kapcsán jegyzi meg, hogy bár a KER kiegészítő kötete mögött egy egészében sikeres kutatás áll, bizonyos skálák olyan mértékben tükröznek társadalmi és kulturális értékeket, amit a nyelvvizsgáztatás már nem tud befogadni. Az okleveleket kiállító intézményektől nem várható el, hogy aszerint minősítsenek jelölteket, hogy például mennyire jó demokratikus állampolgárok (Dunlea 2017). A plurikulturalitás és a másság tisztelete az európai demokráciákban elviekben mérhető lehet, azonban azt is látnunk kell, hogy ezek az értékek is kontextusba ágyazottak és számos ország és társadalom számára nem elfogadhatóak. A KER a felhasználó tudatos döntésére bízta, a diskurzus mely jellemzőit kell a nyelvhasználónak elsajátítania, mely mikro- és makrofunkciókat kell kezelnie, és ugyanígy mi képezi a nyelvi mérés részét (Európa Tanács 2002: 157).

A nyelvtudás szerepe a társadalmi mobilitásban és integrációban

A 2017-ben a franciaországi Sèvres-ben rendezett éves EALTA-konferencia témája arra a kérdésre kereste a választ, mi a nyelvi mérés szerepe a társadalmi mobilitásban és integrációban. A nyelvvizsga-bizonyítványt kiállító intézmények mindig valamilyen állami vagy szakmai akkreditáción esnek át, ami a minőségbiztosítás egyik eszköze. A szintminősítést csak nagyon ritkán célozza meg a nyelvtanuló pusztán azért, hogy nyelvi szintjével hivatalosan is tisztában legyen. A nyelvvizsgák minőségi garanciáira éppen azért van szükség, mert a bizonyítvány megszerzése felsőoktatási felvételi eljárásban, bérkiegészítés megítélésekor vagy munkaviszonyban pozíció betöltésekor elvárásként jelenhet meg. A tanulmányi és a munkahelyi előmenetel a társadalmi mobilitás alapesetei. Számos ország valamilyen nyelvvizsga megszerzését várja el az állampolgárságért folyamodóktól, így a hivatalosan igazolt nyelvtudásnak a társadalmi integrációban is határozott szerepe van. A KER (2001: 1) kifejezett célja az európai mobilitás fejlesztése, ennek azonban vannak erkölcsi, etikai, jogi, társadalmi igazságossági és számtalan egyéb vonatkozásai. Mindezekről a nyelvvizsga-szolgáltatók nem zárkozhatnak el szakmai alapon, hiszen társadalmi és politikai szerepük vitathatatlan (McNamara 2012).

A nyelvvizsgák célzottan kirekesztő aspektusa, hogy inherens különbséget tesznek egy adott nyelvet első nyelvként beszélők és ugyanezt a nyelvet második vagy idegenként beszélő nyelvhasználók között. A gyakorlatban az idegennyelv-tudást mérő nyelvvizsga akkor ad értelmezhető eredményt, ha idegennyelv-tanulók töltik ki. Hasonlóképpen, az anyanyelvi olvasási és szövegértési képesség mérésére fejlesztett feladatsor anyanyelvi populáción működik. Ez a kétosztatúság a felhasználásban is megmarad, és létrehozza a nyelvi mérés „kapuőr” szerepét (Extra–Spotti–van Avermaet 2009). A születésünkkel együtt jár egy vitán felül álló jogosultság az állampolgárságra. Ha az anyaállamtól különböző ország teljes jogú polgárává kívánunk válni, akkor a számos elvárás egyike a nyelvvizsga lehet.

A korlátozás talán legalacsonyabb foka a nemzetközi gyakorlatban a tanulmányokat idegen nyelven folytatni szándékozókkal szemben támasztott nyelvvizsga-követelmény. Számos faktor befolyásolja az idegen nyelvi környezetben történő tanulás sikerességét (Duff 1995). Ilyen a kognitív és tanulmányi érettség, a műveltség, a nyelvi szocializáltság foka, az anyagi háttér, és más nyelvi, pszichológiai, szociális tényezők. Ezek közé tartozik az elért idegen nyelvi szint is, ami természetesen fontos, de nem képes arra, hogy minden mást felülírjon. Azonban a nyelvvizsga adott szintje vagy pontértéke egyszerűen ellenőrizhető egy olyan versenyvizsga esetén, mint a felvételi eljárás. Az alkalmazott nyelvészeti kutatások egy része éppen ezért azt vizsgálja, milyen idegen nyelvi szinttel lehet sikeresen teljesíteni egy oktatási programot idegen nyelven (Dimova 2017, Lukácsi–Berkovics 2017, Moe–Wiger 2019, Newbold 2017, Tervola 2016).

A társadalmi mobilitás és integráció következő nagy érdeklődésre számot tartó területe a munka világa. Az anyanyelviek és a célország nyelvét idegen nyelvként beszélők közötti mesterséges különbségtétel a munkavállalás során talán még élesebb kontrasztot teremt, mint az oktatásban. Ennek oka szakmai természetű, hiszen a hivatását elsőrangúan végző zeneszerzőt, óvodapedagógust, erdészt vagy katonát nem elsősorban az különbözteti meg szaktársaitól, hogy melyik ország közhiteles nyilvántartásában anyakönyvezték. A gyakorlat mégis azt mutatja, hogy sok esetben a munkavállalási céllal érkező és szakmai kompetenciáit hitelesen igazoló külföldieknek egy nyelvi akadályt is le kell győzniük (Pill 2018).

A foglalkoztatásban a nyelvvizsga „kapuőr” szerepét két területen kutatták mélységében. Az egyik ezek közül a légi közlekedés, különösen a repülésirányítás (Alderson 2011, Bullock 2018). A *Versant Aviation English Test* (Downey–Suzuki–van Moere 2010) a pilóták és repülőgép-irányítók számára fejlesztett, teljesen számítógépesített beszédkésztséget és beszédértést mérő vizsga, ahol az interakció nem vizsgázói oldalának szerepkörét és az értékelést is automatizált folyamatok vezérlik. Figyelembe véve, hogy a megcélzott felhasználók a légiközlekedés biztonsága érdekében kizárólag egy előre nyomtatott forgatókönyv szerint kommunikálhatnak, a számítógép valóban képes az ilyen interakciót hűen modellálni. Ugyanakkor természetszerűen vetődik fel az a kérdés is, hogy a kreativitás ilyen mértékű elfojtása, illetve a kodifikáció ilyen foka mellett mennyire indokolható az anyanyelvi nyelvhasználók felmentése a vizsga alól (Kim–Elder 2009). Kim (2018) szerint a nyelvtudás csak akkor válik a hatékony kommunikáció eszközévé, amennyiben a nyelvhasználó rendelkezik általános szakmai kompetenciákkal, követi a repülésirányítási konvenciók szabályrendszerét, és megéri a pilóta, illetve a légiirányítás céljait és viselkedését.

Az egészségügy a munkaerőpiac másik sokat kutatott területe, ahol a szakképesítés megléte mellett az alkalmazás feltétele sok esetben valamilyen nyelvvizsga (Pill 2016, Wette 2011). Vitán felül áll az orvos vagy ápoló és a beteg, illetve a szakmai személyzet közötti interakció esetenként a szó szoros értelmében is életbevágó jelentősége. Kutatást igényel azonban annak feltárása, hogy mi teszi sikeressé az egészségügyi kommunikációt (Hegedüs 2014, O’Hagan–Manias–Elder–Pill–Woodward–Kron–McNamara–Webb–McCull 2014). A szakmabeliek ítéleteire építve Pill és McNamara (2016) megalkotta a klinikai kommunikációs kompetencia fogalmát, amely így az anyanyelvi és az idegen nyelvi nyelvhasználókra egyaránt alkalmazható.

A munkavállalás a migráció egyik fő mozgatórugója, a migráció összetett kérdésköre mégsem szűkíthető le a külföldi álláskeresésre. Az alkalmazott nyelvészeti diskurzusból is gyakran keveredik a migráció alanyainak pontos definiálása és használata. Coste és Cavalli (2015) rendre elkülönítés nélkül, felcserélhető fogalmakként használja a „gazdasági menekült”, a „bevándorló” és a „menekült” kifejezéseket. A fogalmak tisztázatlansága jellemzi az Extra, Spotti és van Avermaet (2009) kötet európai esettanulmányait

és a demokratikus kultúra kompetenciáit tárgyaló Európa Tanács-kiadványt is (Council of Europe: 2016). A fogalomazonosság és fogalomállandóság hiányára mutat rá Block (2010) írása, melyben a szerző kísérletet tesz arra, hogy a migrációs önazonosság csoportjainak leírását adja a klasszikus bevándorlóktól, a menekülteken és marginalizált gazdasági bevándorlókon át a távoli nemzettársakig (Anderson 1998, Block 2010: 489). A fogalmak világos elkülönítése azért különösen fontos, mert a menekültek nemzetközi jogi védelem alatt állnak (1951. évi Genfi Egyezmény), a bevándorlók viszont nem képeznek egyetlen jól körülírható csoportot, így egységes szabályozás sem vonatkozik rájuk a nemzetközi jogban.

Az európai nyelvi mérés és a migráció kapcsolata a nyelvvizsgák „kapuőr” szerepének extrém szélsőségéhez vezet (Shohamy 2001, 2006). Számos elsősorban skandináv állam a migrációt nem politikai eszközzel kívánja szabályozni, hanem adott nyelvből vagy nyelvekből adott szinten teljesített nyelvvizsga-követelményt támaszt a befogadásra várókkal szemben (Hamnes Carlsen–Moe 2016, Gysen–Kuijper–van Avermaet 2009, Milani 2008). A norvég álláspont sajátos őszintesége, hogy miközben 2000 kontaktórási nyelvi kurzusokat biztosít ingyenesen a menekült státusszal rendelkezőknek, ezzel nem segíteni kívánja a nyelvtanulást, hanem eltántorítani akarja a kevésbé kitartókat, így állva útját az integrációnak (Hamnes Carlsen 2017). Krumm (2012: 53) szerint egyetlen nyelvvizsga sem ad pontos iránymutatást a letelepedni kívánó egyén társadalmi integrációjának fokáról. Pochon-Berger és Lenz (2014: 20) óvatosabban fogalmaz, amikor azt állítja, kevés nyelvészeti bizonyíték támasztja alá azt az állítást, hogy a bevándorlókkal szemben támasztott magasabb célnyelvi követelmények és kulturális elvárások előnyösek lennének. Mindez nem feltétlenül jelenti a nyelv mint kulturális alapérték relativizálását (Smolicz 1980). Az idegen nyelvi mérés szakmai keretein belül értelmezve ez egy validitási, érvényességi probléma, jelesül a nyelvvizsgát nem abból a célból fejlesztették ki, amire később az eredményeket használják (Kane 2019).

A technológia szerepe a nyelvi mérésben

Hasonlóan hatalmas és teljesen más okokból megkerülhetetlenül fontos kutatási terület az elmúlt évtizedben a technológia szerepe a nyelvi mérésben. A számítógépek és a számítástechnika 1935-től van jelen a nyelvvizsgáztatásban az adatfeldolgozás, eredményszámítás, korrekciós számítások, tesztelméletek, bankolás, teszttervezés részeként (Fulcher 2000: 93). Ennek megfelelően az EALTA éves konferenciáját megelőző műhelyek között rendre megjelennek az adatfeldolgozásra és nyelvvizsga-eredményszámításra tervezett számítógépes programcsomagokat bemutató, használatuk elsajátítását megcélzó oktatások.

A technológia szerepe és jelentősége eltérő a tesztközpontú alkalmazott nyelvészeti kutatásokban. Klasszikus igény mutatkozik a papír-ceruza alapon létező nyelvvizsgák

számítógépes platformra történő adaptációjára, ami magában hordozza egy sor összehasonlító vizsgálat lehetőségét és szükségességét (Brunfaut–Harding–Batty 2018, Endres 2012, Yu–Livingston–Larkin–Bonett 2004).

Az elmúlt évtizedek automatizált értékelést vizsgáló kutatásai egy váratlan, de megnyugtató eredménnyel szolgáltak. Mára általánossá vált az a gyakorlat, hogy a jelölt íráskészségét közvetlenül, tényleges írásbeli produkcióval tesztelik: a vizsgázó levelet ír, vagy esszében elmélkedik egy problémáról. Az írásproduktumot képzett értékelők olvassák, akik valamilyen előre meghatározott rendszer szerint pontozzák a művet. Az értékelési útmutatók és skálák mellett is a produktív nyelvi tevékenységek közvetlen nyelvi mérésében az elért pontértéket nagymértékben befolyásolják olyan tényezők, mint a műfaj, az értékelési szempontok, az értékelői magatartás, vagy az értékelés fizikai feltételei (Weigle 2002). Hamp-Lyons (1990) az értékelést befolyásoló tényezők közül a vizsgáztatói szigor emeli ki. A vizsgáztatás szabályozottsága és a vizsgáztatóképzések ellenére az értékelők bizonyos fokú eltérést mutatnak. A szigor okozta pontértékelterések eredményt befolyásoló hatásának csökkentésére számos matematikai modellt dolgozott ki a szakma. McNamara és Knoch (2012) szerint a Rasch-módszer tan az ún. igazságos pontértékkel képes korrigálni az értékelői hibák jelentős részét, ezért a két szerző a szubjektív értékelés problémakörét megoldottnak tekinti.

Visszatekintve ma már látható, hogy a szubjektivitást kezelő eljárások helyett a szolgáltatók a gépi értékelés felé fordulnak, így az emberi tényező meg sem jelenik a pontértékek megállapításánál. Számos kutatás témája az automatizált értékelés és a természetes nyelvfeldolgozás olyan szoftverekkel, mint az eRater, a Coh-Metrix, az Intellimetric, a Project Essay Grade, az Intelligent Essay Assessor, a SpeechRater vagy a Versant (Attali 2007, McNamara–Graesser–McCarthy–Cai 2014, Wang–Zechner–Sun 2018). A Pearson PTE Academic vizsgája egy teljes egészében algoritmusok alapján értékelt mérőeszköz (Pearson 2009). A Cambridge Assessment az automatizált vizsgában megtartotta a beszédkészség videokonferenciás humán értékelését, jóllehet, ezt a döntést az értékelőként foglalkoztatott vizsgáztatók véleményére alapozták (Nakatsuhara–Berry–Inoue–Galaczi 2017). Furcsa lenne a gépi értékelés támogatását várni attól a vizsgáztatótól, akinek ez a tevékenység az elsődleges bevételi forrása (Khabbazzbashi–Lam–Nakatsuhara 2019).

A videokonferencia már az online technológia világa. A világháló előnyeit a nyelvoktatás és a nyelvi mérés is kreatív megoldásokat keresve vizsgálja. Dimova (2018b) az online szerepét kutatja az íráskészség fejlesztésében és mérésében, Devine, Pickles és Rogerson (2018) pedig fiatalok nyelvhasználók számára fejlesztett online játékokról számol be mint a nyelvi mérés új eszközéről. A digitális nyilvánosság persze vizsgabiztonsági kihívásokat is jelent (Dimova 2018a).

A technológiai fejlődés új eszközöket és módszereket is elérhetővé tett az alkalmazott nyelvészeti kutatások számára. Ma már a figyelem fókuszát kereső és annak változásait

feltáró szemmozgás-követéses vizsgálatok is elérhetőek a nyugat-európai kutatók számára (Unaldi 2019, Zimmermann 2019). Ilyen vizsgálattal megállapítható, hogy a vizsgázó arra figyel-e, amire a vizsga a figyelmét irányítani próbálja, milyen stratégiák mozgósításával jut el a vizsgázó a válaszig, vagy mennyi időt tölt egy-egy itemmel.

A magyar gyakorlat

Az európai nyelvi mérés 21. századi fejlődési irányvonalainak helyi adaptációja szempontjából Magyarország összetett képet mutat. Nincs olyan új és haladó téma, ami teljes egészében elkerülné a magyar kutatók figyelmét, azonban a pénzügyi források korlátozottsága és a jogszabályi keretek erősen behatárolják a szakma művelőinek mozgásterét.

A hazai akkreditált nyelvvizsgáztatás fontos elvárása a szintillesztés a KER-hez, amit a népszerű, sok vizsgázót értékelő vizsgarendszereknek két évente igazolniuk kell (Oktatási Hivatal 2019). A jogszabályi háttér számos ponton sajátos ellentmondásban áll a KER-rel. Szóhasználatát tekintve a 137/2008 Kormányrendelet 2. § (6) nem követi a KER-szintek elnevezését, mivel az A2 (alapszint) a „belépőfok”, a B1 (küszöbszint) az „alapfok”, a B2 (középszint) a „középfok”, a C1 (haladószint) pedig a „felsőfok” nevet viseli. A fogalmi következetlenségen túl ez behatárolja az akkreditálható szintek valamikori bővíthetőségét is, hiszen mi következhetne a belépőfok előtt vagy a felsőfokon túl? A terminológián túlmutat a megközelítés különbözősége. A KER központi motívuma a nyelvi profil, ami a KER kiegészítő kötetében önálló fejezetet kap (Council of Europe 2018b: 36). A referenciakeret álláspontja szerint a szintek szükségszerűen, de művileg egyszerűsítik le a nyelvhasználó összetett szükségleteit, céljait és eredményeit. A javaslat szerint a nyelvtudás pillanatnyi képét olyan sugárdiagrammal célszerű ábrázolni, amely a nyelvhasználat számos területét egyszerre láttatja, valamint jó áttekintést ad az egyén fejlődéséről. A nyelvvizsgázó tehát nem egyszerűen egy „sikeres B2” minősítést kap, hanem egy részletes képet arról, hogy melyik nyelvi tevékenység tekintetében mennyire sikeres.

Komolyabb és egyre szélesedő törésvonal húzódik a jogszabály közvetítői készsége és a KER átdolgozott közvetítés felfogása között. A 137/2008 kormányrendelet 2. § (3) szerint az egynyelvű vizsgáknak nem része a közvetítés. A KER kiegészítő kötetében (Council of Europe 2018b: 104) mégis számos olyan feladatot találunk az előadás jegyzetelésétől a grafikonelemzésig, ami magyar államilag akkreditált egynyelvű vizsgák részét képezi, illetve kétnyelvű vizsgákon nem a közvetítés vizsgarészbe számít. Ennek a kibékíthetetlen ellentétnek a háttérében több dolog áll. Egyrészt a magyar államilag akkreditált vizsgáztatás rendszerének kidolgozása megelőzte a KER első megjelenését. A szintillesztés mint követelmény évekkel később ékelődött bele a már működő hatósági eljárásba. Másrészt a jogszabály nem a KER nyelvi cselekvéseit tekinti a rendszerezés egységének, hanem a ladói (1961) alapkészségeket.

Ennek ellenére a közvetítés skálarendszerének a kidolgozásában több magyar munkacsoport is aktív részt vállalt, ahogyan a fonológia és a fiatalkorú nyelvhasználókra vonatkozó részek kidolgozásában is (Council of Europe 2018b: 15). A technológiai újítások elsősorban a számítógépes vizsgákban öltenek testet, ami minden bizonnyal magával hozza a terület részletes és áttekinthető szabályozásának az igényét.

A korai idegennyelv-tanulás és a fiatalkori idegen nyelvi mérés helyzete dinamikusan változik hazánkban. A közoktatásban tanulók idegennyelv-tudásának mérését a 35/2014. (IV. 30.), illetve két tanítási nyelvű oktatási intézmények esetén a 20/2012. (VIII. 31.) EMMI-rendelet szabályozza. A 14 év alattiak a 137/2008 kormányrendelet 3. § (1) szerint nem tehetnek államilag akkreditált nyelvvizsgát, mégis számos junior nyelvvizsga rendelkezik akkreditációval (lásd: https://nyak.oh.gov.hu/doc/akk_vizsgarendszer.asp). A feszültség a nyelvi mérés alkalmazott nyelvészeti, illetve jogi megközelítése között a nyelvvizsga-szavatosságban érhető leginkább tetten. A sikeres akkreditált nyelvvizsga egy életre szól még akkor is, ha junior nyelvvizsgát szerez a jelölt. A junior nyelvvizsgák iránti igény növekedésével bizonyára egyre nagyobb hangsúlyt kap majd a nyelvvizsga érvényességi ideje.

Összegzésként

A tanulmány az elmúlt évtizedben fokozott érdeklődésre igényt tartó európai nyelvtudásmérési irányokat mutatta be. Foglalkozott a KER megszületésével és elfoglalt helyével az európai és a kontinensen túli nyelvi mérésben, majd áttekintette a referenciakeret kibővítésére irányuló kutatásokat. Kitért a közvetítés módosított konstruktumára, a fonológia skálarendszerére, valamint a plurilingvális és plurikulturális kompetenciákra is. Áttekintette a nyelvi mérés szerepét a társadalmi mobilitásban és integrációban, végül bemutatta a technológia nyelvtudásmérésre gyakorolt hatását is.

A tanulmány bevezetőjében említett jóslatról a fentiek fényében azt mondhatjuk, John Trimnek ebben a tekintetben nem volt igaza: a 21. századi nyelvi mérés központjában az elmúlt tíz évben nem a grammatikai szerveződés állt. A kutatások mögött sokkal inkább egy interdiszciplináris hatás látszik kirajzolódni, ami egyszerre hozza a felszínre a lokalitások sokfélesége és az egységességre való törekvés közötti feszülteket és összhangot.

IRODALOM

- Alderson, J. C. (2007): The CEFR and the need for more research. *The Modern Language Journal*, 91/4, 659–663. doi: 10.1111/j.1540-4781.2007.00627_4.x
- Alderson, J. C. (2011): The politics of aviation English testing. *Language Assessment Quarterly*, 8/4, 386–403. doi: 10.1080/15434303.2011.622017
- Alderson, J. C. (2012): *Language testing and the CEFR*. English Teachers' Day conference in Luxembourg 2012. október 18.
- Alderson, J. C. (2017): Foreword to the Special Issue “The Common European Framework of Reference for Languages (CEFR) for English language assessment in China” of *Language Testing in Asia*. *Language Testing in Asia*, 7/20, 1–9. doi: 10.1186/s40468-017-0049-9
- Anderson, B. (1998): Long-distance nationalism. In B. Anderson (Ed.), *The spectre of comparisons: Nationalism, southeast Asia, and the world* (pp. 58–77). New York, NY: Verso.
- Attali, Y. (2007): *Construct validity of e-rater® in scoring TOEFL essays* (Research Report No. RR-07-21). Princeton, NJ: Educational Testing Service.
- Barras, M. – Brunfaut, T. – Caudwell, G. – Hasselgreen, A. – Karges, K. – Lenz, P. – Papp, Sz. (2018): Computer-based testing of integrated skills: the role of young learners' working memory and task motivation. In: Tsagari, D. (Moderator), *Technology-based language assessment for young learners*. Szimpózium a 15. EALTA konferencián, Bochum, Németország 2018. május 31.
- Benigno, V. – de Jong, J. (2019): Linking vocabulary to the CEFR and the Global Scale of English: A psychometric model. In: Huhta, A. – Erickson, G. – Figueras, N. (eds.), *Developments in language education: A memorial volume in honour of Sauli Takala*. Jyväskylä: University Printing House, 8–29.
- Block, D. (2010): Problems portraying migrants in Applied Linguistics research. *Language Teaching*, 43/4, 480–493, doi:10.1017/S0261444809990280
- Boyd, E. (2019): *Are the new mediation descriptors likely to alleviate concerns about the dominance of academic English in research papers?* Előadás a 8. EALTA KER kutatócsoport találkozón, Dublin, Írország.
- Brunfaut, T. – Harding, L. – Batty, A. O. (2018): Going online: The effect of mode of delivery on performances and perceptions on an English L2 writing test suite. *Assessing Writing*, 36, 3–18. doi: 10.1016/j.asw.2018.02.003
- Bullock, N. (2018): *Language proficiency testing in aviation: 10 years of challenges and solutions*. Előadás a Language Testing Forum konferencián, Luton, Egyesült Királyság.
- Cambridge Assessment (Producer) (2011): *Critique of the CEFR* (Video 2). <https://www.youtube.com/watch?v=0j97x5PoCJE>
- Coste, D. – Cavalli, M. (2015): *Education, mobility, otherness: The mediation functions of schools*. Strasbourg: Council of Europe.
- Council of Europe (2001): *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

- Council of Europe (2016): *Competences for democratic culture: Living together as equals in culturally diverse democratic societies*. Strasbourg: Cedex.
- Council of Europe (2018a): *Collated representative samples of descriptors of language competences developed for young learners: Resource for educators* (Vols. 1–2). Strasbourg: Language Policy Division.
- Council of Europe (2018b): *Common European framework of reference for languages: Learning, teaching, assessment. Companion volume with new descriptors*. Strasbourg: Language Policy Division.
- Devine, A. – Pickles, M. – Rogerson, S. (2018): *Game-based assessment for the 21st century language learner*. Előadás a 15. EALTA konferencián, Bochum, Németország.
- Dimova, S. (2017): *Lecturer certification in European higher education: What English to assess?* Előadás a 14. EALTA konferencián, Sèvres, Franciaország.
- Dimova, S. (2018a.): *EMI lecturer assessment: Developing secure test management system*. Poszter a 15. EALTA konferencián, Bochum, Németország.
- Dimova, S. (2018b): *L2 writing with internet access: Implications for teaching and assessment*. Előadás a 15. EALTA konferencián, Bochum, Németország.
- Downey, R. – Suzuki, M. – van Moere, A. (2010): High-stakes English language assessments for aviation professionals: Supporting the use of a fully automated test of spoken-language proficiency. *IEEE Transactions on Professional Communication*, 53, 18–32.
- Duff, P. (1995): An ethnography of communication in immersion classrooms in Hungary. *TESOL Quarterly*, 29/3, 505–537.
- Dunlea, J. (2017): Filling the gaps in the CEFR: strengthening the description of the construct. Előadás az 5. EALTA KER-kutatócsoport találkozón, Sèvres, Franciaország.
- Endres, H. (2012): A comparability study of computer-based and paper-based writing tests. *Research Notes*, 49, 26–33.
- Erickson, G. (2019): Holistic peer analyses of National Tests in relation to the CEFR. In: Huhta, A. – Erickson, G. – Figueras, N. (eds.), *Developments in language education: A memorial volume in honour of Sauli Takala*. Jyväskylä: University Printing House, 49–66.
- Európa Tanács (2002): *Közös Európai Referenciakeret: Nyelvtanulás, nyelvtanítás, értékelés*. Pedagógus-továbbképzési Módszertani és Információs Központ.
- Extra, G. – Spotti, M. – van Avermaet, P. (2009): Testing regimes for newcomers. In: Extra, G. – Spotti, M. – van Avermaet, P. (eds.), *Language testing, migration and citizenship*. New York, NY: Continuum, 3–33.
- Figueras, N. – North, B. – Takala, S. – Verhelst, N. – van Avermaet, P. (2005): Relating examinations to the Common European Framework: A manual. *Language Testing*, 22/3, 262–279. doi: 10.1191/0265532205lt308oa
- Folny, V. (2018): *Concurrent validation CEFR scale (for productive skills)*. Előadás a 7. EALTA KER-kutatócsoport találkozón, Bochum, Németország.
- Fulcher, G. (2000): Computers in language testing. In: Brett, P. – Motteram, G. (eds.), *A special interest in computers: Learning and teaching with information and communications technologies*. Manchester: IATEFL Publications, 93–107.
- Galaczi, E. – Taylor, L. (2018): *Interactional competence in the workplace: Challenges and opportunities*. Előadás a Language Testing Forum konferencián, Luton, Egyesült Királyság.

- Gysen, S. – Kuijper, H. – van Avermaet, P. (2009): Language testing in the context of immigration and citizenship: The case of the Netherlands and Flanders (Belgium). *Language Assessment Quarterly*, 6/1, 98–105, doi: 10.1080/15434300802606655
- Hamp-Lyons, L. 1990. Second language writing: Assessment issues. In B. Kroll (Ed.), *Second language writing: Research insights from the classroom*. New York, NY: Cambridge University Press, 69–87.
- Hammes Carlsen, C. (2017): *Language testing with a bitter aftertaste*. Előadás a 14. EALTA konferencián, Sèvres, Franciaország.
- Hammes Carlsen, C. – Moe, E. (2016): Language testing as part of integration policy in Norway. *Kieli, koulutus ja yhteiskunta*, 7/6.
- Harangus, K. (2014): Anyanyelvi kompetenciák: jelen és jövő a mérések tükrében. In: Pletl, R. (szerk.) *Anyanyelvoktatás: Irányok és törésvonalak*. Kolozsvár: Ábel Kiadó, 63–84.
- Hasselgreen, A. – Torsheim, T. (2016): *Construct validation in a test of reading for young learners*. Előadás a 13. EALTA konferencián, Valencia, Spanyolország.
- Hegedűs, A. (2014): *The construct of speaking ability in a bilingual English for medical purposes exam*. Előadás a 3. EALTA Beszédkészség kutatócsoport találkozón, Valencia, Spanyolország.
- Hsieh, C.-N. – Wang, Y. (2016): *Investigating the construct of speaking proficiency for young language learners: A discourse analytic study*. Előadás a 13. EALTA konferencián, Valencia, Spanyolország.
- Kallio, H. (2017): *Revised CEFR descriptor scale for phonological control: A small-scale pilot*. Előadás az 5. EALTA KER-kutatócsoport találkozón, Sèvres, Franciaország.
- Kane, M. (2019): *Flexibility and utility in assessment and validation*. Előadás a 16. EALTA konferencián, Dublin, Írország.
- Kanistra, V. (2018): *Exploring the feasibility of using the Item Descriptor Matching (ID) method to set cut scores on a B2 CEFR level examination*. Előadás a 7. EALTA KER-kutatócsoport találkozón, Bochum, Németország.
- Khabbazzbashi, N. – Lam, D. – Nakatsuhara, F. (2019): *Technology and authenticity in the IELTS Speaking Test: What do the examiners say?* Előadás a 16. EALTA konferencián, Dublin, Írország.
- Kim, H. (2018): What constitutes professional communication in aviation: Is language proficiency enough for testing purposes? *Language Testing*, 35/3, 403–426, doi: 10.1177/0265532218758127
- Kim, H. – Elder, C. (2009): Understanding aviation English as a lingua franca: Perceptions of Korean aviation personnel. *Australian Review of Applied Linguistics*, 32/3, 23.01–23.17.
- Krumm, H.-J. (201): Multilingualism, heterogeneity and the monolingual policies of the linguistic integration of migrants. In: Messer, M. – Schroeder, R. – Wodak, R. (eds.) *Migrations: Interdisciplinary perspectives*. New York, NY: Springer, 43–54.
- Lado, R. (1961): *Language testing*. London: Longman.
- Lankina, O. (2018): *Assessment of oral mediation: Global and analytical marks*. Előadás a 7. EALTA KER-kutatócsoport találkozón, Bochum, Németország.
- Liontou, J. – Tzagari, D. (2017): *A computational linguistic analysis of the development of L2 writing proficiency: Evidence from EFL young learners' written essays*. Előadás a 14. EALTA-konferencián, Sèvres, Franciaország.

- Lukácsi, Z. (2016): Lost in translation: A local interpretation of mediation. Előadás a 2. EALTA KER-kutatócsoport találkozón, Valencia, Spanyolország.
- Lukácsi-Berkovics, I. (2017): *Chinese immigrants in the Hungarian system of public education*. Előadás a 14. EALTA-konferencián, Sèvres, Franciaország.
- McNamara, D. S. – Graesser, A. C. – McCarthy, P. M. – Cai, Z. (2014): *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge: Cambridge University Press.
- McNamara, T. (2012): Language assessments as Shibboleths: A poststructuralist perspective. *Applied Linguistics*, 33/5, 564–581, doi: 10.1093/applin/ams052
- McNamara, T. – Knoch, U. (2012): The Rasch wars: The emergence of Rasch measurement in language testing. *Language Testing*, 29/4, 555–576, doi: 10.1177/0265532211430367
- Milani, T. (2008): Language testing and citizenship: A language ideological debate in Sweden. *Language in Society*, 37/1, 27–59, doi: 10.1017/S0047404508080020
- Moe, E. – Wiger, S. (2019): *Using the CEFR to determine language levels necessary for subject learning*. Előadás a 8. EALTA KER-kutatócsoport találkozón, Dublin, Írország.
- Nakatsuhara, F. – Berry, V. – Inoue, C. – Galaczi, E. (2017): Exploring the use of video-conferencing technology in the assessment of spoken language: A mixed-methods study, *Language Assessment Quarterly*, 14/1, 1–18. doi: 10.1080/15434303.2016.1263637
- Newbold, D. (2017): *Student mobility, ELF, and the role of language certification*. Előadás a 14. EALTA-konferencián, Sèvres, Franciaország.
- North, B. (1992): European Language Portfolio: Some options for a working approach to design scales for proficiency. In: Schärer, R. – North, B. (eds.), *Towards a common European framework for reporting language competency*. Washington, DC: NFLC Occasional Paper, National Foreign Language Center, 158–172.
- North, B. – Docherty, C. (2016): Validating a set of CEFR illustrative descriptors for mediation. *Cambridge Research Notes*, 63, 24–30.
- North, B. – Panthier, J. (2016): Updating the CEFR descriptors: The context. *Cambridge Research Notes*, 63, 16–23.
- O’Hagan, S. – Manias, E. – Elder, C. – Pill, J. – Woodward-Kron, R. – McNamara, T. – Webb, G. – McColl, G. (2014): What counts as effective communication in nursing? Evidence from nurse educators’ and clinicians’ feedback on nurse interactions with simulated patients. *Journal of Advanced Nursing*, 70/6, 1344–1354. doi: 10.1111/jan.12296.
- O’Sullivan, B. (2018): *Localisation: The ultimate goal in language testing?* Előadás a 15. EALTA-konferencián, Bochum, Németország.
- Oktatási Hivatal. (2019): Akkreditációs kézikönyv. <https://nyak.oh.gov.hu/nyat/doc/ak2019/ak2019.htm>
- Pearson (2009): *Official guide to Pearson Test of English Academic*. London: Longman.
- Pham, T. H. N. (2017): Applying the CEFR to renew a general English curriculum: Successes, remaining issues and lessons from Vietnam. In: O’Dwyer, F. et al. (eds.) *Critical, constructive assessment of CEFR-informed language teaching in Japan and beyond*. Cambridge: Cambridge University Press, pp. 97–117.
- Piccardo, E. (2016): *Common European framework of reference for languages: Learning, teaching, assessment. Phonological scale revision process report*. Strasbourg: Cedex.

- Pill, J. (2016): Drawing on indigenous criteria for more authentic assessment in a specific-purpose language test: Health professionals interacting with patients. *Language Testing*, 33/2, 175–193, doi:10.1177/0265532215607400
- Pill, J. (2018): *How policy makers view language tests for professional registration*. Előadás a Language Testing Forum konferencián, Luton, Egyesült Királyság.
- Pill, J. – McNamara, T. (2016): How much is enough? Involving occupational experts in setting standards on a specific-purpose language test for health professionals. *Language Testing*, 33/2, 217–234, doi: 10.1177/0265532215607402
- Pochon-Berger, E. – Lenz, P. (2014): *Language requirements and language testing for immigration and integration purposes* (Report of the Research Centre on Multilingualism). Fribourg: Institute of Multilingualism.
- Rodriguez, C. (2018): *Incorporating mediation descriptors in a C1 speaking test*. Előadás a 7. EALTA KER-kutatócsoport találkozáson, Bochum, Németország.
- Shohamy, E. (2001): *The power of tests: A critical perspective on the uses of language tests*. New York: Longman.
- Shohamy, E. (2006): *Language policy: Hidden agendas and new approaches*. London: Routledge.
- Smolicz, J. J. (1980): Language as a core value of culture. *RELC Journal*, 11(1), 1–13, doi: 10.1177/003368828001100101
- Trevola, M. (2016): *Finnish language proficiency of immigrant physicians in medical licensure examinations*. Előadás a 12. EALTA-konferencián, Kopenhága, Dánia.
- Unaldi, A. (2019): *The use of eye-tracking and verbal protocols in construct validation: Multiple-text reading tasks in EAP tests*. Előadás a 16. EALTA-konferencián, Dublin, Írország.
- Verhelst, N. – Figueras, N. – Prokhorova, E. – Takala, S. – Timofeeva, T. (2019): Standard setting for writing and speaking: The Saint Petersburg experience. In: Huhta, A. – Erickson, G. – Figueras, N. (eds.) *Developments in language education: A memorial volume in honour of Sauli Takala*. Jyväskylä: University Printing House, 278–301.
- Wang, Z. – Zechner, K. – Sun, Y. (2018): Monitoring the performance of human and automated scores for spoken responses. *Language Testing*, 35/1, 101–120, doi: 10.1177/0265532216679451
- Weigle, S. C. (2002): *Assessing writing*. Cambridge: Cambridge University Press.
- Wette, R. (2011): English proficiency tests and communication skills training for overseas-qualified health professionals in Australia and New Zealand. *Language Assessment Quarterly*, 8/2, 200–210, doi: 10.1080/15434303.2011.565439
- Wilkins, D. A. (1978): Proposal for levels definition. In: Trim, (ed.) *Some possible lines of development of an overall structure for a European unit / credit scheme for foreign language learning by adults*. Strasbourg: Council of Europe, (71–78).
- Yu, L. – Livingston, S. A. – Larkin, K. C. – Bonett, J. (2004): *Investigating differences in examinee performance between computer-based and handwritten essays* (ETS Research Report RR-04-18). Princeton, NJ: Educational Testing Service.
- Zimmermann, S. (2019): *Transitions in L2 writing assessment: Insights from eye tracking and stimulated recall*. Előadás a 16. EALTA-konferencián, Dublin, Írország.

TANKÓ GYULA

A validálási folyamat érvelésalapú megközelítéseinek áttekintése

An overview of argument-based assessment validation

The aim of the article is to present an overview of the evolution of validity and of the argument-based approaches to language assessment validation. It is argued that there have been three seminal paradigm shifts in language assessment (the validity types taxonomy elaborated by Cronbach – Meehl 1955; the unitary view of validity proposed by Messick, 1989; and the assessment use argument developed by Bachman – Palmer 2010). The article briefly outlines and contrasts the argument-based validation procedures that make use of Toulmin's (1958) argument model (Mislevy's evidence-centered assessment design, Mislevy – Steinberg – Almond 2003; Kane's interpretation/use argument, Kane 2013; and Bachman & Palmer's assessment use argument published in 2010). It is contended that the last one proposes the most user friendly and straightforward justification procedure, which has real potential to empower not only assessment specialists but also further stakeholders.

Keywords: evolution of validity, informal reasoning, argument-based validation procedures, contrastive analysis of justification procedures, empowering assessment specialists

1. Bevezetés

Minden tesztfejlesztőnek vagy egy tesztet valamilyen célra felhasználó tesztalkalmazónak, aki lehet vizsgáztató, nyelvtanár, az adatgyűjtéshez egy mérőeszközt használó kutató vagy akár egy nyelvvizsgaközpont, arra kell törekednie, hogy a nyelvtudást nemcsak a lehető leghatékonyabb, hanem egyben a legmegbízhatóbb módon mérje. A tesztfejlesztő vagy tesztalkalmazó sokszor ilyen kérdéseket fogalmaz meg:

- Mit jelent a teszten elért pontszám?
- Miért értelmezhetem úgy a pontszámot, hogy a nyelvtanuló egy a Közös európai referenciakeret által meghatározott szintnek megfelelő nyelvtudással rendelkezik; képes egy képzést idegen nyelven sikeresen elvégezni; el tudja látni egy nyelvígé-nyes munkakör feladatait; egy bizonyos mértékig tanulta meg a nyelvkönyv egyik fejezetének tananyagát; egyes nyelvtani szerkezeteket és szókincset megbízhatóan, míg másokat bizonytalanul vagy egyáltalán nem képes használni?

- Miért dönthetek úgy a tesztpontszám alapján, hogy a nyelvtanuló egy bizonyos nyelvvizsga-bizonyítványt megkaphat; felvételt nyer egy képzésre vagy egy munkakörbe; készen áll, hogy elkezdje a következő fejezet tananyagát, vagy újra át kell néznie még valamit a nyelvkönyv éppen befejezett fejezetében; esetleg tudását egy középfokú nyelvtanfolyamon tudja a leghatékonyabban továbbfejleszteni?

Ezekkel a kérdésekkel a szakember tudatosan vagy sem, az általa végzett mérés vagy használt teszt validitását boncolgatja. A validitás márpedig vitathatatlanul minden mérési folyamat sarokköve.

Felmerül tehát egy sor kérdés: Mi a validitás? Mi a validálás? Hogyan validálható egy teszt a gyakorlatban? Hogyan ellenőrizhető egy teszt validitása? Ragaszkodnunk kell a validitáshoz mint nehezen átlátható elvont fogalomhoz, vagy észszerűbb az egyes jellemzői mentén gondolkodni a tesztfejlesztés és -felhasználás során? Validálni kell egy tesztet, vagy elég igazolni egy meghatározott mérési célra való alkalmazhatóságát? Mivel a validitás nehezen meghatározható, illetve a mérés és értékelés gyakorlatára nehezen lefordítható fogalom, ezért a validálás mint folyamat csak kevesek által érthető eljárásnak, a validitás meglétének vizsgálata pedig megvalósíthatatlan feladatnak tűnhet. Viszont a nyelvi mérési és értékelési elmélet legfrissebb fejleményei egyértelműen afelé mutatnak, hogy ennek nem kell feltétlenül így lennie. Ha a validitás fogalmát nem a hagyományos módon közelítjük meg, akkor az általa leírt tartalom közérthetővé válik, és a validálási folyamat érvelés alapú megközelítésével a validálás menete világos eljárássá szelídül, melyet a vizsgálfejlesztő és felhasználó is meg tud érteni, illetve hatékonyan tud kivitelezni.

A validitás jelentésével és validálási folyamattal kapcsolatos fenti kérdésekre keresve a választ, ennek a tanulmánynak a céljai a következők: (i) a validitás fogalmi evolúciójának és az érvelés alapú validálás eredetének, illetve strukturális felépítésének rövid áttekintése, hogy megértsük, miből fakad a validitás fogalmának komplexitása, illetve hogy honnan származik és mi képezi az érvelésen alapuló validálási eljárást; (ii) a Toulmin-féle érvelési modellre alapozott validálási eljárásmodok bemutatása és rövid kontrasztív elemzése annak érdekében, hogy az olvasó nagy vonalakban megértse azokat, és így össze tudja tudja őket hasonlítani.

A tanulmány e célok mentén először a validitás és validálás fogalmakat tárgyalja, azután leírja az érvelés alapú validálási eljárások kiindulópontját képező érvelési modellt, majd röviden bemutatja és összehasonlítja az érvelési modellen alapuló validálási eljárásokat.

2. A validitás fogalmának fejlődése

A validitás fogalmának fejlődése önmagában egy tanulmány lehetne, de most csupán a tanulmány szempontjából fontos mozzanatokat emelem ki. A szakemberek már az ötvenes éveket megelőzően a validitás több típusát különböztették meg (a kritérium-alapú validitás prediktív és megegyező altípusait, valamint a tartalmi validitást), majd Cronbach és Meehl (1955) egy új validitástípust, a szerkezeti validitás fogalmát vezették be, mely a teszteredmény és a teszt által mért mentális konstruktum(ok) közötti kapcsolatot írta le. Ezzel elindítottak egy paradigmaváltást a mérés és értékelés tudományágban.

De hamarosan megkérdőjeleződött a validitás három egyaránt fontos és jól elkülönülő típusa alapján történő meghatározása (Cronbach 1980, 1988, Ebel 1961). A kutatók amellet kezdtek állást foglalni, hogy a validitástípusok egy összetett méréselméleti fogalom szorosan összefüggő attribútumai. Ezt a gondolatmenetet formalizálta Messick (1989), aki az egységes konstruktumvaliditás-fogalom megalkotásával a validitás fogalmának új jelleget adott úgy, hogy a kritériumalapú és a tartalmi validitást a szerkezeti validitás komponenseinek nyilvánította, mivel mindkettő a megszerzett pontszám jelentéstartalmához járul hozzá. Így egy fő validitástípus keletkezett: a szerkezeti validitás, melynek fontos attribútumává vált a teszt megbízhatósága (Chapelle 1999).

Messick (1989) előírta annak a vizsgálatát is, hogy az eredmény értelmezése döntéshozatali szempontból nézve mennyire releváns, milyen következményei vannak, illetve azok mennyire hasznosak. Így az egységes konstruktumvaliditás-fogalom szerkezetéből adódóan a pontszámértelmezés mellett a pontszámalkalmazás és annak következményeinek vizsgálata is a modern validálási folyamat kötelező elemévé vált. A validitás fogalmának Messick-féle újragondolása egy újabb paradigmaváltást eredményezett a mérés és értékelés tudományágban.

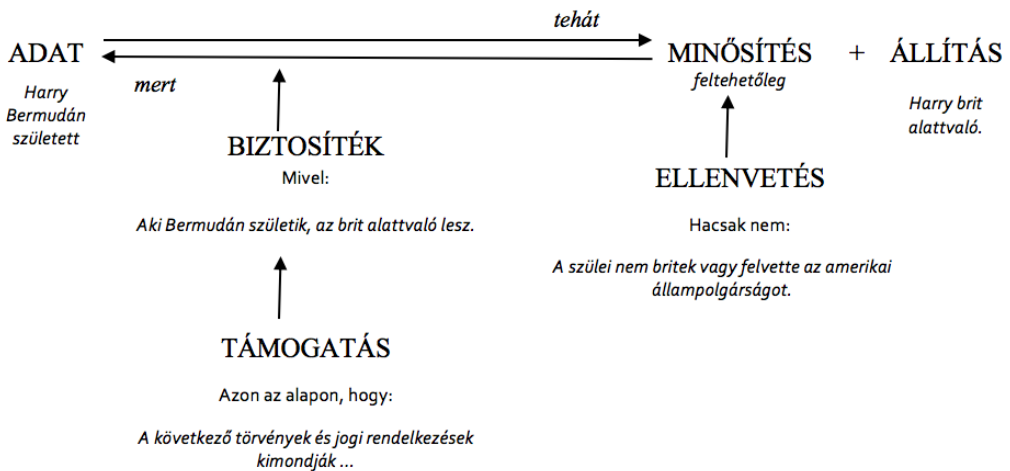
3. Az érvelésalapú validálás eredete és szerkezeti alapjai

A felismerés, miszerint a validitást a validitás igazolására felépített érvrendszer határozza meg (Cronbach–Meehl 1955, House 1980, 2014) a validáláselmélet további fontos fejleménye volt. Így Cronbach (1988) a validálás gyakorlatát egy érvelésen alapuló, és a validálási körülmények változása esetében ismételten elvégezendő formatív értékelési folyamatként határozta meg, melynek célja a teszteredmény magyarázattal történő *értelmezése*. Az értelmezés lehet leíró értékelés, előrejelzés vagy akár egy döntési javaslat (Cronbach 1971), és az értelmezés validitási fokának különböző szintjei lehetnek (Cronbach–Meehl 1955), mivel egy érvrendszer meggyőző ereje a begyűjtött bizonyítékok függvényében változik. Ugyan néhányan még sokáig ragaszkodtak hozzá (Alderson–Clapham – Wall 1995, Henning 1987, Hughes 1989, Kontra–Kormos 2007), illetve teszi azt a OH-NyAK (2019) a mai napig (lásd Fogalomtár), a validitás fogalmának újraértelmezése végérvényesen megcáfolta azt a korábbi meghatározást, miszerint

az a valid vizsga, ami azt méri, amit mérni szándékozik, és így a *valid teszt* fogalom elveszítette időszerűségét.

Cronbach (1988) megközelítésével megegyező módon Messick (1995) is az érvelést tartja a validálás alapfolyamatának. Szerinte a validálás egy érvelésen alapuló tudományos vizsgálódás, melynek során különböző jellegű bizonyítékokat sorakoztatunk fel annak megvizsgálása érdekében, hogy a begyűjtött empirikus bizonyítékok és elméleti okfejtések milyen mértékben támasztják alá a tesztpontszám alapján megfogalmazott értelmezéseket és azok felhasználását, illetve az értelmezések és azok alkalmazásának potenciális következményeit.

Az érvelési folyamat részeként Messick (1995) a pontszám alapján megfogalmazott értelmezésekkel összefüggő következtetéseket és a pontszám alkalmazásával, más szóval a pontszám alapján meghozott döntésekkel, illetve az azokból származó cselekvésekkel összefüggő következtetéseket különböztet meg. Szerinte a validálási folyamat során alá kell támasztani minden egyes következtetést, illetve bizonyítani kell, hogy a rendelkezésre álló bizonyítékok a lehetséges alternatív következtetéseket kevésbé támasztják alá. Ezért az érvelési folyamat részének tekinti az igazolást és a cáfolatot is.



1. ábra. Toulmin (1958) argumentációs modellje kiegészítve az okozat → ok (mert) relációval

A validálás érvelésen alapuló folyamatként való meghatározása szükségessé tette az érvelés menetének pontos és gyakorlatban alkalmazható leírását. Erre a célra a validálással foglalkozó szakemberek a legalkalmasabbnak Toulmin (1958, 2003) argumentációs modelljét találták. Ez az argumentációs modell az informális logikából ered, és a valós életben előforduló érvszerkezeteket mint döntésekhez vezető racionális folyamatokat írja le, lehetővé téve azok elemzését vagy akár létrehozását. A modell (1. ábra) egy következtetést

eredményező érv séma (mindegyik nyíl egy következtetési műveletet jelöl), melynek hat komponense van. A sémát balról jobbra haladva szokás olvasni (ok \rightarrow okozat reláció), de ellenkező irányba is működik, például egy állítás bizonyítási hátterének létrehozása vagy feltárása során.

Az alapértelmezett kiindulási pontot a tények vagy *adatok* képezik. Ez a tényállás, amelyet nem vonunk kétségbe, vagy amennyiben igen, akkor az könnyen verifikálható. Egy teszt esetében a nyelvvizsgáló megfigyelt viselkedése minősül adatnak (például a feladatlapon megjelölt válaszok, vagy egy vizsgán megírt panaszlevél). Az adat alapján egy *állítást* fogalmazunk meg, ami nem más, mint egy következtetés. A logikai lépést az adattól az állításig egy *biztosíték* garantálja, ami nem más, mint egy, a mindennapi életben vagy diszciplínaszpecifikusan elfogadott igazoló állítás vagy általánosítás, ami lehet egy törvény, elv vagy szabály (ezt bővebben Toulmin–Rieke–Janik 1984 tárgyalják). Ez a három komponens megegyezik a szillogizmus alkotóelemeivel: az adat a kis premissza, a biztosíték a nagy premissza, és az állítás a következtetés.

A Toulmin-modell további három komponenst határoz meg. Amennyiben szükséges, a biztosítékot alá lehet támasztani egy *támogatással*. Az érvrendszer fontos része továbbá az *ellenvetés*, mely kivételeket, kizáró okokat vagy az állítás elvetésének feltételeit, vagyis azon körülményeket írja le, melyek az állítás erejét korlátozzák vagy az állítást megcáfolják, ezáltal egyben az érvrendszer egészét gyengítve vagy cáfolva. Közvetlenül ehhez kapcsolódó komponens a *minősítés*, mely általában egy megszorító vagy nyomatékosító határozószó vagy határozói szerkezet, mely vagy az adat, biztosíték és támogatás összeadott hatása alapján, vagy az ellenvetés függvényében jelöli az állítás erősségi fokát (Toulmin–Rieke–Janik 1984).

Nem véletlenül esett a validálási folyamatot vizsgáló kutatók választása Toulmin argumentációs modelljére. A modell egy logikusan felépített sémát kínál, ami az érvelési folyamat minden komponensét tartalmazza, és az igazolást a cáfolattal ötvözi. Egyaránt alkalmas a modell a mindennapi életben, illetve a különböző tudományágakban előforduló érvelési folyamatok leírására, elemzésére vagy létrehozására. Teret enged az állításokkal kapcsolatos bizonytalanságnak, ami a validálási folyamat szempontjából nagyon fontos, hiszen bármely teszt a felsorakoztatott validitást alátámasztó bizonyítékok függvényében egy bizonyos fokú validitással rendelkezik (Cronbach–Meehl 1955, House 2014, Messick 1989).

A validitásfogalom fejlődése, és az érvelésalapú validálás eredetének, illetve érvszerkezetének áttekintése azt mutatja, modern megközelítésben a validitás egy összetett és egységes fogalom, melynek a mindent magában foglaló központi komponense a szerkezeti validitás. A validálás pedig a teszteredmények értelmezésének, alkalmazásának és következményeinek indokolására szolgáló logikán és empirikus bizonyítékokon alapuló koherens érvelési folyamat, melynek operacionalizált leírása Toulmin (2003) argumentációs modellje segítségével valósítható meg. A validálási folyamatban

nemcsak a tesztfeljesztő, hanem a tesztet használó minden egyes félnek részt kell vennie (Chapelle 1999). Végezetül, a validálás potenciálisan egy lezáratlan folyamat, ugyanis amennyiben a validálási körülmények változnak, a megváltozott körülményeket figyelembe véve a validálást újra el kell végezni (Cronbach 1988).

Következésképpen, a validitás egy többszörösen összetett fogalom, és a validálás egy bonyolult folyamat, melynek gyakorlati megvalósítása sok kérdést vet fel, és sokat vitatott. A validálási folyamat operacionalizálására több kutató dolgozott ki validálási eljárás módokat. Ezek közül a legismertebbek és legbefolyásosabbak a Mislevy (2003), Kane (2013), illetve Bachman és Palmer (2010) által javasolt eljárás módok¹.

Ezekben az eljárás módokban több közös vonás felismerhető. Mindegyik célja a pontszámok értelmezésének és alkalmazásának validálása. A validálás érdekében mindegyik egy érvrendszer felépítését írja elő, mely egy érvelési láncból áll, és amelyben minden érv a Toulmin érvelési modell komponenseivel írható le. A retorikából ismert struktúra szerint (van Eemeren–Grootendorst–Snoeck 2002) az érvelési láncban egy érv kimeneti pontja egy másik érv bemeneti pontját képezi. Tehát az érvelési lánc szerkezete és az érvek struktúrája biztosítja a validáláshoz szükséges érvelési keretrendszert, de azt az egyes eljárás módok máshogy és eltérő sikerességgel töltik fel tartalommal a validálási folyamat kivitelezésére javasolt eljárás módok között. A validálási eljárás módok áttekintése azt mutatja, hogy azok egy evolúciós láncot képeznek, ugyanis egymásra épülnek, és egymáshoz képest kifinomultabbá válnak. Annak megítélése érdekében, hogy a validálási eljárás módok közül melyik ültethető át a gyakorlatba a legegyszerűbben, a következő részben bemutatásra kerül a három legbefolyásosabb modell legújabb változata és azok rövid, a gyakorlati alkalmazhatóságuk szempontjából elvégzett elemzése.

4. Érvelésalapú validálási eljárás módok

4.1. A Mislevy-féle bizonyítékalapú tesztfeljesztési keretrendszer

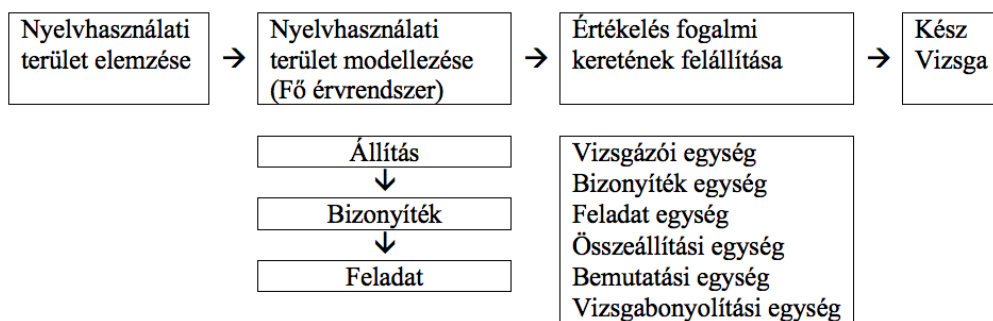
A bizonyítékalapú tesztfeljesztési keretrendszert a kidolgozói (Mislevy 2003, Mislevy–Steinberg–Almond 2002, 2003) tesztek létrehozására és mérőeszközként használt tesztek elemzésére alkalmas módszertani útmutatónak szánták a pedagógiai értékelésben érdekelttek számára². Míg a keretrendszer elsősorban a tesztfeljesztési folyamatot mutatja be, a mérőeszközök vizsgálatának kulcsfontosságú részét képező érvelésen alapuló validálási folyamatokat is körvonalazza. Ugyanis Mislevy (2003) szerint a

¹ Más érvelésalapú validálási eljárás módok is léteznek, de azok nem a Toulmin-féle argumentációs modellel épülnek (lásd pl. Weir 2005, melynek összefoglalását Tankó 2015, validálási tanulmánya tartalmazza).

² Fontos megjegyezni, hogy ebben a tanulmányban a Mislevy-féle keretrendszert a nyelvvizsgáztatásban érdekelt felek szempontjából írom le, és ezt az egyes alkotóelemek specifikusabb megnevezése is tükrözheti (pl. az eredeti „terület” összetevő megnevezése ebben a tanulmányban a „nyelvhasználati terület”, mivel a téma a nyelvi mérés, nem a pedagógiai célú mérések összessége).

pedagógiai értékelés nem más, mint egy okfejtés, melynek során pár speciális helyzetben nyújtott teljesítmény alapján arra következtetünk, a vizsgázó milyen ismeretekkel rendelkezik, vagy általában mire képes.

A keretrendszer három egymásra épülő összetevőből áll (2. ábra). A *Nyelvhasználati terület elemzése* fázisban, a célcsoport várható nyelvhasználati területeinek szükségletelemzése történik. Ebben a fázisban gyűjtik össze a tesztfelkészítők a nyelvhasználati területtel kapcsolatos alapvető információkat, melyek forrása lehet például tananyag, kutatási eredmények vagy elméletek.



2. ábra. A bizonyítékalapú vizsgafelkészítési keretrendszer felépítése

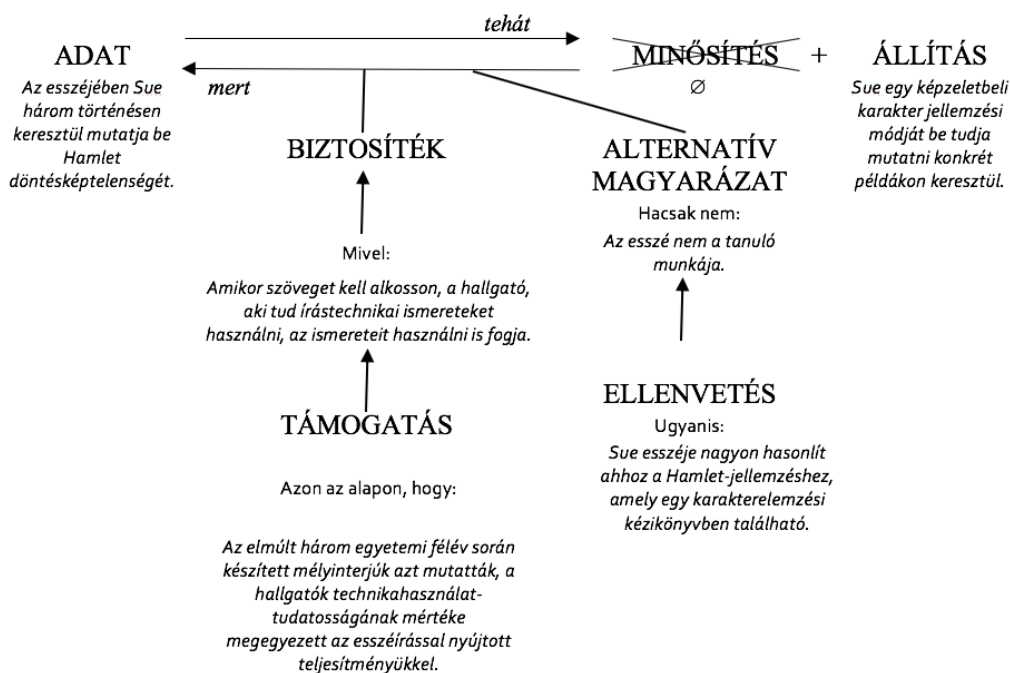
A második összetevő, a központi szerepet betöltő *Nyelvhasználati terület modellezése*, három tesztfelkészítési produktumból áll, melyek együtt a keretrendszer fő érvrendszerét képező érvelési láncot alkotják. Az első ezek közül az *Állítás*, melynek célja a nyelvvizsgázó nyelvtudásáról a vizsgateljesítmény alapján megfogalmazott kijelentéseknek pontos és érthető leírása (pl. a nyelvvizsgázó több forrásból származó információt és érvelést tud a tudományos próza követelményeinek megfelelő formában szintetizálni.). A második produktum a *Bizonyíték*, melynek célja annak meghatározása, hogy a vizsgázóról megfogalmazott állítások alátámasztásához pontosan milyen jellegű információra van szükség (pl. a fentebbi szemléltető állítás esetében a nyelvvizsgázó által megírt tudományos szövegre). Az utolsó produktumot a *Feladat* képezi, melynek célja azon nyelvhasználati helyzetek meghatározása, melyek révén az előző lépésben leírt vizsgázói tudás meglétét igazoló bizonyítékok megszerzése megvalósulhat (pl. a nyelvvizsgázó meghallgat egy előadásrészletet, majd elolvas egy szöveget, mely az előadás témájához kapcsolódik, és a két forrásból kijegyzetelt tartalom alapján megír egy esszét).

A keretrendszer harmadik összetevője az *Értékelés fogalmi kerete*, ami nem más, mint egy technikai vizsgaleírás, mely tartalmazza például a teszt szerkezetére, tartalmára, szintjére vagy az utasításokra vonatkozó előírásokat. Ezen belül hat meghatározott funkciójú egység (angolul *model/design object*) található. A *Vizsgázói egység* azt írja

le, a vizsga mit mér. A *Bizonyíték egység* a mentális konstruktumok méréséhez szükséges bizonyítékokat, illetve a pontozási eljárást részletezi. A *Feladat egység* szerepe a bizonyítékok megszerzését lehetővé tevő nyelvhasználati helyzetek meghatározása. Az *Összeállítási egység* azokat a kritériumokat tartalmazza, melyek segítségével a vizsgában szereplő feladatokat kiválasztják (pl. egy készséget hány darab és milyen típusú itemmel kell mérni). A *Bemutatói egység* az itemek és feladatok elrendezését, a vizsgázóval történő interakció jellegét, és a vizsgázói produktumok rögzítési módját szabályozza. És végül a *Vizsgalebonyolítási egység* tartalmazza a vizsga formátumára (pl. számítógépes adaptív vagy hagyományos papíralapú vizsga), és a vizsgabiztonságra vonatkozó információkat, vagy például az egyes feladatokra és vizsgarészekre megszabott időkorlátokat.

A mérőeszköz validitásának vizsgálata szempontjából Mislevy (2003) az *értékelési érvelést* tartja kulcsfontosságúnak, ami a Nyelvhasználatiterület-modellezés komponensátívelő érvelési lánc. Ez a fő érvrendszer a mérőeszközzel elicított teljesítményt, vagyis amit a nyelvvizsgáló teszhelyzetben mond, ír vagy tesz, a nyelvvizsgálóról megfogalmazott állításokkal köti össze. Az állítás a teljesítmény alapján megfogalmazott következtetés (pl. ha a nyelvvizsgáló által alkotott szöveg tartalmazza a forrásanyagokban szereplő tartalom szintézisét, a szöveg logikusan felépített, kohezív és nyelvileg jól megformált, illetve lexikogrammatikai jellemzői és stílusa tekintetében megfelel az „x” nyelvű tudományos próza sajátosságainak, akkor az arra enged következtetni, hogy a nyelvvizsgáló több forrásból származó információt és érvelést tud a tudományos próza követelményeinek megfelelő formában szintetizálni). Az így létrejött állítások alapján a nyelvvizsgálóról döntések születnek (pl. egy bizonyos osztályzatot kap, átmegy vagy megbukik egy vizsgán, felveszik vagy sem egyetemre). Mislevy (2003) szerint egy teszt validitását a teszt fő érvrendszerének validitása határozza meg.

Mislevy (2003) az érvelési lánc validitásának vizsgálatára egy érvszerkezetet leíró modellt javasol, melynek szerkezeti felépítéséhez Toulmin (2003) argumentációs modelljét vette alapul, illetve egészítette ki egy további elemmel a jogi bizonyítási eljárás során felmerülő bizonyítékok elemzésére kidolgozott ábrázolási módszerből (lásd Anderson–Schum–Twining 2005, Wigmore 1913). Mislevy adaptációja két szempontból kifogásolható: elhagyta a Toulmin által javasolt Minősítés elemet, amivel a validitás fokának jelölése szűnt meg. Továbbá az érvszerkezetet leíró modelljének kiegészített komponenseinek megnevezése hibás. A modell leírásában Mislevy (2003) az Ellenvetést alárendeli az Alternatív magyarázatnak (3. ábra), ami nem logikus, ugyanis Toulmin (2003, Toulmin, Rieke és Janik, 1984) értelmezése alapján az Ellenvetés az állítás erejét korlátozó *körülményeket* hivatott leírni. A modellt bemutató példa vizsgálatakor egyértelművé válik, hogy Mislevy Ellenvetés és Alternatív magyarázat elemei hibásan elnevezett, de funkcionális elemek. Az Alternatív magyarázat a következtetést megcáfoló érvrendszer állítása, az Ellenvetés pedig annak az alátámasztása.



3. ábra. Az argumentációs modellt bemutató példa kiegészítve a hiányzó komponenssel (Mislevy–Steinberg–Almond 2003: 12)

Az adaptált argumentációs modell állításának erejét szabályzó része azért is kifogásolható, mert Mislevy és társszerzői (2003) az Alternatív magyarázatot alternatív hipotézisnek hívják. Ha tehát az ellenérv egyik típusa egy alternatív hipotézis, akkor az érvrendszer állítása egy nullhipotézis kell legyen. Ez viszont azért értelmetlen, mert a szerzők szerint is az érvrendszer két változó, pontosabban a pontszám és annak az értelmezése közötti összefüggés meglétét hivatott igazolni. Az Ellenvetés Toulmin (2003) szerint sem egy hipotézis, csak az érvrendszeren belüli következtetés erejét gyengítő vagy a következtetést megcáfoló körülmény.

Azt a felvetést, hogy a validálási folyamat részeként nemcsak az állítások alátámasztására érdemes törekedni, hanem azok megcáfolási lehetőségeit is vizsgálni kell, korábban már Cronbach és Meehl (1955) is megfogalmazták. Azért kritizálták a validálási kutatásokat, mert azok kizárólag a megfogalmazott állítások alátámasztására törekedtek, megfelelően az állítás alternatív magyarázatainak megvizsgálásáról és, amennyiben lehetséges, azok falszifikációjáról. Messick (1989) hasonlóképpen azt hangsúlyozta, hogy a megszerzett pontszámon alapuló következtetések validálásának kétféle módon kell megtörténnie, mégpedig annak egyidejű bemutatásával, hogy amíg

egy állítást egymástól független bizonyítékok támasztanak alá, addig a lehetséges alternatív következtetések kevésbé meggyőzően bizonyíthatók. A bizonyítás, illetve cáfolás a validitási folyamat szerves és egyenrangú elemeiként kezelendők, és az alternatív következtetések az Ellenvetés egyik lehetséges típusát képezik, ahogy azt a következő részben tárgyalt érvelésalapú validálási eljárás mód esetében látni is fogjuk.

A bizonyítékalapú vizsgafejlesztési keretrendszer további kétséges pontjai közé sorolható a validitás szemlélete, mely vitathatóan leszűkített. A keretrendszer ugyanis a pedagógiai értékelés azon megközelítésén alapszik, miszerint az értékelés mindössze a vizsgahelyzetben, illetve a valós vizsgahelyzeten kívül történő nyelvhasználat összefüggéseit vizsgálja, ami az előrejelző validitás problémakörébe tartozik, és csupán a validitás egyik szelete. Az általánosíthatóság Messick (1995) egységes konstruktumvaliditás-fogalmában is csak egyik a hat szerkezeti validitást meghatározó aspektus közül.

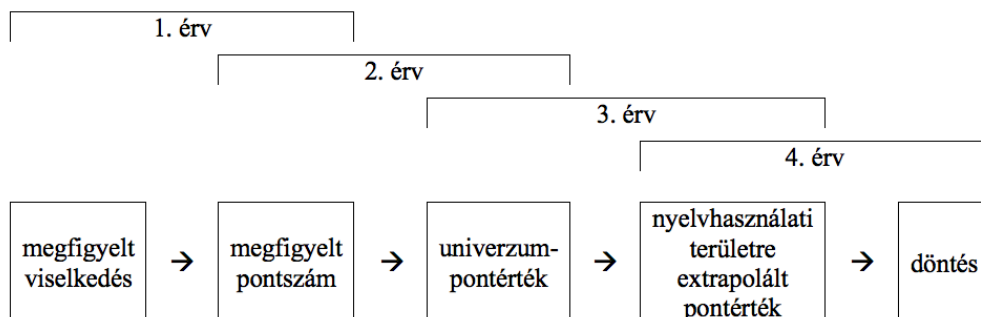
Továbbá, ahogy arra McNamara és Roever (2006) rámutattak, a keretrendszer nem foglalkozik a vizsgapontszámok felhasználásával a vizsgázóról megfogalmazott állításokon és az azokra alapozott döntéseken túlmenően. Márpedig a döntések következményeinek figyelembevételét az egységes konstruktumvaliditás fogalmának meghatározásakor már Messick (1989) is előírta.

Végezetül azt is érdemes kiemelni, hogy ugyan a keretrendszerben felépített érvelések elemzésére javasolt argumentációs modell egy logikus és jól kidolgozott elméleti modell, de az érvstruktúra vizsgálata itemszintű elemzést ír elő (Mislevy 2003), ami a gyakorlatban gazdaságosan nem kivitelezhető. Mislevy az értékelési érvelésre gondosan kidolgozott példákat hoz több elméleti megközelítés alapján (trait-elmélet, behaviorizmus, kognitív pszichológia és szociokulturális elmélet), de azok bonyolultsági foka az argumentációs modell valós életbeli felhasználhatóságát negatívan befolyásolja.

4.2. A Kane-féle tesztpontszám értelmezési és alkalmazási érvrendszer

Kane (2010, 2012, 2013) a pontszámértelmezés és alkalmazás validálási folyamatát egy olyan keretrendszerként írja le, mely segítségével a pontszám alapján megfogalmazott állítások, illetve az állítások alapján meghozott, pl. diagnosztikai vagy válogatási döntések elfogadhatóságát lehet vizsgálni. A keretrendszer alapját a Toulmin-féle argumentációs modellre (2003) épülő érvrendszer képezi. A többször módosított keretrendszerét Kane (1999, 2006) végül a pontszámok értelmezési és alkalmazási érvrendszerének nevezte el (Kane 2013). A keretrendszer négy érvből áll, ugyanis négy állítást tartalmaz, és mindegyik állításhoz önálló alátámasztás tartozik. A négy érv mindegyike kötelező elem, és együtt egy érvláncot képeznek, melyen mind a tesztfejlesztés, mind a validálási folyamat során szigorú sorrendben az elsőtől az utolsóig kell végighaladni. Ha egy állításnak a bizonyítása nem sikerül, a tesztfejlesztés vagy

validálási folyamat az állítást tartalmazó érvnél megszakad. Tehát egy valid teszt mindig egyik állítását alá kell támasztani, és a teszt validitásának erősségi fokát a leggyengébb érv határozza meg.



4. ábra. A pontszámok értelmezési és alkalmazási érvrendszere (Kane 2013)

Az első érv adat összetevője a *megfigyelt viselkedés* és az állítás összetevője a *megfigyelt pontszám* (4. ábra). Az első következtetés, amit a validálás részeként bizonyítani kell az, hogy a megfigyelt tesztpontszám a megfigyelt viselkedés, vagyis a megfigyelt vizsgázói nyelvi teljesítmény, értékeléssel megvalósított leképezése. Az érvelési láncban a második érv adat elemét a megfigyelt pontszám, az állítását pedig az *univerzum-pontérték* képezi. A második bizonyítandó következtetés arra vonatkozik, hogy a nyelvtanuló megfigyelt pontszáma megfelelőképpen reprezentálja az ő univerzum-pontértékét. Az univerzum-pontérték az itemek univerzumából véletlenül vett minta alapján kiszámított várható vizsgázói pontérték (Verhelst 2006). Ez a következtetés tehát akkor elfogadható, ha bizonyítást nyer, hogy a megfigyelt vizsgapontszám nem módosul (i) a vizsgafeladatok vagy értékelők változásával, tehát azok csereszabatosak, illetve (ii) a pontozási szabályokon vagy (iii) a vizsgaalkalmakon keresztül. A harmadik érv adata az univerzum-pontérték és állítása a *nyelvhasználati területre extrapolált pontérték*. A harmadik érvnek azt kell bizonyítania, hogy a nyelvtanuló univerzum-pontértéke megfelelően tükrözi az ő nyelvhasználati területre extrapolált pontértékét, vagyis az ő vizsgahelyzeten kívüli, de a vizsgált készségre vonatkozó nyelvhasználatát. A negyedik érv adatát a nyelvhasználati területre extrapolált pontérték, míg állítását a rajta alapuló *döntés* képezi. Az érv azt bizonyítja, hogy egy adott döntést (például hogy a nyelvtanuló készen áll folytatni tanulmányait egy egyetemen, vagy el tudja látni a munkaköri feladatait egy adott idegen nyelven) azért lehet meghozni, mert az extrapolált pontérték azt a bizonyos döntést indokolja. A négy érv által leírt következtetések ha-akkor szerkezetűek, és mindössze az állítások valószínűségét garantálják, nem minden kétséget kizáróan igazolják azokat.

Kane (2010) megközelítése alapján a pontszámértelmezési és -alkalmazási érvrendszer felállítása a tesztfejlesztéssel párhuzamosan történik. Ezt követi a meghatározott célra, egy bizonyos kontextusra, illetve egy adott populáció számára megfogalmazott pontszámértelmezési és -alkalmazási érvrendszer kritikai értékelése. Ez az érvrendszer-értékelés képezi magát a validálási folyamatot.

Ebben a megközelítésben a validálás egy ismétlődő folyamat, melynek során először a megfogalmazott értelmezési és alkalmazási érvrendszert kell megvizsgálni, majd szükség szerint újrafogalmazni. Ezt követően, az újrafogalmazott érvrendszert ismételtelen meg kell vizsgálni. A folyamat addig ismétlődik, amíg nem marad kétség afelől, hogy a pontszám alapján megfogalmazott pontszámértelmezési, illetve -alkalmazási állítások megfelelően alá vannak támasztva, nincs olyan bizonyíték, amely azokat megcáfolhatja, illetve nincs a pontszámok értelmezésének további elfogadható alternatívája (Kane 2012, 2013). Fontos kiemelni, hogy a pontszámok javasolt értelmezése és alkalmazása idővel módosulhat a változó szükségletek és tudományos magyarázatok fényében, ugyanis azok a validáláshoz szükséges bizonyítékok megváltozását eredményezhetik (Kane 2013).

Az értelmezési és alkalmazási érvrendszer vizsgálata ugyancsak több lépésből áll (Kane 2011, 2013). Először azt kell ellenőrizni, hogy az érvrendszer állításai érthetően vannak-e megfogalmazva. Ezt követi az egyes érvek kritikai értékelése. Az érveket annyira részletesen kell megfogalmazni, hogy a hihetőségük értékelése egy független tesztminősítő számára is lehetségessé váljon. Egyértelműen beazonosíthatóknak kell lenniük mind a következtetéseknek, mind pedig az alapfeltételezéseknek, különösképpen azoknak, melyek megkérdőjelezhetők. Továbbá az alapfeltételezések természetük-nél fogva maguktól értetődők, vagy pedig bizonyítékokkal megfelelően alátámasztott tartalmak lehetnek. Az alapfeltételezéseken alapuló következtetésekkel szemben támasztott követelmény az észszerűség. Nem utolsósorban a teljes érvelési lánc akkor elfogadható, ha logikus felépítésű, és kimerítő részletességgel írja le a pontszámok javasolt értelmezését vagy felhasználási módját.

Kane (2013) pontszám értelmezési és alkalmazási érvrendszerében a Toulmin-féle (2003) érvelési modell *Ellenvetés* elemének a megnevezése *Kivétel*. A megváltoztatott megnevezés viszont ugyanazt a funkciót takarja, mint az eredeti modellben. Kane tehát visszaállítja a Mislavy (2003) által kevésbé sikeresen újraértelmezett *Ellenvetés* összetevőt az eredeti állapotába. Kane validálási módszere a pontszámok alkalmazásának hangsúlyos fókuszba állítása miatt is túlmutat a Mislavy-féle megközelítésen. Kane (2011, 2013) rámutat ugyanis arra, hogy mivel az általa javasolt megközelítésben igyekezett Messick (1989) egységes konstruktumvaliditás-fogalmának átfogó jellegét megőrizni, fontosnak tartja a tesztalkalmazás hatásainak vizsgálatát. Ezáltal ebben az újragondolt validálási eljárás módban megjelennek a Messick által is fontosnak tartott szociális és politikai dimenziók.

Ellenben Kane (2013) validálási módszere sem könnyen alkalmazható a gyakorlatban. Nem írja le ugyanis azt, hogy pontosan mi mindent kell bizonyítani a validitás bizonyítási folyamat részeként. Tehát a validitás bizonyításához szükséges bizonyítékok pontos típusa továbbra sincs meghatározva, illetve az a kérdés is nyitott marad, hogy a bizonyítékok mindegyike szükséges-e a validitás bizonyításához.

Továbbá, ahogy az a pontszámok értelmezési és alkalmazási érvrendszerének bemutatása során tapasztalható (4. ábra), az érvelési lánc elemeinek értelmezése is erősen technikai jellegű, és széles körű pszichometrikai ismereteket feltételez. Emiatt Kane (2011) állítása ellenére a javasolt validálási eljárás mód csak látszólag praktikus és egyszerű. Ez lehet a magyarázata annak, hogy a széles körű pozitív fogadtatása ellenére ezt a validálási módszert nem sokan használták, és különösen kevesen alkalmazták sok vizsgázóval dolgozó pedagógiai értékelések esetében (Shaw–Crisp 2015).

4.3. A Bachman és Palmer-féle tesztalkalmazási érvrendszer

Az értékelési folyamat igazolása nevű tesztminősítési megközelítést Bachman és Palmer (2010) dolgozták ki, majd Bachman és Damböck (2018) tovább pontosították, illetve az osztálytermi értékelésre átdolgozták. Az „igazolás” megnevezés megfelel Bachman és kollégái által tudatosan kommunikálni kívánt közlésnek. Ahogy később látni fogjuk, Bachman és kollégái célirányosan kerülnek egyes méréselméleti fogalmakat. Az ő megközelítésükben az értékelési folyamat igazolása egy olyan eljárás, melynek segítségével bemutatható a tesztalkalmazás igazolhatóságának mértéke. Segítségével tehát tudományosan vizsgálható, hogy egy teszt alkalmazása mennyire helyénvaló.

Bachman és Palmer (2010) szerint bármely értékelési folyamat igazolásának feltétele egy tesztalkalmazási érvrendszer megfogalmazása, illetve az azt alátámasztó bizonyítékok összegyűjtése. A tesztalkalmazási érvrendszer egy olyan fogalmi keretrendszer, mely a tesztfejlesztés és -alkalmazás módszerét a méréselméleti követelményeknek megfelelően, de ugyanakkor a tesztfejlesztésben és -alkalmazásban érintett minden szakember számára érthető, követhető és a gyakorlatban kivitelezhető módon írja le. Az érintett szakemberek között első alkalommal jelenik meg a pontszámértelmezés történelmében explicit módon a nyelvtanár mint pedagógiai tesztfejlesztő, vagy mint minősítő döntéseket hozó tesztfelhasználó.

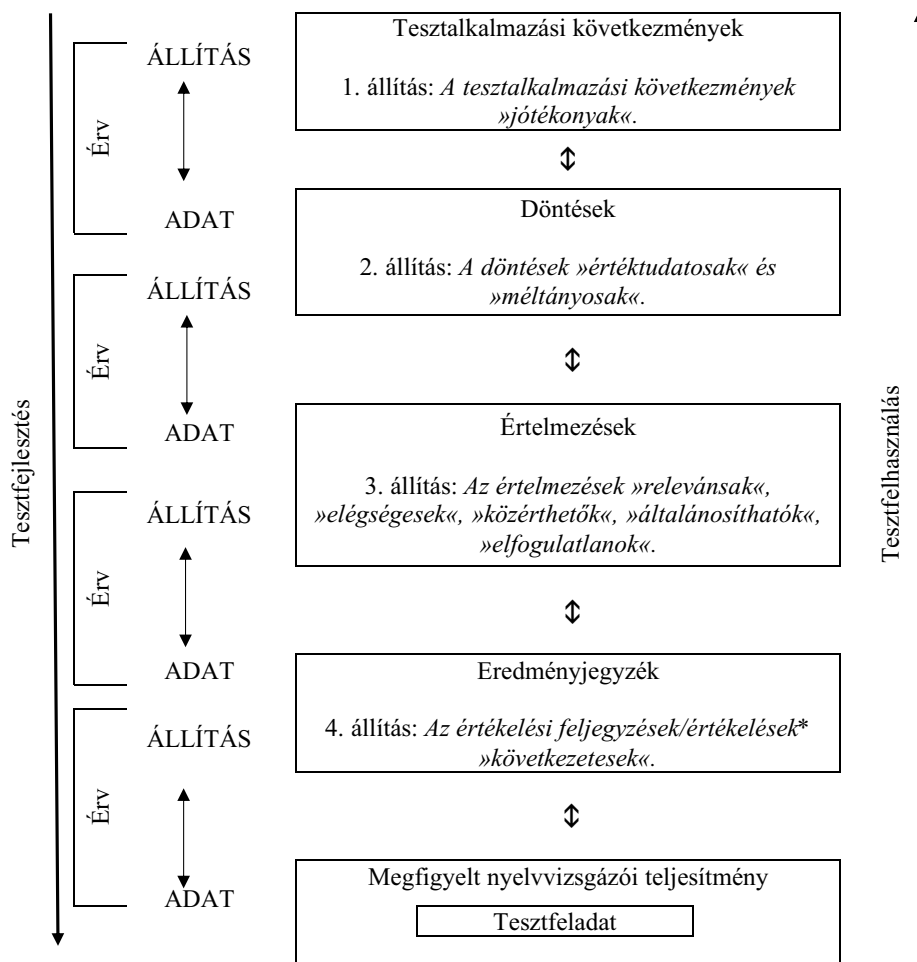
A tesztalkalmazási érvrendszer négy állítást tartalmaz (5. ábra), melyek (i) a teszt-eredmények alkalmazásának következményeire, (ii) a következményeket eredményező döntésekre, (iii) a döntéseket alátámasztó értelmezésekre, és (iv) az értelmezések alapját képező eredményjegyzék (*assessment record*) formájában (pl. összpontszám, írásos vagy szóbeli minősítés, osztályzat) történő megjelenítésére vonatkozó kijelentéseket fogalmaznak meg. Az állítások tehát a megfigyelt nyelvvizsgálói teljesítményt a tesztalkalmazás következményeivel kötik össze, és négy egymásra épülő érvből álló érvrendszert képeznek. Amint ebben az érvrendszerben az egyik érv állítása igazolást

nyer, az a következő érv adatává válik. Az érvrendszer azzal arányos mértékben elégti ki az igazságosság kritériumát, amilyen mértékben az állítások közötti következtetések a tesztben érdekelt felek számára meggyőzően alátámaszthatók és igazolhatók. A tesztfejlesztő az érvrendszerben fentről lefelé, a teszthasználó pedig az ellenkező irányba haladva építi fel az okfejtését, melynek eredménye a tesztalkalmazás igazolása úgy a tesztfejlesztő által létrehozott, mint a vizsgáztató által használt teszt esetében.

Mindegyik állításnak két fő összetevője van: az *eredmény* és egy vagy több *eredményminősítő*. Az 5. ábrán az eredmények az állítások felett látható megnevezések, míg az eredményminősítők az állítások idézőjelekkel (»...«) kiemelt részei. Így a tesztalkalmazási következményekről megfogalmazott állításban a „tesztalkalmazási következmények” az eredmény és a „jótékonyosság” az eredményminősítő. A tesztalkalmazási érvrendszer igazolhatóságának egyik fontos feltétele, hogy abban az összes eredményminősítőnek, illetve a hozzájuk kapcsolódó bizonyítékoknak szerepelniük kell (személyes beszélgetés L. Bachmannal, 2019. március 23.). Ez a megközelítés, szemben Kane érvrendszerével, pontosan meghatározza, mit kell bizonyítani az igazolási folyamat során. A tesztalkalmazási érvrendszerben használt eredményminősítők meghatározásai az 1. mellékletben található.

Azon túl, hogy mit kell bizonyítani a tesztalkalmazási érvrendszer igazolása során, Bachman és Palmer a bizonyítás menetét is leírják, mégpedig az általuk javasolt érvrendszerre specifikusan, így ebből a szempontból is felülmúlják Mislevy és Kane megközelítéseit, akik az érvek sematikus leírására ugyancsak Toulmin (2003) érvelési modelljét használják. Továbbá Bachman és Palmer a modellt hatékonyan igazították saját céljaikhoz. Ellentétben Kane-nel, aki a Toulmin-modellt mindössze a validálási keretrendszere érvelésstruktúrájának általános leírásaként használja, illetve Mislevyvel, aki hibásan módosította a modell Ellenvetés komponensét, Bachman és Palmer (2010) adaptációja a céljaiknak megfelelő több kiegészítéssel sikeresen és meggyőzően pontosítja a Toulmin-modellt, bár egyik elemét nem a klasszikus módon használja.

Bachman és Palmer (2010) az érvelési modell *Állítás* összetevőjét úgy pontosítják, hogy a kijelentés topikja az eredmény, fókusz pedig az eredményminősítő. Az *Adat* számukra a megfigyelt vizsgáztatói viselkedés vagy az érvrendszer részét képező és az érvelési láncban korábban szereplő érv állítása. A *Biztosíték* összetevő a Toulmin-modellben betöltött funkcióval rendelkezik, és ebben az érvrendszerben van a legpontosabban leírva: azazhogy a jellegét az Állítás tartalma határozza meg, vagyis mindegyik Állításhoz egyedi Biztosíték és ahhoz egyedi *Támogatás* tartozik. A szerzők az *Ellenvetés* összetevő két formáját különböztetik meg, és egy ellenérvrendszerrel egészítik ki. Az egyszerűbb változata a cáfolat, tehát például amíg az érv állítása az, hogy „A tesztalkalmazási következmények jótékonyak”, a cáfolat az ellenérvet fogalmazza meg, miszerint „A tesztalkalmazási következmények *nem* jótékonyak”. A tesztfejlesztő feladata ez esetben az ellenérv cáfolata egy cáfoló érvrendszer (lásd Tankó 2012: 83) felállításával. Az ellenvetés összetettebb változata



* a megfigyelt vizsgázói teljesítmény rögzítése pl. osztályzat, szóveges értékelés vagy százalékos értékelés formájában

5. ábra. Tesztalkalmazási érvrendszer (Bachman-Damböck 2018: 31)

a tesztben érdekelt felek által megfogalmazott kételyek, melyek eloszlatása ugyancsak cáfoló érvrendszerek segítségével valósítható meg.

Ami viszont nem a hagyományos módon van jelen a Toulmin-modell Bachman és Palmer adaptációjában, de szemben Kane megközelítésével szervesen bele van építve az érvelési folyamatba, az a *Minősítés* összetevő. Ez nem összetévesztendő az érvelési rendszerben szereplő *eredményminősítő* fogalommal, de azzal összefüggésben van.

A *Minősítés* összetevő Toulmin értelmezése alapján az állítás erősségét jelöli, és emiatt kifejezetten fontos szerepe van, ugyanis a tudományos igazolás korszerű értelmezése szerint is (pl. Popper 1997) az állítás valószínűségi értékkel rendelkezik, amely a hamis és igaz végpontok között helyezkedik el. Ha egy ellenérvet sikerül megcáfolni, akkor elvileg egyáltalán nincs erre a komponensre szükség. Ellenben, ha az érdekelt felek által megfogalmazott kételyeket csak egy bizonyos mértékig lehet eloszlatni, akkor azt egy minősítéssel jelölni kell. Bachman és Palmer az állítások valószínűségi fokának klasszikus módon történő jelölése helyett a négy érv eredményminősítőibe építik bele a valószínűség fogalmát. Ahogy az *1. mellékletben* látható, minden meghatározás azzal a kifejezéssel kezdődik, hogy az eredményminősítő valaminek a *mértéke*.

Bachman és Palmer tesztalkalmazási érvrendszere nemcsak gyakorlati, de elméleti szempontból is felülmúlja a Mislevy-féle bizonyíték alapú tesztfejlesztési keretrendszer, illetve a Kane-féle tesztpontszám értelmezési és alkalmazási érvrendszer által leírt érvelés alapú validálási eljárás módokat. A tesztalkalmazási érvrendszer a tesztfejlesztési és -alkalmazási folyamatokat minden tesztben érdekelt számára átlátható és követhető módon írja le. Bachman és Damböck (2018) útmutatásai alapján akár a nyelvtanár is készíthet tanóráira egyedül vagy kollégáival közösen elméletileg jól megalapozott, az igazságosság kritériumának megfelelő és igazolható tesztek. Akár a nyelvtanár is megvizsgálhatja, mennyire igazolható egy mások által készített teszt, melyet a nyelvórán szeretne használni, vagy a tesztalkalmazási érvrendszer segítségével felül is vizsgálhatja a saját mérési és értékelési gyakorlatát.

Mivel a tesztalkalmazási érvrendszer nyelvtanárok számára is értelmezhető nyelvezetet használ, a benne szereplő fogalmak egyértelműen vannak meghatározva, és logikusan strukturált, segítségével bármelyik tesztfejlesztési és -alkalmazási vagy mérési és értékelési folyamat érthetővé és egyben ellenőrizhetővé válik. Popper azt állítja, „egy igazolás akkor *objektív*, ha elvileg bárki ellenőrizheti és megértheti” (1997: 54–55). Tehát az érvrendszer a tesztfejlesztésben és -alkalmazásban érdekeltnek lényegesen szélesebb köre számára teremt lehetőséget az objektív igazolásra. Ezáltal több szempontból megszűnik a méréselméleti szakembereknek kiváltságos státusza, ami pozitív visszacsatolást eredményezhet nemcsak a nyelvi mérés és értékelés gyakorlatára, hanem a nyelvoktatásra is. Például sokan azt vallják, rossz gyakorlat, ha a nyelvtanárok arra készítik fel a nyelvtanulókat, hogy jól teljesítsenek standardizált teszteken. Ez valóban lehet rossz gyakorlat, de csak akkor, ha például egy teszt nem életszerű nyelvhasználatot mér, nem életszerű feladatokat tartalmaz, vagy olyan ismereteket, képességeket, készségeket vagy stratégiákat mér, melyek nem teszik lehetővé az idegen nyelv valós élethelyzetekben történő célszerű és sikeres alkalmazását. Ellenben, ha a nyelvtanár a tesztalkalmazási érvrendszer segítségével kritikusan megvizsgál egy tesztet, akkor tájékozott döntést tud hozni, miszerint azt használni fogja vagy sem az oktatási gyakorlatában, arra ’trenirozza’ a tanulóit vagy sem. Ugyanis a jó tesztfeladat a

valós élethelyzetek kommunikációs feladatainak megoldásához szükséges stratégiákat ugyanúgy aktivizálja, ahogy azokat a nyelvtanuló egy valós élethelyzetben alkalmazná. Az érvrendszert használva olyan tesztet tud fejleszteni, melyek a sikeres nyelvelsajátításhoz szükséges nyelvtanulói és nyelvtanári viselkedési mintákat erősítik. Az idegen nyelvi érettségi feladatsorok értő vizsgálata pedig lehetőséget ad a nyelvtanárnak, hogy rámutasson a feladatsorok gyengéire, és építő kritikát fogalmazzon meg annak reményében, hogy az illetékesek változtatnak a tesztfejlesztési gyakorlatukon, így a tesztek javulnak, és előnyösen hatnak vissza a nyelvtanításra.

A tesztalkalmazási érvrendszer közérthetősége részben annak tudható be, hogy a szerzők tudatosan kerülnek a vitatottságuk miatt kompromittált és túlságosan technikai jellegű *validitás*, *megbízhatóság*, *autentikusság* vagy *teszthatás* fogalmakat (Bachman–Palmer 2010: 132). Az ezek helyett megjelenő eredményminősítők közérthetőbbek, és egyben a méréselméleti követelményeknek is megfelelnek. A figyelmes olvasó bizonyára ráismert a relevancia és elégségesség eredményminősítők mögött a tartalmi validitás jellemzőire, vagy a következetesség meghatározásakor a megbízhatóság fogalmának operacionalizált változatára. Ugyancsak felismerhetők például az általánosíthatóság, illetve jótékonyág eredményminősítők sajátosságaiban az általánosíthatóság aspektus vagy a következményes validitás formában is ismert hatásaspektus Messick (1989) egységes konstruktumvaliditás-fogalmának leírásából.

A tesztalkalmazási érvrendszer közérthetősége annak is köszönhető, hogy Bachman és Palmer (2010) kérdés-felelet formájában írják le a tesztfejlesztési és -használati igazolás menetét (164–166), illetve az egyes általánosított állításokhoz tartozó biztosítékokat, támogatásokat és ellenvetéseket gyakorlatban közvetlenül felhasználható módon részletezik a tesztfejlesztő vagy felhasználó számára (158–161). Továbbá Bachman és Palmer (2010), illetve Bachman és Damböck (2018) is a tesztalkalmazási érvrendszert számos konkrét és jól követhető példán keresztül mutatják be, melyek között egyaránt megtalálhatók kis és nagy vizsgáztói létszámú, az óvodától kezdve az egyetemig használható tesztek.

A tesztalkalmazási érvrendszert az is felhasználóbaráttá teszi, hogy a tesztfejlesztést egy olyan rendszerként írja le, melyben jól elkülöníthetőek a tesztfejlesztési fázisok a tesztfejlesztési folyamat két produktumától, melyek a tesztigazolás és a teszt-előállítás (Bachman–Palmer 2010: 140). A tesztigazolás részét képezi a fentebb leírt tesztalkalmazási érvrendszer. A Mislevy- és Kane-féle megközelítések közül egyik sem tudja a tesztfejlesztés és -alkalmazást egy ilyen jól áttekinthető rendszerként bemutatni. Bachman és Palmer megközelítésében az egyes fázisok, folyamatok, valamint az azok közötti kapcsolódási pontok is vizuálisan szemléltetve vannak.

A klasszikus fogalomtár kiváltása és az azok mögöttes tartalmának egységes rendszerbe foglalása egy újabb paradigmaváltás kezdetét jelzi a mérés és értékelés tudományágban. Ugyancsak a változás jele, egy olyan egyértelműségre és közérthetőségre való törekedés a nyelvi mérés és értékelés elméletében, ami ezt a tudományágot sikeresen gyakorlatközelbe tudja hozni.

5. Konklúzió

A tanulmányban tárgyaltak szerint egyértelművé vált, hogy a mérés és értékelés korszerű elméletei szerint a validitás a teszteredmények értelmezésének és felhasználhatóságának jellemzője. Ugyancsak egyértelművé vált, hogy a validálás egy érvelési folyamat, melynek szerepe a teszteredmények értelmezésének és alkalmazásának indoklása. Az ezek által jelzett paradigmaváltás már 1955-ben megkezdődött, amikor Cronbach és Meehl kijelentették, „Nem a tesztet validáljuk, hanem azt az elvet, melynek segítségével következtetéseket fogalmazunk meg” (Cronbach–Meehl: 297). Ehhez képest az újdonságot Bachman és Palmer (2010) értékelési folyamat igazolása nevű tesztminősítési megközelítése jelenti. A tanulmányban felvázolt három megközelítés közül nemhogy az övük a legígéretesebb, hanem az egyben egy újabb paradigmaváltást is jelent a mérés és értékelés elméletében.

Bachman és Palmer megközelítésében a tesztalkalmazás a kimeneti pont, következőképpen az értékelési folyamat igazolása a mérést egészében és folyamatában vizsgálja. Továbbá, az igazolhatóság feltételeként azt is vizsgálja, hogy egy teszt megfelel-e minden eredményminősítő által leírt feltételnek. Ezáltal pontosan teljesíti azokat a követelményeket, melyeket Messick (1989) progresszív mátrixaként ismert folyamatábrája támaszt. Az eddigi legmeggyőzőbb és legteljesebb módon képezi le ugyanis az egységes konstruktumvaliditás-fogalmának osztatlan és folyamatszerű jellegét. Emiatt kijelenthető, hogy Bachman és Palmer (2010) számára sikerült eddig a leghűebben operacionalizálni Messick (1989) validitáselméletét.

A Bachman és Palmer-féle megközelítés központi elemét képező tesztalkalmazási érvrendszer egy bonyolult elméletet képez le egyértelműen, és tesz minden tesztben érdekelt fél számára követhetővé és kivitelezhetővé egy okfejtési folyamat formájában. Az értékelési folyamat igazolása megközelítés tesztalkalmazási érvrendszere révén egy rég várt paradigmaváltás tanúi vagyunk, mely a méréselméletnek és a nyelvi mérés és értékelés elméletének felhasználóközelivé válását teszi lehetővé. Felmerül viszont a kérdés, hogy ha egyes kontextusokban az 1955-ben elkezdődött paradigmaváltás sem érezteti egyelőre a teljes hatását, mikorra várható érzékelése és gyakorlatba való átültetése az újnak? Ez a tanulmány reményeim szerint megtette azokat a lépéseket, melyek segítségével ez az újabb paradigmaváltás a magyar kontextusban is megvalósulhat.

Ami Bachman és Palmer (2010) nyelvi értékelési megközelítését potenciálisan a létező leghatásosabbá és ezáltal a legfigyelemreméltóbbá teszi az az, hogy tudást oszt meg, és ezáltal erősít. Ha a szülő, a nyelvtanár, az iskolaigazgató, a minisztériumi döntéshozó ennek a tudásnak a birtokában van, akkor annak pozitív visszacsatolási hatása várható nemcsak a mérési és értékelés gyakorlatra, hanem az oktatási tevékenységekre is. Ezt a megközelítést tehát a tanárképzési felsőoktatási programok részeként és tanártovábbképzéseken oktatni kell.

IRODALOM

- Alderson, J. C. – Clapham, C. – Wall, D. (1995): *Language test construction and evaluation*. Cambridge, UK: Cambridge University Press.
- Anderson, T. – Schum, D. – Twining, W. (2005): *Analysis of evidence (2nd ed.)*. Cambridge, UK: Cambridge University Press.
- Bachman, L. – Damböck, B. (2018): *Language assessment for classroom teachers*. Oxford, UK: Oxford University Press.
- Bachman, L. – Palmer, A. (2010): *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford, UK: Oxford University Press.
- Chapelle, C. A. (1999): Validation in language assessment. *Annual Review of Applied Linguistics*, 19, 254–272.
- Cronbach, L. J. (1971): ‘Test validation.’ In Thorndike, R. L. (ed.) *Educational Measurement*. Washington, DC: American Council on Education, 443–507.
- Cronbach, L. J. (1980): Validity on parole: How can we go straight? New directions for testing and measurement – Measuring achievement over a decade – Proceedings of the 1979 ETS Invitational Conference. San Francisco: Jossey-Bass, 99–108.
- Cronbach, L. J. (1988): Five perspectives on the validity argument. In: Wainer, H. – Braun, (eds.) *Test validity*. Hillsdale, NJ, US: Lawrence Erlbaum Associates, 3–17.
- Cronbach, L. J. – Meehl, P. E. (1955): Construct validity in psychological tests. *Psychological Bulletin*, 52/4, 281–302.
- Ebel, R. L. (1961): Must all tests be valid? *American Psychologist*, 16, 640–647.
- Henning, G. (1987): *A guide to language testing: Development, evaluation, research*. Cambridge, MA: Newbury House.
- House, E. R. (1980): *Evaluating with validity*. Beverly Hills, CA: Sage.
- House, E. R. (2014): Origins of the ideas in evaluating with validity. In: Griffith, – Montross-Moorhead, B. (eds.) *Revisiting truth, beauty, and justice: Evaluating with validity in the 21st century*. New Directions for Evaluation, 142, 9–15.
- Hughes, A. (1989): *Testing for language teachers* (1st ed). Cambridge, UK: Cambridge University Press.
- Kane, M. (1999): Validating measures of performance. *Educational measurement: Issues and practice*, 18/2, 5–17.
- Kane, M. (2006): Validation. In: Brennan, R. (ed.), *Educational measurement* (4th ed.). Westport, CT: American Council on Education and Praeger, 17–64.
- Kane, M. (2010): Validity and fairness. *Language Testing*, 27/2, 177–182.
- Kane, M. (2011): Validating score interpretations and uses: Messick lecture, language testing research colloquium, Cambridge, April 2010. *Language Testing*, 29/1, 3–17.
- Kane, M. (2012): Articulating a validity argument. In: Fulcher, G. – Davidson, F. (eds.) *The Routledge handbook of language testing*. New York, NY: Routledge, 34–47.
- Kane, M. T. (2013): Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50/1, 1–73.
- Kontra, H. E. – Kormos, J. (2007): *An introduction to language testing for teachers of English*. Budapest: Okker.
- McNamara, T. F. – Roever, C. (2006): *Language testing: The social dimension*. Malden, MA: Blackwell.

- Messick, S. (1989): 'Validity.' In: Linn, R. L. (ed.) *Educational Measurement*. New York, NY: Macmillan/American Council on Education, 13–103.
- Messick, S. (1995): Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50/9, 741–749.
- Mislevy, R. J. (2003): Substance and structure in assessment arguments. *Law, Probability, and Risk*, 2/4, 237–258.
- Mislevy, R. J. – Steinberg, L. S. – Almond, R. G. (2002): Design and analysis in task-based language assessment. *Language Testing*, 19/4, 477–496.
- Mislevy, R. J. – Steinberg, L. S. – Almond, R. G. (2003): On the structure of assessment arguments. *Measurement: Interdisciplinary Research and Perspectives*, 1/1, 3–62.
- OH-HYAK (2019): Fogalomtár. *Akkreditációs Kézikönyv (53–65)*.
https://nyak.oh.gov.hu/nyat/doc/ak2018/Akkreditacios_Kezikonyv_2018.pdf
- Popper, K. R. (1997): *A tudományos kutatás logikája*. Budapest: Európa Könyvkiadó.
- Shaw, S. – Crisp, V. (2015): *Reflections on a framework for validation – Five years on*. Research Matters: A Cambridge Assessment Publication, 19, 31–37.
- Tankó, Gy. (2012): *Professional writing: The academic context* (2nd ed.). Budapest: Eötvös University Press.
- Tankó, Gy. (2015): Investigating the context validity of an English academic writing test. In: Holló, D. & Károly, K. (eds.) *Inspirations in foreign language teaching*. London, UK: Pearson Education, 128–150.
- Toulmin, S. E. (1958): *The uses of argument*. Cambridge, UK: Cambridge University Press.
- Toulmin, S. E. (2003): *The uses of argument* (Updated ed.). Cambridge, UK: Cambridge University Press.
- Toulmin, S. E. – Rieke, R. – Janik, A. (1984): *An introduction to reasoning* (2nd ed.). New York, NY: MacMillan.
- Van Eemeren, F. H. – Grootendorst, R. – Snoeck-Henkemans, A. F. (2002): *Argumentation analysis, evaluation, presentation*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Verhelst, N. (2006): Általánosíthatósági elmélet. In: Banerjee, J. – Kaftandjieva, F. – Takala, S. – Verhelst, N. *Szintillesztési módszertani segédlet. A nyelvvizsgák illesztése a Közös európai referenciakerethez című kézikönyv előzetes, kísérleti verziójához*. Budapest: Nyelvvizsgát Akkreditáló Testület – Nyelvvizsgáztatási Akkreditációs Központ, 128–150.
- Weir, C. (2005): *Language testing and validation*. Basingstoke, UK: Palgrave Macmillan.
- Wigmore, J. H. (1913): *The science of judicial proof*. Norwood, MA: J. S. Cushing.

1. MELLÉKLET

Állítástípus	Eredményminősítő
Tesztalkalmazási következmény	<i>Jótekonyság:</i> Annak a mértéke, hogy a teszt alapján meghozott döntések és a teszthasználát következményei mennyire kedvező hatásúak, és mennyire nem károsítják a tesztben érdekelt feleket.
Döntés	<i>Értéktudatosság:</i> Annak a mértéke, hogy a teszt alapján meghozott döntések és a teszthasználát következményei mennyire veszik figyelembe az aktuális nevelési és társadalmi értékeket, valamint a vonatkozó törvényeket, szabályokat és szabályozásokat. <i>Méltányosság:</i> Annak a mértéke, hogy a mért képesség azonos szintjén levő vizsgázókat milyen eséllyel sorolják be ugyanabba a csoportba.
Értelmezés	<i>Relevancia:</i> Annak a mértéke, hogy az értelmezések a tesztfelhasználó számára mennyire biztosítják a döntéshozáshoz szükséges információkat. <i>Elégségesség:</i> Annak a mértéke, hogy az értelmezések a tesztfelhasználó számára mennyire biztosítják a döntéshez szükséges mennyiségű információt. <i>Közérthetőség:</i> Annak a mértéke, hogy az értelmezések 1. mennyire tájékoztatják a tesztben érdekelt feleket a teszt által mért készség(ek)ről, 2. olyan terminológiával vannak-e megfogalmazva, ami a tesztben érdekelt felek számára érthető, és amit magukra tudnak vonatkoztatni. <i>Általánosíthatóság:</i> Annak a mértéke, hogy az értékelés alapján megfogalmazott értelmezések mennyire alkalmazhatók a vizsgázók nyelvhasználati területén belül, vagy mennyire terjeszthetők ki azokra a területekre. <i>Elfogulatlanság:</i> Annak a mértéke, hogy a feladatok formátuma és tartalma, illetve a vizsgaadminisztráció összes aspektusa mennyire mentes a vizsgázónak kedvezhető vagy ártatható részrehajlástól.
Eredményjegyzék	<i>Következetesség:</i> Annak a mértéke, hogy a diákok különböző mérések alkalmával (pl. különböző vizsgaalkalmak során, eltérő feladatokat tartalmazó tesztekkel vagy különböző javítókkal/értékelőkkel folytatott mérés esetén) azonos nyelvi képességen belül mért teljesítménye lényegében ugyanazokat az értékelési jegyzékeket eredményezi (a vizsgázó ugyanazt a pontszám, osztályzat, százalékos érték vagy szöveges értékelés formájában kifejezett eredményt éri el).

Bachman és Palmer (2010: 253) tesztalkalmazási érvrendszerében használt eredményminősítői

CSIZÉR KATA - ÖVEGES ENIKŐ

Idegennyelv-tanulási motivációs tényezők és a nyelvi vizsgák Magyarországon: összefüggések vizsgálata egy kérdőíves kutatás segítségével

The relationship between foreign language learning motivation and language exams: the results of a questionnaire study

Despite the fact that many investigations have targeted the foreign language learning motivation of secondary schools students in Hungary or the role of language exams in the compulsory Hungarian educational system, there are no recent studies specifically exploring possible relationships between variables related to foreign language learning motivation and students' dispositions towards various language exams. The aim our study is thus to investigate students' motivational and exam-related dispositions, and to map the possible relationships between these with a questionnaire study among secondary school students in Budapest. Our main results are as follows: 1. the construct that included language exam-related items proved to be a complex and multi-dimensional concept, 2. the scales related to internal motivation yielded higher mean values than the other scales, and 3. the language exam scale did not reveal any significant relationships to any of the others. As a result, it seems that the motivational impact of language exams can be only short-term at best.

Keywords: language exams, foreign language learning motivation, contact, autonomy

Bevezető

Az elmúlt évtizedben számos kutatás született középiskolások idegennyelv-tanulási motivációjáról és a közoktatás helyzetéről. Ezekből a kutatásokból egyfelől tudjuk, hogy a diákok nyelvtanulással kapcsolatos motivációja nagyrészt megfelelő mértékű, másfelől viszont az általános helyzetet vizsgáló felmérésekből az is látható, hogy a nyelvtanítás a köznevelésben még mindig nagyon vizsgaközpontú. Ezt alátámasztja Öveges és Csizér 2017/2018-ban lefolytatott, a köznevelésben folyó nyelvtanítás sajátosságait feltáró kutatása is, melyben több ponton hangsúlyos elemként jelennek meg a válaszokban a nyelvi vizsgák (Öveges–Csizér 2018). A nagymintás, a hazai általános és középiskolákra reprezentatív felmérésben egy intézményvezetői kérdőív keretében adott válaszok alapján a középiskolai igazgatók iskolájuk három legfontosabb, nyelvtanítással kapcsolatos eredménye között a sikeres vizsgákat szinte mindig megemlítik (Tartsay–Tiboldi–Katona 2018). Az intézményvezetők által felsorolt összes legfontosabb eredmény közel kétharmadát tették ki az érettségire, nyelvvizsgára vagy egyéb nyelvi mérésekre vonatkozó említések (például gimnáziumban az összes kiemelt eredmény

65%-a), de már általános iskolában is megjelenik ez a vonal (az összes siker 15%-a). A felmérésben rákérdeztek az intézményen belül, de tanórán kívül megvalósított nyelvtanulási lehetőségekre is, és az eredményekből jól láthatóan kirajzolódott az a tény, hogy a formális nyelvi megmérettetésekre történő felkészítés az összes ilyen alkalom igen jelentős hányadát teszi ki (például gimnáziumban az összes idevágó válasz 34%-a adott meg érettségire, valamint 29%-a nyelvvizsgára történő gyakorlást, ugyanez a szakgimnáziumokban 55% és 24%).

Az intézményvezetői válaszok mellett a kutatás egy papíralapú kérdőívvel a 7. és 11. osztályos nyelvtanulók véleményét is feltárta (Albert–Tankó–Piniel 2018). A kapott adatok fényében elmondható, hogy a nyelvi vizsgák szempontjából releváns tizenegyedikes évfolyam nagy szerepet tulajdonít a nyelvvizsgának: egyötödük már letette és 67 %-uk tervezi ezt teljesíteni a közeljövőben. A tanulói válaszok tükrében a sikeres vizsgákat a diákok fontos, de nem elsődleges célnak tekintik. A megkérdezett nyelvtanárok véleménye megerősíti az előbbieket: a nyelvi vizsgákat (mind az érettségit, mind a nyelvvizsgát) fontos motiváló tényezőnek látják. Emellett azonban a pedagógusok megjelenítik szakmai kétségeiket azzal kapcsolatban, hogy valóban megfelelő eszköze-e a nyelvtudás mérésének a nyelvvizsga (Csizér–Illés 2018). A fentiekkel összhangban lévő eredményeket találtak korábbi idevágó empirikus kutatások is. Egy a középfokú szakképző intézményekben folyó nyelvoktatást és nyelvtanulást vizsgáló jelentésből szintén az derül ki, hogy mind az intézményvezetők, mind a nyelvtanárok a sikeres nyelvi érettségiket és nyelvvizsgákat sorolják a legfontosabb elért eredmények közé (Nikolov–Ottó–Öveges 2009).

Mindezen eredmények ellenére nem tudunk olyan kurrens vizsgálatról, amely kutatná, hogy a nyelvi vizsgákkal kapcsolatos hozzáállás milyen kapcsolatot mutat magával a motivációval a diákok szemszögéből. Ezért jelen tanulmányunkban a következő kutatási kérdésekre keressük a választ:

- Hogyan jellemezhetőek a középiskolások nyelvi vizsgákkal kapcsolatos véleményei?
- Milyen szerepet játszik a nyelvvizsgákkal kapcsolatos beállítódás az idegennyelv-tanulási motivációval kapcsolatosan?

A kérdések megválaszolására egy kérdőíves kutatást terveztünk, amely során középiskolások diákoktól gyűjtöttünk adatokat. A jelen tanulmányban közölt eredmények egy nagyobb kutatás részét képezik, illetve közvetlenül kapcsolódnak Albert, Piniel és Lajtai (2019) tanulmányához. Az irodalmi áttekintésünkben először a nyelvi vizsgák szerepét, majd a nyelvtanulási motivációval kapcsolatos legfontosabb eredményeket foglaljuk össze. A módszer és az eredmények ismertetése után nemcsak további kutatásokra teszünk javaslatot, hanem nyelvpedagógiai és nyelvpolitikai következtetéseket is levonunk.

A nyelvi vizsgák szerepe a hazai nyelvvoktatásban

A vizsgák szerepe a hazai nyelvvoktatásban régóta meghatározónak tűnik, és a sikeres vizsga ma is „útlevél” lehet a továbbtanuláshoz és a szakmai boldoguláshoz (Fekete–Csépes 2019). Két közelmúltban született oktatáspolitikai döntés még tovább erősítette ezt annak ellenére, hogy az egyik időközben visszavonásra került. Az azóta törölt szabályozás a felsőoktatásba történő bemenetet érintette: 2020-tól előírták a B2 szintű nyelvtudás igazolását mint a felsőoktatásba történő felvételi követelményt, azaz csak azok jelentkezhettek volna egyetemre, akik letettek egy emelt szintű nyelvi érettségi vagy középfokú államilag elismert nyelvvizsgát. Az intézkedést 2019 novemberében hatálytalanították, a sajtóban megjelent nyilatkozatok alapján véglegesen, a szakmai és hallgatói szervezetek észrevételei alapján. A másik, érvényben maradt rendelet (53/2017. számú kormányrendelet) szerint 2018-tól a fiatal korosztály számára egy sikeresen teljesített akkreditált komplex nyelvvizsga díját az állam visszatéríti, és ugyanez vonatkozik a középiskolai tanulmányokon túl letett emelt szintű érettségi térítésére is. Mindkettő üzenete további lökést ad ahhoz, hogy a nyelvtanulás fő céljaként a hazai nyelvtanulók a sikeres nyelvi vizsgát tekintsek, és ennek feltétlenül hatása van a nyelvórákon folyó munkára. Nem csak nálunk van ez azonban így: az Európai Bizottság egy nemrégiben kiadott tanulmánya (2015) arra a megállapításra jutott az európai országokban zajló országos nyelvi mérések összehasonlító vizsgálata alapján, hogy az államok többsége folytat formális nyelvi mérést mind általános, mind középiskolában, és ezek jelentősége folyamatosan erősödik. A mérések általában nagy tétellel, meghatározó erővel bírnak a nyelvtanulók további előrehaladásában, és sokszor az eredmények külön kiemelten, tanúsítványban is megjelennek. A különböző országokban a méréseket jellemzően a Közös európai referenciakeret (KER) (Európa Tanács 2002) szintrendszeréhez illesztették, és azok mind a négy alapkészséget mérik. A szintek vonatkozásában a hazai mérések többnyire szinkronban vannak más országok gyakorlatával, bár az általános iskolai vizsgák sorában máshol megjelenik a hazai idevágó vizsgák palettáján nem szereplő B1 szint, és olyan országok is vannak, ahol a középiskolában a kötelező vizsgák kimeneti szintjét C1-re teszik, amely szintén nem jelenik meg a mi középiskoláinkban a normál nyelvi képzésben (Öveges 2018).

Hazánkban a köznevelésben tanulók számára több kötelező és opcionális formális nyelvi mérés létezik. Általános iskolában 2015 óta minden hatodikos (A1+ szint) és nyolcadikos (A2 szint) részt vesz a tanév végén megszervezett idegen nyelvi mérésekben (kivéve a két tanítási nyelvű intézmények diákjait, számukra a 2014-ben bevezetett célnyelvi mérés a kötelező), melynek eredményeit az Oktatási Hivatal a következő tanév végén teszi elérhetővé honlapján (https://www.oktatas.hu/koznevelles/meresek/idegen_nyelvi_meres/eredmenyek). Bár ezeken a méréseken mindenkinek részt kell vennie, hiányzás esetén a tesztet pótolni nem kell, és az elért pontszámokat nem rendelik hozzá a diákok egyéb törzsadataihoz, azaz azok nem hatnak ki formálisan

az előrehaladásukra. Az eredményeiket ellenőrizhetik az adatbázisban, de erről tanúsítványt nem kapnak. Összesítéseket ezek vonatkozásában az Oktatási Hivatal készít évente (Kákonyi 2016, Oktatási Hivatal 2018a, 2018b), valamint a célnyelvi mérések bevezetésének tapasztalatairól egy tanulmány is készült (Nikolov–Szabó 2015). Az idegen nyelvi és célnyelvi mérések elsődleges célja a haladásról történő mintavétel a további tanulási folyamat érdekében. Bár az akkreditált nyelvvizsgák 14 éves kortól letehetőek, az általános iskolai korosztály, azaz a 14 év alattiak részére is egyre több külső nyelvvizsga áll rendelkezésre, mint például az ELTE Origo Nyelvi Centrum Junior nyelvvizsgálója angolból vagy németből a Goethe-Zertifikat A1: Fit in Deutsch 1 vizsga, melyeket 10 éves kortól lehet letenni. Ezek a junior vizsgák nem részei az akkreditált nyelvvizsgák rendszerének, mert államilag elismert vizsgára csak 14 éves kortól lehet jelentkezni. Az eredmények nem beszámíthatóak a későbbi továbbtanulásban például, mégis egyre népszerűbbnek tűnnek. Pontos és nyilvános számadatok azonban a részvételről nem állnak rendelkezésre.

Ezzel szemben részletes adatok érhetőek el az államilag elismert nyelvvizsgák vizsgázóiról, melyekből egyértelműen kirajzolódik, hogy a vizsgázói populációban szignifikáns a köznevelésben tanulók aránya: 2018-ban például az Oktatási Hivatal Nyelvvizsgáztatási Akkreditációs Központjának (NYAK) honlapján található statisztikai adatok alapján az összes vizsgázó 45 százaléka tartozott a 14–19 éves korosztályhoz. A fent leírt intézkedésekkel (a B2 szint igazolása a felvételi eljáráshoz és a vizsgadíj visszatérítése) feltehetőleg ez az arány tovább tolódik majd az ilyen korú nyelvtanulók felé, amely tovább erősítheti majd azt a visszahatást, amellyel a felnőtt korosztály számára kifejlesztett nyelvvizsgákra történő felkészülés, felkészítés gyakorol az iskolai nyelvi órákra és a tanítási-tanulási folyamatra (Víg 2012, 2013). Az akkreditált vagy más néven államilag elismert nyelvvizsgák rendszere 2000 óta létezik (Bárdos 2015), immár több tucat vizsgaközpontban lehet nyelvvizsgát tenni az élő és klasszikus nyelvek széles köréből, általános és szakmai nyelvre vonatkozóan. A NYAK adatai szerint legtöbbször angol nyelvből tesznek vizsgát, ennek aránya 2018-ban például az összes vizsgázó vonatkozásában 74 százalék volt, de nagyon hasonlóak az egyéb években és az egyes korosztályok vonatkozásában tapasztalt arányok is. Érdekes még a nyelvekkel kapcsolatban, hogy a 14–19 éves korosztályban elenyésző az eszperantó választók száma az idősebb vizsgázókhöz képest. Ennek több oka lehet, például az, hogy a nyelvválasztást a középiskolásoknál még kevésbé pragmatikus okok határozzák meg, mint esetleg később, amikor az egyes élethelyzetek miatt az elért nyelvtudás fontossága háttérbe szorul a megszerzett nyelvvizsgához képest. A vizsgák szóbeli és írásbeli részekből állnak, a kettő együtt testesíti meg a komplex vizsgát, mely várhatóan bemeneti követelmény lesz a felsőoktatásba (jelenleg kimeneti követelmény a diplomához), és amely most is előnyökhöz juttathatja a sikeres vizsgázót (például többletpont a felvételi eljárásban vagy többletjuttatás a fizetésben). A vizsgákat alap- (KER szerinti B1),

közép- (B2) és felsőfokon (C1) lehet letenni, ezek közül ma Magyarországon a B2 szinten jelenik meg a legtöbb vizsgázó: a NYAK számai alapján 2018-ban az összes vizsgaesemény 81 százaléka történt középfokon.

A középfokú oktatásban a közneveléshez képest külső, opcionálisan letehető nyelv-vizsgák mellett kötelező nyelvi mérés az érettségi. 2004-től az érettségi vizsga két-szintűvé vált, a vizsgareformról több összegzés készült (Alderson–Nagy–Öveges 2000, 2001, Einhorn 2007, 2008). A két szintből a középszintű vizsga (KER szerinti B1) le-tétele mindenki számára kötelező. Idegen nyelvekből is választható emelt szinten az érettségi, mely 2005-től (a korábbi gyakorlatot gyökeresen megváltoztatva) 60 százalék feletti eredmény esetén egyenértékű a B2 szintű akkreditált nyelvvizsgával az adott nyelvből. Emelt szinten a vizsgák aránya jellemzően alacsony, az Oktatási Hivatal ada-tai alapján (Érettségi 2018) 2018-ban például az összes, különböző tantárgyakból letett érettséginek csak alig több mint 10 százaléka volt emelt szintű. A vizsga szintjének ki-választását idegen nyelvből több tényező, például a továbbtanulási szándék vagy a ta-nári elvárások befolyásolják (Vígh 2012). Mindkét szinten központilag összeállított írásbeli feladatsorral mérnek, a szóbelinél azonban ez csak az emelt szintre igaz. Az értékelés középszinten belső, azaz a diák saját nyelvtanára végzi, míg az emelt szintű vizsga külső vizsgáztatókkal folyik. A kétszintű rendszert idegen nyelvekből 2017-ben finomították azzal a céllal, hogy a KER szerinti szinteket megbízhatóbban közelítsék, és a különböző nyelvekből a követelmények egybeessenek. Az érettségi eredmények az Oktatási Hivatal honlapján elérhetőek, de statisztikai elemzésükről kevés összesítés készült (Vígh 2013). Idevágó tanulmányok leginkább az angol és a német nyelvvel kapcsolatban születtek (Einhorn 2007, 2008, Halápi–Hegedűs 2014, Jilly–Nagy 2010, Sominé–Vígh 2010, Vígh 2013). Az érettségi vizsga mindkét szinten letehető előre-hozott formában, azaz a középiskolai tanulmányokat lezáró félév vége előtt.

Egy 2018-as kvalitatív kutatásban a szerzők 32 hazai köznevelési intézmény helyi tantervét, annak is idegen nyelvi részeit vetették alá dokumentumelemzésnek (Öveges–Csizér–Dér 2019). A szövegrészeket több szempontból vizsgálták meg, köztük feltárták a nyelvi vizsgák szerepét, súlyát is ezekben a helyi szintű szabályozókban. Megállapí-tották, hogy az egyéb felmérésekből kirajzolódó domináns szerepük ellenére, az iskolai tantervekben igen kevés említés esik a nyelvi vizsgákról és mérésekről. A leggyakrabban említett vizsga a nyelvi érettségi volt, különös tekintettel az erre történő felkészülés le-hetőségeire, míg a nyelvvizsga összesen két középiskolánál szerepelt, az idegen nyelvi mérés pedig egyáltalán nem jelent meg a vizsgált általános iskolák helyi tanterveiben. A diagnosztikus idegen nyelvi és célnyelvi mérések vonatkozásában, természetükből adódóan helyénvaló, hogy ezekre történő felkészülés nem található meg a tantervekben, és bár más szempontból, annak külső, az iskola és a köznevelés keretein kívülsége miatt ugyanez igaz a nyelvvizsgákra is, azaz az érettségi mint leggyakoribb elem az in-tézményeknek a nyelvi mérésekhez való megfelelő hozzáállását tükrözheti.

Az idegennyelv-tanulási motiváció vizsgálata hazánkban

Magyarországon évek óta népszerű kutatási téma a nyelvtanulók motivációjának vizsgálata, amely során a kutatók azt próbálják feltérképezni, hogy milyen tényezők befolyásolhatják a nyelvtanulás eredményességét, hogy a nyelvtanulók mennyi erőfeszítést hajlandóak az idegen nyelvek tanulására fordítani, azaz hogyan írható le a motivált tanulási viselkedésük. Középiskolások körében számos adatfelvétel áll rendelkezésünkre az elmúlt évtizedekből, amelyekben nemcsak általános motivációs eredményeket, hanem több motivációval kapcsolatos határterületet is feltártak a kutatók. Az alábbi rövid összefoglalóban a teljesség igénye nélkül bemutatjuk azokat a nyelvtanulási motivációval és autonómiával, illetve a kontaktusélményekkel kapcsolatos vizsgálatokat, amelyek előzményeit jelentik a jelen felmérésünknek.

Az egyik első tanulmány, amely a magyarországi középiskolások nyelvtanulási motivációjával foglalkozott, Clément, Dörnyei és Noels (1994) munkája volt, amelyben a kutatók 11. osztályos diákok budapesti mintáján mérték, hogy a nyelvtanulók attitűdjei, szorongása, illetve az osztálytermi folyamatokkal kapcsolatos véleményük miként befolyásolja a nyelvtanulási motivációjukat. Eredményeik alapján a nyelvtanulási motivációra nemcsak az attitűdök, hanem a nyelvi önbizalom és az összetartó csoport megléte is pozitívan hatott. Egy kicsit későbbi vizsgálat már az eszközjellegű motívumokra koncentrált hasonló korcsoportnál. A Csapó Benő és Nikolov Marianne nevével fémjelzett országos reprezentatív adatfelvétel során, 10. osztályosok köréből szerzett eredmények alapján, a diákok nyelvtanulási motivációját vizsgákkal kapcsolatos eszközjellegű motívumok határozták meg. Az érettségi vizsga, a középfokú nyelvvizsga és a felsőfokú nyelvvizsga megszerzése motiválja leginkább a legtöbb diákot (Csapó 2001, Nikolov 2003a, 2003b).

Jelen munkánk szempontjából fontosak Dörnyei Zoltán (2005, 2009) „második nyelvi motivációs érendszerével” kapcsolatos magyarországi eredményei is. Az elmélet szerint a diákok motivációját három tényező befolyásolja: 1. Az ideális második nyelvi énképük, azaz azok az elképzelések és vágyak, amelyeket a nyelvtanuló szeretne elérni. 2. A szükséges második nyelvi énképük, amely énkép külső motívumokat foglal magában. 3. A nyelvtanulási tapasztalatok, azaz olyan nyelvtanuláshoz kapcsolódó motívumok, amelyek leírják, milyen nyelvtanulási élményekkel és tapasztalatokkal rendelkeznek a különböző diákok. Bár ez az elmélet nagyon sok kutatást inspirált a világ számos országában az évek során (Boo–Dörnyei–Ryan 2015), ma már egyre több kritika is éri az eredeti elképzelést (Csizér megjelenés alatt). A magyarországi első vizsgálat sem tudta mindhárom tényező jelenlétét bizonyítani: Kormos és Csizér (2008) regresszióelemzésen alapuló eredményei alapján csak az ideális második nyelvi én és a nyelvtanulási élmények járultak hozzá szignifikánsan a középiskolások idegen nyelvi motivációjának alakulásához, a szükséges nyelvi ént nem sikerült megbízhatóan mérni. Későbbi vizsgálatok azonban már sikerrel jártak mindhárom tényező mérését illetően

hazánkban, és bemutatták azt is, hogyan hat a tanult nyelv a motivációra (Csizér–Lukács 2010), illetve milyen hatása van a tanároknak és szülőknek ezekre a tényezőkre (Csizér–Galántai 2012). Nem tudunk viszont olyan vizsgálatokról, amelyek a nyelvi vizsgák szerepét tárnák fel ezen motivációs tényezőkkel kapcsolatban.

A tanulói autonómiával kapcsolatos vizsgálatok is foglalkoztak középiskolásokkal Magyarországon, hiszen a nyelvórákon eltöltött idő nem ad elég lehetőséget a nyelv gyakorlására, ezért mindenképpen fontos lenne, hogy a diákok felelősséget vállaljanak saját nyelvtanulási folyamatukért, és különböző tanulási és nyelvhasználati stratégiák segítségével önállóan, az iskolán kívül is fejlesszék a nyelvtudásukat (Little 1999). Különösen fontos az autonómia vizsgálata középiskolások körében, mert több kutatás is megmutatta, hogy ebben az életkorban csökken a diákok autonómiája (D. Molnár 2014). Albert, Tankó és Piniel (2018) általános és középiskolások idegennyelv-tanulási autonómiáját illetően találtak szignifikáns különbséget: a 11. évfolyamos diákok eredményei alacsonyabbak voltak, mint a 7. évfolyamosoké.

Egy korábbi országos kutatás bemutatta azt is, hogy az autonómia vizsgálata nemcsak önmagában fontos, hanem azért is szükségszerű, mert az autonómia kapcsolatban áll más olyan tényezőkkel is, amelyek pozitívan befolyásolhatják a nyelvtanulás sikerességét. Csizér és Kormos (Csizér–Kormos 2012, 2014, Kormos–Csizér 2014) első lépésben korrelációs elemzéssel bizonyította a motiváció és az autonómia szoros kapcsolatát, majd strukturális elemzéssel azt, hogy a motiváció közvetlenül befolyásolja az autonóm tanulási viselkedést is. Szükséges lenne azonban azt is megérteni, hogy a meglehetősen vizsgaközpontú oktatási rendszerünkben (Öveges–Csizér 2018) milyen kapcsolat van a diákok nyelvi vizsgákra vonatkozó elképzelései és az autonóm tanulási viselkedésük között.

A nyelvtanulás sikerességét illetően fontos feltérképezni a diákok kontaktusélményeit is: mennyit használják az idegen nyelvet az osztálytermen kívül, amelyek kapcsolódnak a motivációhoz is, és befolyásolhatják a tanulás sikerességét is. Bár Kormos és Csizér (2007) végzős általános iskolások körében végzett interjúkutatást arról, hogy milyen típusú kontaktusélményekről számolnak be a diákok, feltételezhetjük, hogy a feltárt tipológia az idősebb korosztályokra is érvényes. Ez a tipológia három fő csoportba osztja a kontaktusélményeket: direkt kontaktus, indirekt kontaktus és kulturális termékeken keresztüli kontaktus. A kutatás további meghatározó eredménye, hogy nemcsak maga a kontaktusélmény befolyásolja a motivációt, hanem az is, hogy a nyelvtanuló milyen fontosságot tulajdonít az élményeknek. Egy nagymintás, angol nyelvvel kapcsolatos vizsgálatban (Csizér–Kormos 2009) kiderült, hogy a motivált tanulási viselkedést csupán két tényező befolyásolta közvetlen módon: a nyelvvel és beszélőikkel kapcsolatos pozitív attitűdök és a kontaktusélményeknek tulajdonított fontosság. Mivel ezek a vizsgálatok már több mint tíz éve készültek Magyarországon, a helyzet természetesen megváltozhatott, így további vizsgálatokra van szükség.

Összefoglalva tehát az alábbi kutatásunkban a motivált tanulási viselkedéshez kapcsolódóan a fenti áttekintés alapján a következő tényezőket mérjük: 1. A második nyelvi motivációs énrendszer három tényezőjét: az ideális és szükséges második nyelvi ént és a nyelvtanulási tapasztalatokat. 2. A kontaktusélmények közül a kulturális kontaktust, direkt kontaktus és a kontaktus fontosságát. 3. Az autonómiával kapcsolatosan pedig a technológiai alapú önálló tanulást és a tanulói autonómiát (autonóm tanulási viselkedést).

A kutatás módszere röviden

A fent leírt két nagy téma közti összefüggések vizsgálatára egy kérdőíves kutatást terveztünk, mert átfogó képet szerettünk volna kapni a változók közti kapcsolatrendszeréről. A felmérés céljaira összeállított kérdőívet egy budapesti középiskola diákjai töltötték ki. Fontos volt, hogy az eredmények egyetlen iskolából érkezzenek, mert így tudtuk az iskolák közti különbségek hatását kiküszöbölni (Csizér–Kormos 2008). A papír alapon készült magyar nyelvű kérdőívünket összesen 53 tanuló töltötte ki, a 10. és 11. évfolyamról, a nemek aránya kiegyenlített volt: 26 fiú és 27 lány került a mintánkba. Átlag életkoruk 17 év volt, és átlagosan 9 éve tanultak angolul. A mintába került diákok nyelvi szintje a KER szerinti B1, illetve B2 szint volt, és megközelítőleg harmaduk rendelkezett többnyire középfokú nyelvvizsgával a felmérés időpontjában.

Az adatgyűjtéshez egy a vizsgálat céljaira kifejlesztett kérdőívet használtunk, ami 124 állítást (öt pontos Likert-skála) és 6 háttérkérdést tartalmazott. Összesen 22 skálába tömörítettük a mérni kívánt konstruktumokat. A nyelvi vizsgákkal kapcsolatos beállítódásokat a jelen kutatásban fejlesztettük ki, nagyban támaszkodva Öveges és Csizér (2018) kutatására. A többi skála esetében korábbi hazai vizsgálatokban már használt állításokat gyűjtöttünk egybe (Csizér–Kormos 2009, 2012, Kormos–Csizér 2008), amelyekből 10-et használunk fel tanulmányunkban. Ezek a következők voltak:

1. *A nyelvi vizsgákkal kapcsolatos beállítódást 9 különböző állítással mértük, hogy a diákok mit gondolnak a nyelvvizsgák és érettségi vizsgák fontosságáról és motiváló szerepéről. Példa: A nyelvvizsga-bizonyítvány nagyon fontos bizonyítéka a nyelvtudásnak.*
2. *A motivált tanulási viselkedés (5 állítás): ez a skála leírja, hogy mennyi erőfeszítést hajlandóak a diákok az angol nyelv tanulásába fektetni. Példa: Bátran mondhatom, hogy mindent megteszek azért, hogy nagyon jól megtanuljak angolul.*
3. *Az ideális második nyelvi énkép (5 állítás) magában foglalja azokat az elképzeléseket, amelyeket a diákok jövőbeli nyelvtudásukról és nyelvhasználatukról alakítanak ki. Példa: Amikor a jövőmre gondolok, az angol nyelv használata fontos része az elképzeléseimnek.*
4. *A szükséges második nyelvi énkép (5 állítás) írja körül azokat a motívumokat, amelyekről a diákok külső hatás miatt gondolják, hogy számukra fontos. Példa: A körülöttem élő emberek számára az angolnyelv-tudás az általános műveltség része.*

5. *A nyelvtanulási tapasztalatok* (4 állítás), azaz olyan nyelvtanuláshoz kapcsolódó motívumok, amelyek leírják, milyen nyelvtanulási élményekkel és tapasztalatokkal rendelkeznek a különböző diákok. Példa: *Tetszenek azok a dolgok, amiket az angolórákon szoktunk csinálni.*
6. *A kulturális kontaktus* (4 állítás) skálája foglalja össze a diákok angol nyelvű kulturális termékekkel kapcsolatos szokásait. Példa: *Szoktam filmeket, sorozatokat, videókat nézni angolul.*
7. *A kontaktus fontossága* (6 állítás): milyen fontosságot tulajdonítanak a diákok az angol nyelv használatának anyanyelvű vagy nem anyanyelvű beszélővel az osztálytermen kívül. Példa: *Szerintem azért jó külföldiekkel beszélgetni, mert megismerhetem beszédstílusukat, kiejtésüket, szokincsiüket.*
8. *A direkt kontaktus* (6 kérdés): milyen gyakran használják az idegen nyelvet a diákok anyanyelvű vagy nem anyanyelvű beszélővel az osztálytermen kívül. Példa: *Milyen gyakran szoktál angolul beszélni külföldi ismerősökkel/barátokkal?*
9. *A technológiai alapú önálló tanulás* (6 állítás): a diákok képessége arra, hogy nyelvtudásuk fejlesztésére önállóan használják az internet és számítástechnika adta lehetőségeket. Példa: *Gyakran használom az internetet, hogy fejlesszem a nyelvtudásomat.*
10. *Az autonóm tanulási viselkedés* (11 állítás): mennyire képesek a nyelvtanulók önállóan tanulni és gyakorolni az angol nyelvet. Példa: *Több időt fordítok annak a gyakorlására, ami nehezebben megy számomra az idegennyelv-tanulásban.*

A terepmunka során először az intézmények vezetőinek engedélyét kértük a kérdőív kitöltésére, amely névtelen és önkéntes volt. A nyelvtanulók a projekt egyik kutatójának jelenlétében válaszoltak a kérdésekre egy iskolai óra keretében 2019 januárjában. A kérdőív kitöltése körülbelül 25 percet vett igénybe. Az adatok elemzéséhez az SPSS-program 20. verzióját használtuk, és a kutatási kérdéseink megválaszolásához a leíró statisztikán túl faktorelemzést és korrelációs elemzéseket végeztünk.

A kutatás eredményei

Nyelvi vizsgákkal kapcsolatos vélemények

Az 1. táblázat tartalmazza a nyelvvizsgákkal kapcsolatos állítások leíró elemzését. Az átlagokból kirajzolódik, hogy a diákok nem gondolják azt, hogy a nyelvvizsga hosszú távon is bizonyítéka lehet a nyelvtudásnak: nyelvhasználat nélkül nem lehet megtartani szerintük a nyelvtudást. A nyelvvizsgákkal és érettségivel kapcsolatos kérdések nem mutatnak egyértelmű eredményt. Egyrészt a páros t-teszt szignifikáns különbséget mutat a következő két állítás között: Sokszor csak azért tanulok angolt, hogy készüljek a nyelvvizsgára. / Azért tanulok angolul, hogy minél jobban sikerüljön az érettségim. ($t = -3,592$; $p = 0,001$). Ez azt mutatja, hogy az érettségít kicsit fontosabbnak tartják a

nyelvvizsgánál, bár még a magasabbik átlag is alacsonynak mondható. Ugyanakkor viszont a nyelvvizsga szintjének motiváló erejét mérő állítás (Azért tanulok angolul, hogy minél magasabb szintű nyelvvizsgát tudjak tenni) átlaga szignifikánsan magasabb a két előző állítás átlagánál, amelyből az egyik előbbi szintén a nyelvvizsgával kapcsolatos állítás volt ($t = 5,227$; $p < 0,001$; $t = -2,627$; $p = 0,011$).

	Átlag	Szórás
Az tud valóban angolul, akinek nyelvvizsgája van.	1,74	0,90
Sokszor csak azért tanulok angolt, hogy készüljek a nyelvvizsgára.	2,11	0,93
Azért tanulok angolul, hogy minél jobban sikerüljön az érettségim.	2,60	0,88
Az angol érettségi nagyon fontos bizonyítéka a nyelvtudásnak.	2,85	1,26
Ha le is teszem az előrehozott érettségit angolból, azért még szeretnék utána angolóra járni.	2,87	1,39
Azért tanulok angolul, hogy minél magasabb szintű nyelvvizsgát tudjak tenni.	3,00	1,09
A nyelvvizsga-bizonyítvány nagyon fontos bizonyítéka a nyelvtudásnak.	3,36	1,08
A nyelvvizsga megszerzése nem jelenti azt, hogy abba lehet hagyni az angoltanulást.	3,93	1,11
Hiába van nyelvvizsgája az embernek, ha nem használja a nyelvet, akkor elfelejti azt.	4,02	0,93

1. táblázat. Nyelvi vizsgákkal kapcsolatos állítások átlaga és szórása (ötfokú skála)

A fenti inkonzisztenciák miatt felmerült a gyanú, hogy a nyelvi vizsgákkal kapcsolatos állítások valójában nem egyetlen látens dimenziót írnak le, hanem többet, így a kilenc állítást faktorelemzésnek vetettük alá, hogy lássuk, milyen struktúra húzódhat meg a háttérben. Mivel kutatásunk eredeti célja nem a nyelvi vizsgákkal kapcsolatos dimenziók feltárása volt, így a faktorelemzésre most csak azért van szükség, hogy lássuk, van-e az állításoknak olyan csoportja, amely megbízhatóan méri a diákok hozzáállását, és amit a további elemzésben felhasználhatunk. Az eredmények valóban azt mutatják, hogy elhamarkodott volt azt gondolnunk, hogy a nyelvi vizsgákkal kapcsolatos beállítódás egyetlen dimenziót alkot.

Mivel a 9 állítás nem alkotott egyetlen dimenziót, így fontos látni, hogyan gondolkodhatnak a nyelvtanulók a nyelvi vizsgákról. A 2. táblázat alapján látszó első faktor a nyelvvizsgák fontos bizonyító erejét mutatják, a második faktor inkább a nyelvi vizsgák motiváló hatását írják le, míg a harmadik faktor azokat a nézeteket tömöríti, amelyek arra utalnak, hogy a nyelvi vizsgák a nyelvtudás pillanatnyi szintjén túl nem

bizonyítanak jövőbeli nyelvtudást. Nyilván további vizsgálatoknak kell feltárnia a pontos gondolkodási sémákat, amelyek a faktorok mögött meghúzódnak, ezen a ponton mi csak a három faktort alkotó állítások megbízhatóságára voltunk kíváncsiak, amit Cronbach-alfával mértünk. Az eredmények: faktor 1 = 0,72; faktor 2 = 0,55; faktor 3 = 0,59. Így a további elemzésbe csak az első faktort vontuk be, amelynek a nyelvi vizsgák elnevezést adtuk.

	Faktorok		
	1	2	3
A nyelvvizsga-bizonyítvány nagyon fontos bizonyítéka a nyelvtudásnak.	0,729		
Azért tanulok angolul, hogy minél magasabb szintű nyelvvizsgát tudjak tenni.	0,478		
Az tud valóban angolul, akinek nyelvvizsgája van.	0,519		
Az angol érettségi nagyon fontos bizonyítéka a nyelvtudásnak.	0,732		
Sokszor csak azért tanulok angolt, hogy készüljek a nyelvvizsgára.		0,372	
Azért tanulok angolul, hogy minél jobban sikerüljön az érettségim.		0,982	
Ha le is teszem az előrehozott érettségit angolból, azért még szeretnék utána angolórára járni.			0,519
A nyelvvizsga megszerzése nem jelenti azt, hogy abba lehet hagyni az angoltanulást.			0,884
Hiába van nyelvvizsgája az embernek, ha nem használja a nyelvet, akkor elfelejtje azt.			0,373

2. táblázat. A nyelvi vizsgákkal kapcsolatos faktorelemzés eredménye (maximum likelihood, varimax forgatás; 0,35-nél nagyobb értékek)

A skálák leíró elemzése

Miután létrehoztuk a nyelvi vizsgákat mérő skálát, bemutatjuk az elemzésbe bevont többi skálát is. A leíró statisztikai eredményeket a 3. táblázat foglalja össze. A táblázatból látszik, hogy a legalacsonyabb átlagot a nyelvi vizsgákkal kapcsolatos skála érte el, átlaga a direkt kontaktust mérő skálán kívül ($t = -1,976$; $p = 0,054$) mindegyik skálánál szignifikánsan alacsonyabb. A táblázatból az is kiténik, hogy a legmagasabb átlaggal a kulturális kontaktus és a diákok ideális második nyelvi énjét mérő skálák rendelkeznek, azaz a belső motivációs tényezők. A diákok nyelvtanulási tapasztalatai nem túl pozitívak, és az autonómia területén inkább a technológiai autonómiára kerül a hangsúly, mint a hagyományosabbnak mondható tanulói autonómiára.

	Cr. alpha	Átlag	Szórás
Nyelvi vizsgák	0,72	2,74	0,81
Direkt kontaktus	0,79	3,01	0,81
Nyelvtanulási tapasztalatok	0,61	3,07	0,91
Autonóm tanulási viselkedés	0,81	3,10	0,64
Technológiai alapú önálló tanulás	0,82	3,84	0,71
Kontaktus fontossága	0,83	3,85	0,80
Motivált tanulási viselkedés	0,83	3,87	0,74
Szükséges második nyelvi én	0,64	4,07	0,62
Ideális második nyelvi én	0,62	4,60	0,54
Kulturális kontaktus	0,66	4,63	0,65

3. táblázat. Az elemzésbe bevont skálák leíró statisztikája

Annak ellenére, hogy a nyelvi vizsgákkal kapcsolatos skála alacsony átlagot mutat, lehet, hogy szorosan összefügg más skálákkal, amiből arra lehetne következtetni, hogy közvetett módon kapcsolódhat a motivációhoz. Korrelációs elemzésünk viszont azt mutatja, hogy a nyelvi vizsgákkal kapcsolatos állításokat tömörítő skála csupán egyetlen másik skálával mutat szignifikáns korrelációt, a technológiai autonómiával ($r = 0,283$; $p = 0,04$). Bár az eredmény szignifikáns, a korreláció nem mondható erősnek vagy jelentősnek. Ezen eredmények alapján viszont nem mondhatjuk, hogy a nyelvi vizsgákkal kapcsolatos beállítódás szorosan kapcsolódik a jelen kutatásban mért motivációs tényezőkhöz. Ha viszont korrelációs összefüggést nem találunk, akkor megkérdőjeleződik a direkt, ok-okozati hatás is. Így levonhatjuk a következtetést, hogy a jelen mintánkban nem találtuk a nyelvi vizsgákkal kapcsolatos állítások motiváló erejét. Természetesen egyetlen vizsgálat alapján nem állíthatjuk, hogy ez a kapcsolat nem létezik, más mintán, más motivációs tényezőkkel és a nyelvi vizsgákkal kapcsolatos vélemények más mérésével eltérő eredményeket kaphatunk.

Összefoglalás

Legfontosabb eredményünk felveti a nyelvi vizsgákkal kapcsolatos beállítódások motiváló erejét a vizsgált mintán, azonban ezt az eredményt magától értetődően csak hasonló budapesti iskolákra lehet általánosítani, és ebből nem lehet országos szinten következtetéseket levonni. Mindazonáltal fontos lehetőségként kell számolni azzal, hogy bár a nyelvi vizsgáknak egyre fontosabb szerepe van a közoktatásban, azok motiváló ereje nem biztos, hogy hosszú távon is megfelelő mértékben alakítja majd a mai középiskolások jövőbeli nyelvtanulását. Ezért nyelvpedagógiai és nyelvpolitikai szempontból fontos lenne, hogy a nyelvi vizsgák rövid távú motiváló erejére a

mindennapi gyakorlatban építsen az iskolai nyelvoktatás, azaz éljenek a nyelvtanárok azzal a lehetőséggel, hogy a felvételi követelmények érdekében a diákok esetleg kevesebb bátorítással is készek a nyelvtudásukat fejleszteni, a nyelvórai munkát komolyabban venni. Fontos azonban elkerülni az ezzel járó veszélyeket, ilyenek például a vizsgára történő felkészülés előtérbe helyezése, az ilyen típusú feladatok dominanciája a nyelvórákon és ezzel együtt, az iskolán kívüli, valós nyelvhasználat és egyéb nyelvtanulási célok háttérbe szorítása, akár teljes figyelmen kívül hagyása. Mindezt tudatosítani kellene mind a gyakorló tanároknál, mind a tanárképzésben tanuló hallgatóknál. Jelentős szempont ez az iskolán túlmutatóan az oktatáspolitikai szintjén is, hiszen az eredmények intelmek lehetnek arra nézve, hogy a szinte kizárólag a nyelvi vizsgák meglétében, számában mért siker nem fog használható, fenntartható nyelvtudást és önálló nyelvtanulókat, nyelvhasználókat eredményezni. A nyelvi vizsgák szerepét a nyelvtanulási motivációban pontosan körül kell írni, a hozadékaikra és az esetleges negatív hatásaikra figyelemmel érdemes szerepüket a köznevelésben folyó nyelvtanításban meghatározni. A tényleges és hatékony idegennyelv-használat elősegítésére a nyelvi vizsgákon túlmutató intézkedésekre is szükség van.

Mivel eredményeink egyetlen középiskola diákjaival történt adatfelvétel alapján készültek, azt gondoljuk, hogy mindenképpen további vizsgálatoknak kellene a felvázolt problémát körüljárni, pontosabb adatokat gyűjteni nemcsak a nyelvi vizsgákkal kapcsolatos beállítódásokról, hanem olyan háttértényezőkről is, amelyek befolyásolhatják a fent leírt kapcsolatrendszer. Fontos lenne különbséget tenni hosszú és rövid távú célok és motiváció között, mert ez is különbözőképpen alakíthatja a nyelvi vizsgákkal kapcsolatos nézeteket. Kvalitatív kutatások segíthetnek feltárni azokat a gondolkodásmódokat, amelyek közelebb vihetik a tanulókat és tanárokat ahhoz, hogy hosszú távon is ki tudják használni a nyelvi vizsgák motiváló erejét.

Megjegyzés

A tanulmány megírását az NKFI-6-K-129149-es számú kutatás támogatta.

IRODALOM

- Albert Ágnes – Piniel Katalin – Lajtai Ádám (2019): Középiskolás diákok angolórákkal kapcsolatos érzelmei egy kérdőíves felmérés tükrében. *Modern Nyelvoktatás*, 25/2, 27–41.
- Albert Ágnes – Piniel Katalin – Lajtai Ádám (2018): A tanulók válasza a 11. évfolyamon. Negyedik fejezet. In: Öveges Enikő – Csizér Kata (szerk.) *Az idegennyelv-oktatás keretei és hatékonysága a köznevelésben. Kutatási jelentés*. Budapest: Emberi Erőforrások Minisztériuma és Oktatási Hivatal, https://www.oktatas.hu/pub_bin/dload/sajtoszoba/nyelvoktatasi_kutatasi_jelentes_2018.pdf.

- Alderson, J. C. – Nagy, E. – Öveges, E. (2000, eds.): *English language education in Hungary, Part II. Examining Hungarian learners' achievements in English*. Budapest: The British Council.
- Alderson, Charles J. – Nagy Edit – Öveges Enikő (2001): Az angol nyelvi érettségi vizsgareform. *Iskolakultúra*, 11/9, 93–87.
- Bárdos, J. (2015): Quality assurance in accrediting foreign language examinations. In: Holló, D. – Károly, K. (EDS.) *Inspirations in foreign language teaching: Studies in applied linguistics, language pedagogy and language teaching*. Harlow: Pearson Education, 207–219.
- Boo, Z. – Dörnyei, Z. – Ryan, S. (2015): L2 motivation research 2005–2014: Understanding a publication surge and a changing landscape. *System* 55, 147–157.
- Clément, R. – Dörnyei, Z. – Noels, K. (1994): Motivation, self-confidence and group cohesion in the foreign language classroom. *Language Learning* 44, 417–448.
- Csapó Benő (2001): A nyelvtanulást és a nyelvtudást befolyásoló tényezők. *Iskolakultúra* 11/8, 25–35.
- Csizér, K. (megjelenés alatt): The L2 Motivational Self System. In: Lamb, M. – Csizér, K. – Henry, A. – Ryan, S. (eds.) *Palgrave Macmillan handbook of motivation for language learning*. Basingstoke, UK: Palgrave.
- Csizér Kata – Galántai Dóra (2012): A tanári és szülői szerepek alakulása a középiskolások idegen nyelvi motivációjának alakításában. Egy strukturális modell tanulságai. In: Németh András (szerk.) *A neveléstudományi doktori iskola programjai: Tudományos arculat, kutatási eredmények*. Budapest: ELTE Eötvös Kiadó, 171–178.
- Csizér Kata – Illés Éva (2018): A nyelvtanárok válaszai. In: Öveges Enikő – Csizér Kata (szerk.) *Vizsgálat a köznevelésben folyó idegennyelv-oktatás kereteiről és hatékonyságáról: kutatási jelentés*. Budapest, Magyarország: OH-EMMI, 29–52.
- Csizér, K. – Kormos, J. (2008): An overview of Hungarian secondary school students' foreign language motivation. In: Knudsen, H. V. (ed.) *Secondary school education*. Hauppauge, NY: Nova Science Publishers, 65–87.
- Csizér, K. – Kormos, J. (2009): Modelling the role of inter-cultural contact in the motivation of learning English as a foreign language. *Applied Linguistics*, 30/2, 166–185.
- Csizér Kata – Kormos Judit (2012): A nyelvtanulási autonómia, az önszabályozó stratégiák és a motiváció kapcsolatának vizsgálata. *Magyar Pedagógia*, 112/1, 3–17.
- Csizér, K. – Kormos, J. (2014): The ideal L2 self, self-regulatory strategies and autonomous learning: A comparison of different groups of English learners. In: Csizér, K. – Magid, M. *The impact of self-concept on language learning*. Bristol, UK: Multilingual Matters. 73–86.
- Csizér, K. – Lukács, G. (2010): The comparative analysis of motivation, attitudes and selves: The case of English and German in Hungary. *System*, 38/1, 1–13.
- D. Molnár Éva (2014): Az önszabályozott tanulás pedagógiai jelentősége. In: Buda András – Golnhofer Erzsébet (szerk.) *Tanulmányok a neveléstudomány köréből, 2013. Tanulás és környezete*. Budapest, Magyarország: MTA Pedagógiai Tudományos Bizottság, 29–54.
- Dörnyei, Z. (2005): *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Dörnyei, Z. (2009): The L2 motivational self system. In: Dörnyei Z. – Ushioda E. (eds.) 2009. *Motivation, language identity and the L2 self*. Bristol: Multilingual Matters. 9–42.
- Einhorn Ágnes (2007): Az idegen nyelvi érettségi vizsga reformja. In Vágó Irén (szerk.) *Fókuszban a nyelvtanulás*. Budapest: Oktatókutatató és Fejlesztő Intézet, 73–105.

- Einhorn Ágnes (2008): Hogy vizsgázott az új érettségi németből? In: Bánkuti Zsuzsanna – Lukács Judot (szerk.) *Tanulmányok az érettségiről. Hatásvizsgálat, tantárgyi vizsgák értékelése, feladatfejlesztés*. Budapest: Oktatókutatató és Fejlesztő Intézet, 207–237.
- Európa Tanács (2002): *Közös Európai Referenciakeret: Nyelvtanulás, nyelvtanítás, értékelés*. Hozzáférés https://nyak.oh.gov.hu/nyat/doc/ker_2002.asp
- Európai Bizottság (2015): *Languages in secondary education: An overview of national tests in Europe – 2014/2015*. https://eacea.ec.europa.eu/national-policies/eurydice/content/languages-secondary-education-overview-national-tests-europe—201415_en
- Fekete Adrienn – Csépes Ildikó (2019): B2-es szintű nyelvvizsga bizonyítvány: útlevel a diplomás élethez, társadalmi mobilitáshoz. *Iskolakultúra*, 28/10–11, 13–24.
- Halápi Magdolna – Hegedűs Kristóf (2014): *Érettségi vizsgatárgyak elemzése. 2009–2012 évi tavaszi vizsgaidőszakok. Angol nyelv*. https://www.oktatas.hu/pub_bin/dload/unios_projektek/tamop318/erettségi_vizsgatargyak_elemzese/angol.pdf
- Jilly Viktor – Nagy Ildikó (2010): *Összehasonlító elemzés angol nyelv vizsgatantárgyból*. https://www.oktatas.hu/pub_bin/dload/unios_projektek/tamop318/ertekelesi_skalak_osszehasonlitasa/ertekelesi_skalak_angol nyelv.pdf
- Kákonyi Lucia (2016): Idegen nyelvi mérések tapasztalatai. *Új Köznevelés*, 2016/5–6. <http://fo-lyoiratok.ofi.hu/uj-kozneveles/idegen-nyelvi-meresek-tapasztalatai>
- Kormos, J. – Csizér, K. (2007): An interview study of inter-ethnic contact and its role in language learning in a foreign language environment. *System*, 35/2, 241–258.
- Kormos, J. – Csizér, K. (2008) Age-related differences in the motivation of learning English as a foreign language: attitudes, selves and motivated learning behaviour. *Language Learning*, 58/2, 327–355.
- Kormos, J. – Csizér, K. (2014): The interaction of motivation, self-regulation and autonomous learner behavior in different learner groups. *TESOL Quarterly*, 48/2, 275–299.
- Little, D. (1999): Learner autonomy is more than a Western cultural construct. In: Cotterall S. – Crabbe D. (eds.) *Learner autonomy in language learning: Defining the field and effecting change*. Frankfurt am Main: Peter Lang, 11–18.
- Nikolov Marianne (2003a): Angolul és németül tanuló diákok nyelvtanulási attitűdje és motivációja. *Iskolakultúra*, 13/8, 61–73.
- Nikolov Marianne (2003b): Az idegennyelv-tanítás megújulásának hatásai. *Új Pedagógiai Szemle*, 53/3. 46–57.
- Nikolov Marianne – Ottó István – Öveges Enikő (2009): *Az idegennyelv-tanítás és tanulás helyzete és fejlesztési lehetőségei a középfokú szakképző intézményekben*. Budapest: Oktatásért Közalapítvány.
- Nikolov, M. – Szabó, G. (2015): A study on Hungarian 6th and 8th graders' proficiency in English and German at dual-language schools. In: Holló, D. – Károly, K. (eds.) *Inspirations in foreign language teaching: Studies in applied linguistics, language pedagogy and language teaching*. Harlow: Pearson Education, 184–206.
- Oktatási Hivatal (2018a): *Eredmények összesítése. Idegen nyelvi mérés, 2018*. https://www.oktatas.hu/pub_bin/dload/kozoktat/meresek/idegen_nyelvi_meres/Idegennyelvimeres2018_Eredmenyekosszesitese.pdf

- Oktatási Hivatal (2018b): *Eredmények összesítése. Célnyelvi mérés, 2018.*
https://www.oktatas.hu/pub_bin/dload/kozoktatas/meresek/celnyelvi_meresek/Celnyelvi2018_Eredmenyekosszesitese.pdf
- Oktatási Hivatal (2018c): *Érettségi 2018. május-június.* https://www.oktatas.hu/kozneveles/erettsegi/prezentaciok_tanulmanyok
- Öveges Enikő (2018): A magyarországi gyermeknyelvi vizsgák tartalmi és jogi szabályozottsága. *Gyermeknevelés*, 6/3, 135–147.
- Öveges Enikő – Csizér Kata (2018): *Vizsgálat a köznevelésben folyó idegennyelv-oktatás kereteiről és hatékonyságáról: kutatási jelentés.* Budapest, Magyarország: OH-EMMI.
- Öveges Enikő – Csizér Kata – Dér Kinga (2019): Az idegennyelv-tanítás iskolai szintű szabályozása: keretek, célok és megközelítések. *Modern Nyelvoktatás*, 25/1, 3–15.
- Sominé Hrebik Olga – Vígh Tibor (2010): *Összehasonlító elemzés német nyelv vizsgatantárgyból.* http://www.oktatas.hu/pub_bin/dload/unios_projektek/tamop318/ertekelesi_skalak_osszehasonlitasa/ertekelesi_skalak_nemet_nyelv.pdf
- Tartsayné Németh Nóra – Tiboldi Tímea – Katona László (2018): Az intézményvezetők válasza. Második fejezet. In: Öveges Enikő – Csizér Kata (szerk.) *Vizsgálat a köznevelésben folyó idegennyelv-oktatás kereteiről és hatékonyságáról: kutatási jelentés.* Budapest, Magyarország: OH-EMMI, 29–52.
- Vígh Tibor (2012): A kétszintű idegen nyelvi érettségi tanítási-tanulási folyamatra gyakorolt hatása. Egy tanári kérdőíves felmérés eredményei. *Iskolakultúra*, 22, 62–79.
- Vígh Tibor (2013): A középszintű angol és német nyelvi érettségiben elért teljesítmények változásai 2007 és 2012 között. *Modern Nyelvoktatás*, 19/1–2, 17–35.

SZABÓ GÁBOR

A szövegnehézség vizsgálata az emelt szintű angol érettségi olvasáskomponensében

Examining text difficulty in the reading component of the advanced level school-leaving exam in English

This paper presents a comparative study targeting the difficulty of texts used in the reading component of the advanced level school-leaving exam in English. The paper introduces the context of the school-leaving exam in Hungary, which is followed by a description of the concept of text difficulty along with a discussion of attempts to measure it. A modern, multi-faceted, computer-based analysis of text difficulty, the Coh-Metrix platform is introduced next, along with a description of the Coh-Metrix Common Core Text Ease and Readability Assessor (TERA) and its components describing five text characteristics: narrativity, syntactic simplicity, word concreteness, referential cohesion and deep cohesion. A specific L2 readability measure of Coh-Metrix is also discussed. These six measures are then utilized to analyze altogether 33 texts used in a total of eight versions of the reading component of the advanced level school-leaving exam in English in order to determine whether significant differences exist across the clusters of texts used in different exam versions. The results reveal that no significant differences were found in terms of any of the measures, despite the fact that individual texts appeared to differ from one another.

Keywords: school-leaving exam, reading comprehension, language testing, text difficulty, Coh-Metrix

Bevezetés

A nyelvtudás mérése kapcsán alapvető fontosságú kérdés a mérés stabilitásának garantálása, azaz annak biztosítása, hogy a különböző vizsgaalkalmak során az aktuális vizsgázók ugyanolyan szintű kihívással kerüljenek szembe. Praktikusan ez azt jelenti, hogy az adott vizsga szintjének minden alkalommal azonosnak kell lennie nem pusztán a vizsgaszpecifikáció kvalitatív szempontjainak megfelelően, hanem a vizsga empirikusan megállapítható nehézsége szempontjából is. Minél nagyobb tétellel bíró vizsgáról van szó, annál nagyobb jelentősége van mindennek, hiszen adott esetben a vizsga sikerességének vagy sikertelenségének a vizsgázók életében komoly következményei lehetnek.

Hazánkban a nyelvtudásmérés nagy tétellel bíró közege egyrészt az akkreditált nyelvvizgák kontextusa, másrészt az akkreditált nyelvvizsgarendszerek körén kívül eső,

ám – megfelelő eredmény esetén – azokkal azonos értékű vizsgabizonyítványt eredményező emelt szintű idegen nyelvi érettségi vizsga [100/1997. (VI. 13.) Korm. rendelet]. Mindkét közegben a legnagyobb az angol nyelvből tett vizsgák száma. A Nyelvvizsgáztatási Akkreditációs Központ adatai szerint 2018-ban összesen 70 971-en tettek középfokú, azaz B2-es szintű angol nyelvvizsgát (https://nyak.oh.gov.hu/doc/statisztika.asp?strId=_43ny), míg ugyanebben az évben az Oktatási Hivatal adatai szerint 12 140 emelt szintű angol érettségi vizsgára került sor (<https://www.ketszintu.hu/publicstat.php>). Mindezek alapján relevánsnak tűnik, hogy az igen jelentős vizsgázói létszámmal bíró emelt szintű angol érettségi működésének bizonyos részleteit tegyük vizsgálatunk tárgyává.

A vizsga egészének működése igen sok elemből áll, s ezek áttekintése természetesen meghaladja a jelen tanulmány kereteit. Irányítsuk tehát figyelmünket a fentebb már jelzett stabilitásra, ezen belül is a vizsga olvasásértést mérő komponensének stabilitására. Az olvasás a nyelvi készségeken belül bizonyos értelemben kiemelt készségterületnek tekinthető, hiszen fejlesztése által nem csupán az olvasás maga, hanem más készségek fejlődése is elérhető (Krashen 2004).

A receptív készségek mérése esetén a vizsgaszint lényeges összetevője a méréshez használt szövegek nehézségi szintje. Éppen ezért nagy fontossággal bír a vizsgaverziók stabilitása szempontjából a különböző verziókban használt szövegek nehézségének nyomon követése. A jelen kutatás célja azon szövegek nehézségének összehasonlító vizsgálata, amelyek az emelt szintű angol érettségi vizsga nyolc egymást követő vizsgaidőszakában az olvasásértést mérő komponensben kerültek felhasználásra.

Az emelt szintű idegen nyelvi érettségi

Mint azt fentebb láthattuk tehát, az emelt szintű idegen nyelvi érettségi jelentős vizsgázói létszámmal rendelkező, nagy tétellel bíró vizsga. A következőkben röviden áttekintjük a vizsga legfontosabb jellemzőit.

A jelenleg is hatályos kétszintű érettségi rendszere hosszas fejlesztőmunkát követően 2005-ben került bevezetésre, s az idegen nyelvi érettségi korszerű mérési elveket követve autentikus szövegekre épül, és számos feladattípust alkalmaz (Einhorn–Major 2006). Az emelt szintű vizsga írásbeli része méri az olvasás- és hallásértést, valamint az íráskészséget és a nyelvhelyességet, és szintje a Közös európai referenciakeret (KER) B2-es szintjét célozza meg. A jelen vizsgálat szempontjából kiemelten fontos angol olvasásértési komponensben négy feladat szerepel (Major 2006).

Az idegen nyelvi érettségi, és ezen belül az emelt szintű érettségi jelentősége nyelvpedagógia szempontból is kétségtelen, s éppen ezért indokoltnak tűnne a vizsga minőségbiztosításának fokozott garantálása. A vizsga működéséről, illetve ennek objektíven mérhető mutatóiról azonban meglepően kevés tanulmány jelent meg. Bár kétségkívül léteznek olyan publikációk, amelyek a vizsga egyes verzióinak, illetve bizonyos

aspektusainak működését vizsgálták (pl. Halápi–Hegedűs 2013, Jilly–Nagy–Enyedi–Wintsche–Morvay 2014, Vigh 2008, 2013), kevés az elérhető információ arról, hogy a vizsga folyamatában, vizsgaidőszakról vizsgaidőszakra mennyiben felel meg a minőségi elvárásoknak, illetve hogy stabilan azonos-e a vizsga nehézsége. A jelen tanulmány az emelt szintű érettségi működéséről alkotott kép teljesebbé tételéhez kíván hozzájárulni az olvasásértési komponensben használt szövegek vizsgálatával.

A szövegnehézség mérése

A KER (PTMIK 2002) széles körű elterjedése kapcsán világszerte megnőtt a nyelvtudás tartalma és szintjei iránti érdeklődés (lásd pl. Kecker–Eckes, 2010). Természetesen minderről a KER megjelenése előtt is sok szó esett (pl. Bachman–Lynch–Mason 1995), ám a KER addig példátlanul részletes szintleírásai megnövelték az ez irányú érdeklődést. Nem volt ez másként az olvasásértés kapcsán sem.

Nem szabad elfeledkeznünk azonban arról, hogy egy olvasásértést mérő feladat kapcsán a feladat nehézsége több tényező együttes hatásaként áll elő, s leggyakrabban a szöveg, a feladat és az olvasó interakciójának eredőjeként határozzák meg (pl. Castello 2008). Maga az olvasás folyamata szempontjából a szövegnek nincs is feltétlenül egyértelmű jelentése, csupán jelentéspotenciálja (Halliday 1978), amelyet a különböző olvasók különböző módon realizálnak, sőt olyan vélemény is akad, mely szerint minden egyes olvasó egyéni és egyedi módon értelmez egy adott szöveget (Alderson 2000).

Mindebből akár arra is következtethetünk, hogy egy adott szöveg esetében nem is lehet annak szintjéről beszélni, mivel a szöveg értelmezése teljesen az adott olvasótól függ. Ha azonban ettől a talán kissé szélsőséges nézettől elvonatkoztatunk is, könnyen belátható, hogy egy szöveget több szinten lehet érteni, értelmezni. Nem mindegy pl., hogy a szöveg globális megértése a cél, vagy részletes szövegértést várunk el. Ez utóbbi megkülönböztetés összhangban áll a KER által képviselt megközelítéssel, mely szerint az egyes deskriptorok szintjén a szövegértés kapcsán tételesen különbséget kell tenni a lényeg, illetve a részletek megértése között (vö. PTMIK 2002).

Egy szöveg nehézségének meghatározása tehát nem egyszerű feladat, ám ennek ellenére számos kísérlet történt már a szövegnehézség valamely objektív paraméterek tükrében történő meghatározására. A nemzetközi szakirodalomban talán a legismertebb szövegnehézségi indexek a *Flesch* olvashatósági index, illetve a *Flesch–Kincaid*-index, amelyek egy a szavak száma, a mondatok száma, illetve a szótagszám közötti feltételezett összefüggésre épülnek (Klare 1974–1975).

Ezek az indexek azonban problematikusnak mutatkoztak, különösen idegen nyelvi szövegértelmezési kontextusokban, mivel igen gyenge összefüggést mutattak tényleges szövegértési vizsgateljesítményekkel (Brown 1998), és további gyengéjük az a fajta megközelítés, amely a szövegnehézséget meghatározó komplex tényezőket egyetlen indexszé egyszerűsíti (Alderson 2000).

Mindezek okán a közelmúltban olyan modelleket is kifejlesztettek, amelyek a szövegnehézséget lényegesen komplexebb módon közelítik meg, s így valósabb képet adnak a szövegek nehézségéről. Ezen próbálkozások közül kiemelkedik a *Coh-Matrix* olvashatósági formula (Graesser–McNamara–Kulikowich 2011). A Coh-Matrix összesen 53 különféle index segítségével írja le egy-egy szöveg jellemzőit, s ezáltal igen részletes képet alkothatunk a szövegről. Ilyen nagyszámú jellemző alkalmazása azonban nem volna praktikus, így főkomponens-analízis segítségével az 53 jellemzőt sikerült nyolc fő komponensre redukálni, melyek a következők: narrativitás, referenciális kohézió, szintaktikai egyszerűség, szókonkrétság, oksági kohézió, igekohézió, logikai kohézió és temporális kohézió. A nyolc komponensre sikeresen kapcsolták össze a Graesser és McNamara (2011) által felállított ötszintű elméleti modellel a következőképpen: *műfaj* (narrativitás), *szituációs modell* (okszági kohézió, igekohézió, logikai kohézió és temporális kohézió), *szövegbázis* (referenciális kohézió), *szintaxis* (szintaktikai egyszerűség) és *szavak* (szókonkrétság).

Ez a modell szolgálta alapul a szövegek nehézségét és olvashatóságát mérő internetes platform, a TERA (*Coh-Matrix Common Core Text Ease and Readability Assessor*) megalkotásához (Jackson–Allen–McNamara 2017). A TERA öt dimenzió mentén méri a szövegek nehézségét, az eredményeket pedig százpontos skálán jeleníti meg. A *narrativitás* a szöveg történetességét vizsgálja a pusztán információközlő szövegekkel szemben. A narratív szövegekben nagyobb arányban szerepelnek gyakori szavak, illetve könnyen érhető igék, valamint olyan névmások, amelyek közelebb hozzák a szöveget az olvasóhoz. A *szintaktikai egyszerűség* a szövegben szereplő mondatok komplexitásán alapul, melyhez a modell számításba veszi többek között a tagmondatok számát, illetve a mondatokban szereplő szavak számát, valamint a főmondatot megelőző szavak számát. A különböző bekezdésekben szereplő mondat szerkesztési hasonlóságok vizsgálata ugyancsak részét képezi az elemzésnek. A *szókonkrétség* a szövegben szereplő absztrakt és konkrét szavak arányára épül. Az absztrakt szavak általában nehezebbek, míg a konkrét szavak könnyítik a szövegértést. A *referenciális kohézió* a szavak, szótövek, illetve fogalmak mondatok közti átfedésének vizsgálatára épül. Ha nagy arányban vannak jelen ilyen átfedések, az könnyíti a szöveg megértését. A *mélykohézió* elsősorban a kötőszókat vizsgálja, mivel így állapítható meg, hogy a szövegben szereplő események és információk mennyire kapcsolódnak egymáshoz. A kötőszók nagy száma erősebb kapcsolódást jelez, s ez könnyíti a szövegértést (Jackson et al. 2017).

A Coh-Matrix eredmények alapján egy speciális *Idegen nyelvi olvashatósági* indexet is kifejlesztettek, amely a szavak gyakoriságán, a szintaktikai hasonlóságokon, illetve a tartalmat kifejező szavak közötti átfedéseken alapul (Crossley–Greenfield–McNamara 2008). Mint bebizonyosodott, ez az index a hagyományos mutatóknál (pl. Flesch–Kincaid-index) lényegesen hatékonyabban alkalmazható az olvashatóság

mérésére, mivel hatékonyabban reprezentálja az olvasás során aktivizált kognitív műveleteket (Crossley–Allen–McNamara 2011).

A kutatás során a Coh-Metrix platform interneten elérhető verziója (McNamara–Louwerse–Cai–Graesser 2013) került alkalmazásra.

A kutatás célja és menete

Amint fentebb jeleztük, a jelen kutatás fókuszában az angol nyelvi emeltszintű érettségi olvasásértést mérő feladatai állnak, még pontosabban az ezen feladatokban felhasznált szövegek nehézsége. A vizsgálat összesen nyolc vizsgaidőszak (2015. május–2018. október) mindösszesen 33 szövegére terjedt ki, és azt vizsgálta, vajon mennyiben mutatható ki különbség a szövegek nehézségében a Coh-Metrix TERA szövegjellemzőket, illetve az *Idegen nyelvi olvashatóságot* leíró mutatók tükrében.

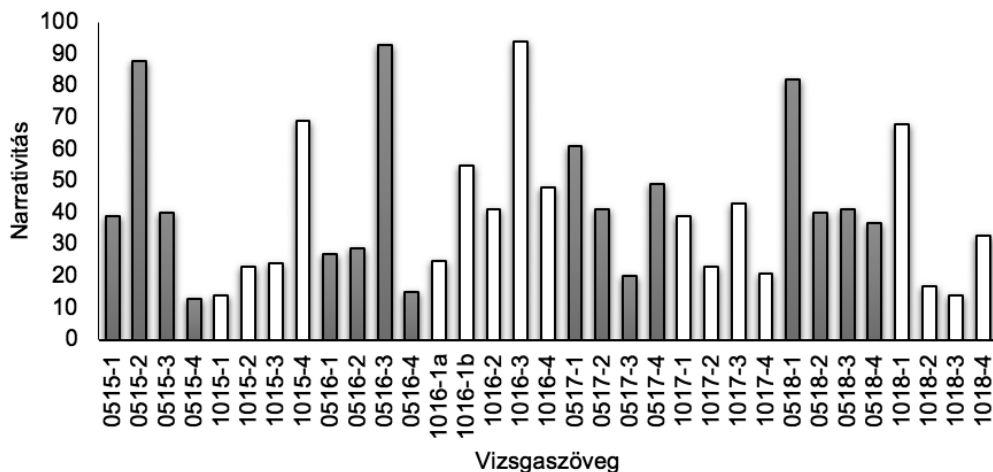
Az elemzés menete a következő volt. Először a feladatokban használt szövegeket elemezhető formátumúvá kellett alakítani. Erre azért volt szükség, mert az érettségi különböző feladattípusai között olyanok is szerepelnek, amelyek esetében a szöveg egyes részei hiányoznak, és a vizsgázók feladata a megfelelő szövegrészek megfelelő helyre történő beillesztése. Mivel ezek a szövegek a feladatban szereplő formátumukban nem alkotnak koherens egészet, értelemszerűen először – a megoldókulcsok segítségével – „érintetlen” formátumúvá kellett őket alakítani. Természetesen az olyan feladatok esetében (pl. feleletválasztás), ahol a szövegek komplett egészként jelennek meg, a feladatokban szereplő szövegformátum közvetlenül alkalmazható volt.

Második lépésként került sor a szövegek Coh-Metrix TERA elemzésére. Ennek során, a rendelkezésre álló online felületen (<http://cohmetrix.com>) minden szöveg esetében meghatároztuk és rögzítettük az öt fő szövegjellemző értékeit, valamint hatodik szempontként a Coh-Metrix *Idegen nyelvi olvashatóság* indexét is. Ezen a ponton valamennyi szövegjellemző mentén összehasonlíthatóvá váltak a szövegek, ám annak megállapításához, hogy léteznek-e a felszíni látszaton túl tényleges, statisztikailag szignifikáns különbségek a szövegjellemzőkben, megfelelő statisztikai próba alkalmazására volt szükség. Tekintettel az adatok jellemzőire, vagyis arra, hogy az adatok nem feltétlenül normáleloszlás formájában jelennek meg, nemparametrikus statisztikai eljárást, a Kruskal–Wallis-próbát alkalmaztuk, mely az egyszempontos varianciaanalízis nemparametrikus alternatívája.

Az elemzés segítségével ezek után megállapíthatóvá vált, hogy vannak-e statisztikailag szignifikáns különbségek az eltérő vizsgaidőszakokban használt szövegek jellemzői között, illetve hogy a szövegjellemzők összesítése, mely a szövegek egészét kompozit pontszámként hivatott leírni, mutat-e szignifikáns különbséget az egyes vizsgákhoz kapcsolódó szövegcsoportok között.

Eredmények

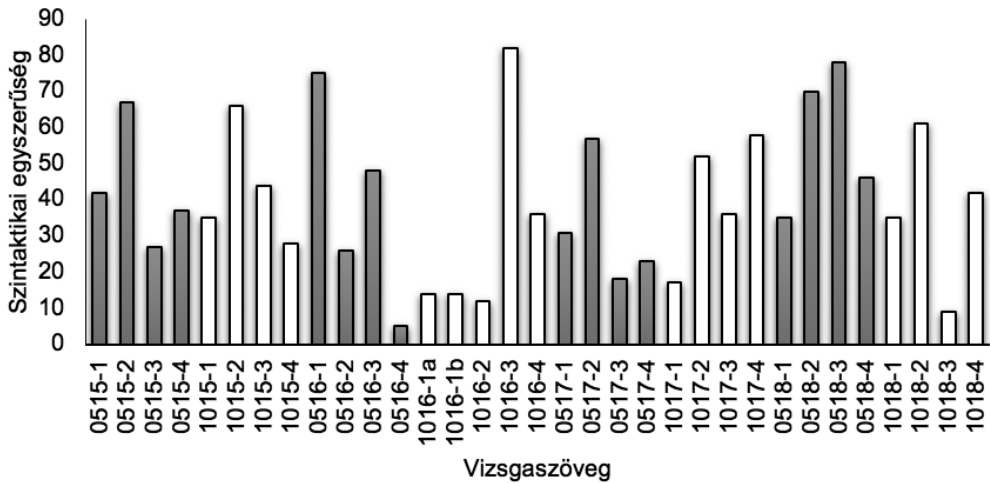
A következőkben áttekintjük a vizsgálat eredményeit. Elsőként vegyük szemügyre a *narrativitás* kapcsán született Coh-Matrix TERA eredményeket, melyeket az 1. ábra mutat be.



1. ábra. Coh-Matrix TERA eredmények: narrativitás

A függőleges tengelyen a *narrativitás* szövegjellemzőhöz kapcsolódó érték látható, míg a vízszintes tengelyen az adott vizsgaszöveg kódja szerepel (hónap, év, szövegszám; pl.: 0515-1: május, 2015., első szöveg). Az oszlopok színezése a vizsgaidőszakok könnyebb elkülöníthetőségét szolgálja. Mint az ábrán látható, a narrativitás szempontjából a szövegek meglehetősen heterogénnek tűnnek. Érdekes azonban megfigyelni, hogy csaknem valamennyi vizsgaidőszakban szerepel magas és alacsony narrativitású szöveg is. Vagyis, bár a tesztekben vizsgaidőszakról vizsgaidőszakra narrativitásukat tekintve meglehetősen eltérőnek tűnő szövegek szerepeltek, a különbségek nem csupán a vizsgaidőszakok között, hanem egy-egy vizsgaidőszakon belül is megfigyelhetők. Fontos ezen a ponton leszögeznünk, hogy ez nem feltétlenül jelent problémát, sőt kifejezetten előnyös lehet. A B2-es szintű olvasáskészség mérésekor ugyanis kifejezetten kívánatos, hogy a szövegek sokfélék legyenek. Ezáltal lehet ugyanis garantálni a konstuktum megfelelő reprezentáltságát, s így módon a mérés tartalmi érvényességét.

Továbbá tekintsük most át a *szintaktikai egyszerűség* vizsgálata kapcsán előállt eredményeket, amelyeket a 2. ábra mutat be.

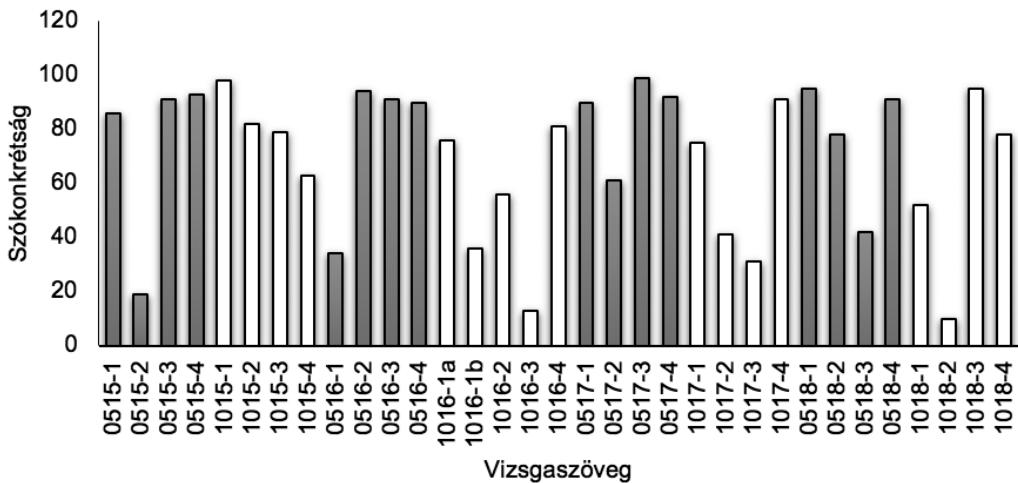


2. ábra. Coh-Metrix TERA eredmények: szintaktikai egyszerűség

Akárcsak az előző szövegjellemző esetében, a szintaktikai egyszerűség kapcsán is azt látjuk, a szövegek a vizsgaidőszakok között és a vizsgaidőszakokon belül is igen eltérőnek mutatkoznak. A 2016. májusi vizsga kapcsán pl. az látható, hogy az első szöveget (0516-1) valamennyi szöveg között a második legmagasabb szintaktikai egyszerűség jellemzi, ugyanakkor a vizsga negyedik szövege (0516-4) valamennyi szöveg között a legalacsonyabb szintű szintaktikai egyszerűséget mutatja. Hasonlóan kontrasztos a 2016. októberi vizsga is, ahol a hármas számú szöveg (1016-3) valamennyi vizsgált szöveg közül a legmagasabb szintaktikai egyszerűséget mutatja, míg az első két feladathoz tartozó három szöveg (az első feladathoz két szöveg tartozik: 1016-1a és 1016-1b) igen alacsony szintaktikai egyszerűséget mutat. Ezen a ponton felmerülhet a kérdés, vajon nem jelent-e problémát az, hogy egy B2-es szintű vizsgában számos olyan szöveg jelenik meg, amely igen magas szintaktikai egyszerűséggel rendelkezik, azaz szintaktikailag igen egyszerű. Érdemes azonban megjegyeznünk, hogy bár a komplex szintaxis nehezíti a megértést, egy szintaktikailag egyszerű szöveg nem automatikusan egyszerűen érthető, ez a szöveg egyéb jellemzőinek is függvénye. Megállapítható tehát, hogy a vizsgált szövegek szintaktikai egyszerűség szempontjából is sokszínűnek bizonyultak, ami ebben az esetben sem jelent feltétlenül problémát.

A harmadik vizsgált szövegjellemző a *szókonkrétság* volt. Az ehhez kapcsolódó eredményeket a 3. ábra szemlélteti.

A korábban, az eddig vizsgált szövegjellemzők kapcsán megfigyelt heterogenitás a szókonkrétság kapcsán is megfigyelhető. Ugyanakkor figyelemre méltó, hogy a vizsgált szövegek többsége magas szókonkrétságot mutat, és két olyan vizsgaidőszak is akad (2015. október, 2017. május), amelyben csak magas szókonkrétségű szövegek



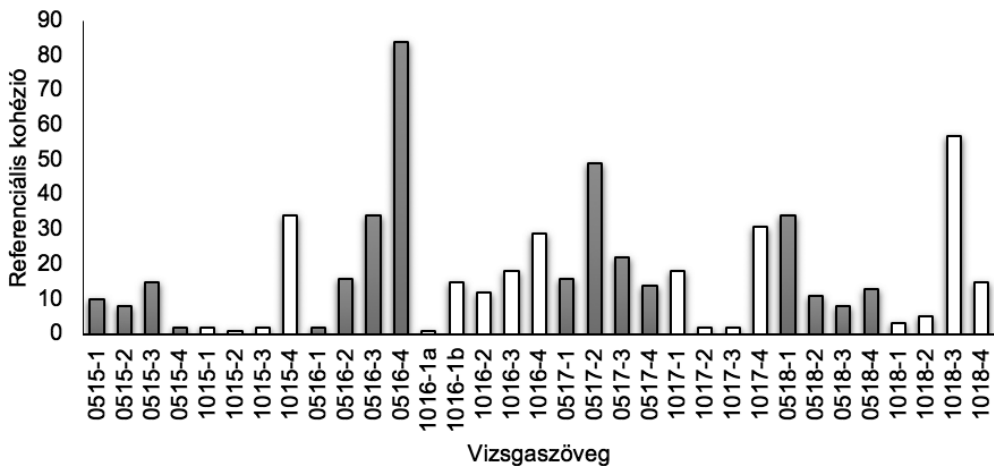
szerepeltek. A szavak konkrétsága általában könnyebbé teszi a szövegek érthetőségét,

3. ábra. Coh-Metrix TERA eredmények: szókonkrétság

ám ebben az esetben is érdemes figyelembe venni, hogy ez pusztán egy szövegjellemző a sok közül. A szókonkrétság szempontjából legmagasabb értéket mutató, tehát elvileg könnyebben érthető, a 2017. májusi vizsgában szereplő harmadik szöveg (0517-3) például mind narrativitását, mind szintaktikai egyszerűségét tekintve alacsony értéket mutat, azaz nehezebben érthető. Vagyis, tekintve, hogy a szövegek tényleges nehézsége a szövegjellemzők összességének függvénye, az egy-egy szövegjellemző esetében előfordul extrém módon magas vagy alacsony értékeket más szövegjellemzők kompenzálhatnak, így biztosítva a szövegek hasonló nehézségét. Hogy ez a vizsgált szövegek esetében így volt-e, azt ezen a ponton még korai eldönteni, de az elsődleges benyomások alapján ez mindenesetre elképzelhető.

A következő vizsgált szövegjellemző a *referenciális kohézió* volt. A 4. ábra az ehhez kapcsolódó eredményeket mutatja be.

Az eddig megvizsgált szövegjellemzőkhöz képest a legfeltűnőbb különbség a szövegek többségének igen alacsony referenciális kohéziója. Mint fentebb már utaltunk rá, az alacsony referenciális kohézió azt jelzi, kicsi az átfedés az egyes mondatokban megjelenő explicit módon kifejezett gondolatok, illetve szavak között. Ebből adódóan az alacsony referenciális kohézió növeli az olvasott szöveg megértésének nehézségét (McNamara–Graesser–Cai–Kulikowich 2011). De vajon indokolt-e az ilyen alacsony referenciális kohézió ezen a szinten? A kérdésre megkísérrelhetünk válaszolni a KER B2-es deskriptorainak tükrében, hiszen a kohézió nyilvánvalóan igen lényeges eleme

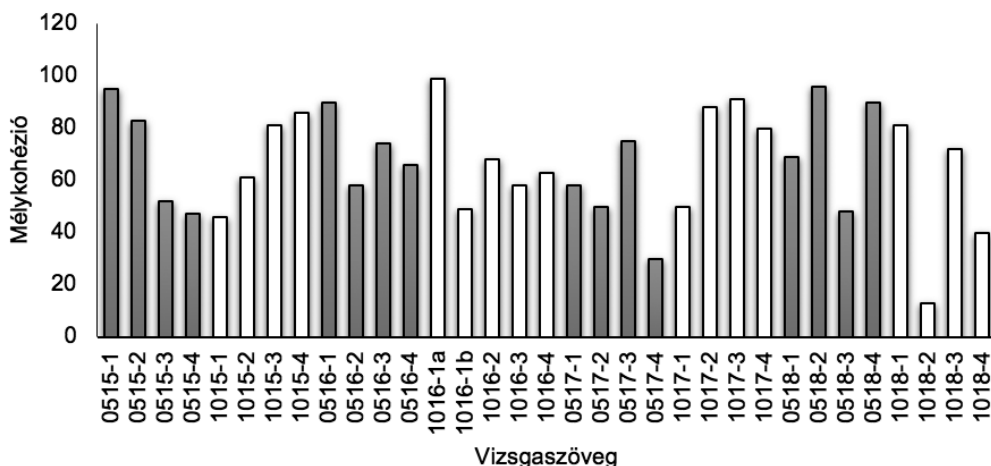


4. ábra. Coh-Metrix TERA eredmények: referenciális kohézió

a szövegek szerkezetének, s így azok megértésének is. Érdekes módon azonban a KER olvasásértéshez kapcsolódó skáláiban a kohézió nem jelenik meg. Bár a KER-ben létezik egy „Kohézió és koherencia” nevű skála, ennek tartalma azonban kizárólag a produktív készségekre vonatkozik. A „Jelek azonosítása és kikövetkeztetés” c. skála B2-es deskriptora szerint azonban a nyelvtanuló ezen a szinten „a megértéshez különböző stratégiákat használ, például szöveghallgatáskor a főbb pontok keresése; a megértés ellenőrzése kontextuális jelek segítségével” (PTMIK, 2002: 89). A „kontextuális jelekre” történő utalás vizsgáldásunk szempontjából releváns lehet, bár nem túlzottan specifikus. Hasonló a helyzet a KER közelmúltban megjelent kísérőkötetében (Council of Europe 2018) szereplő megfogalmazásával. Az „Olvasás információszerzés és érvek keresése céljából” c. skálában szereplő új deskriptor szerint egy B2-es szintű nyelvtanuló „fel tudja ismerni a diszkurzív szövegekben szereplő különböző struktúrákat: az egymással szemben álló érveket, a problémák és megoldások bemutatását és az ok-okozati összefüggéseket” (Council of Europe 2018: 63). A szövegekben szereplő struktúrák felismerése köthető a referenciális kohézió fogalmához, ám semmiképp sem azonos vele. Mint a fentiekből látható, a KER deskriptorai alapján tehát nem igazán indokolható a vizsgált szövegekre általában jellemző alacsony referenciális kohézió, ám, némileg paradox módon, éppen ez az egyik olyan szövegjellemző, amelynek tekintetében a szövegek leginkább hasonlítanak egymásra.

Az ötödik vizsgált szövegjellemző a *mélykohézió* volt. Az ezzel kapcsolatban született eredményeket az 5. ábra szemlélteti.

Mint látható, ezen szövegjellemző esetében is megfigyelhető a heterogenitás, azonban észrevehetően ritka a kifejezetten alacsony szintű mélykohézió. Tekintettel arra, hogy ez a komponens a szövegben megjelenő események, fogalmak és információk

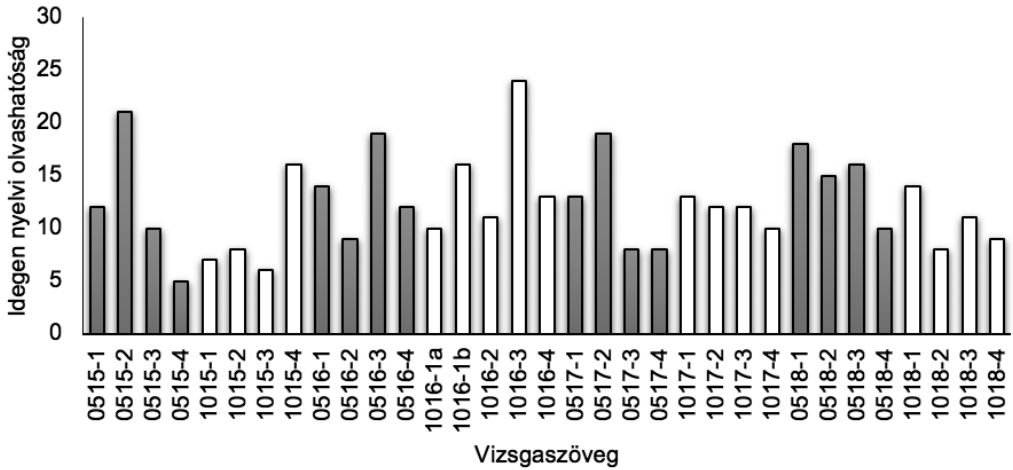


5. ábra. Coh-Matrix TERA eredmények: mélykohézió

összekapcsolódását elemzi, és elsősorban a kötőszókon alapul, az eredményen nem kell csodálkoznunk. Érdekes megfigyelni ugyanakkor, hogy a 2018. októberi vizsga második szövegének (1018-2) mennyire alacsony ez a mutatója, ami azt sejteti, hogy ebben a szövegben sokkal kevésbé kapcsolódnak össze az egyes események és fogalmak. Ez a szöveg más szövegjellemzőket tekintve is alacsony indexekkel rendelkezik, kivéve a szintaktikai egyszerűséget. Tekintve, hogy ezt a szöveget egy olyan feladathoz használták, amelyben a vizsgázóknak a szövegből törölt részeket kellett a szövegbe illeszteni, elképzelhető, hogy a szintaktikai szerkezet nagyobb befolyást gyakorolhatott a megoldások helyességére. Mindemellett tény, hogy ez a szöveg határozottan különbözni látszik a többitől.

Az utolsó vizsgált szövegjellemző az *idegen nyelvi olvashatóság* volt, mely ugyan nem része a TERA-elemzésnek, vizsgálódásunk szempontjából azonban kiemelt jelentőséggel bír, tekintve hogy az érintett szövegek idegen nyelvi szövegértés méréséhez kerültek felhasználásra. Az eredményeket a 6. ábra mutatja be.

Első ránézésre itt is felismerhető a korábban már sokszor megfigyelt heterogenitás. Azonban figyelembe véve a tényleges értékeket, némileg más kép bontakozik ki. Mint látható, valójában valamennyi szöveg viszonylag alacsony értékkel bír. A mérőszám még a legmagasabb értéket mutató szöveg (1016-3) esetében is csupán 24 a 100-as skálán. Ha felidézzük, hogy ez a szövegjellemző a szavak átfedésén, a szavak gyakoriságán, illetve a szintaktikai hasonlóságon alapul, világossá válik, hogy itt egyértelműen nem „könnyű” szövegekről van szó, ami tekintettel a vizsga B2-es szintjére, indokoltnak tűnik. Az is valószínűsíthető, hogy ez a viszonylagosan alacsony idegen nyelvi olvashatóság számos tényező együttes hatásaként jelenik meg. Természetesen



6. ábra. Coh-Matrix Idegen nyelvi olvashatóság eredmények

a legérdekesebb kérdés az, vajon ezek a szövegek megfelelnek-e a B2-es szintnek. Mint korábban jeleztük, egy szöveg szintjét nem feltétlenül lehet egyértelműen megállapítani, mégis érdemes lehet a KER egy releváns B2-es deskriptorát felidézni. Az „Általános szövegértés” B2-es deskriptorainak egyike így szól: „Széles körű aktív olvasási szókincsel rendelkezik, de ritkán előforduló idiómák esetében nehézségekbe ütközhet” (PTMIK 2002: 85). A „széles körű aktív olvasási szókincs” tehát elvárható ezen a szinten, s ez adott esetben kevésbé gyakori szavak ismeretét is szükségessé teheti. Tény az is ugyanakkor, hogy a KER deskriptorok közvetlenül nem érintik a szavak átfedését, illetve a szintaktikai hasonlóságot, így közvetlen összehasonlítás a KER-rel ezen szövegjellemző kapcsán sem tehető.

Az egyes szövegjellemzőket érintő áttekintést követően megállapítható, hogy a szövegjellemzők többsége kapcsán a szövegek között látványos különbségek detektálhatók. De hogy ezek a különbségek véletlenszerűnek tekinthetők, vagy tényleges figyelmet érdemelnek, azt az elemzés következő szakaszában, a Kruskal–Wallis-próba eredményeinek tükrében tudjuk eldönteni. Mivel vizsgálatunk célja elsődlegesen annak megállapítása volt, hogy az egyes vizsgaalkalmak kapcsán használt szövegek között létezik-e szignifikáns különbség, az egyes vizsgaalkalmak során használt négy (illetve a 2016. októberi vizsgán használt öt) szöveget egy-egy egységnek tekintettük. Ily módon összesen nyolc vizsgaidőszak, azaz nyolc szövegcsoportot hasonlítottunk össze. Ez alkalommal is célszerű először is szövegjellemzőnként megvizsgálni a különbségek szignifikáns mivoltát. Először most is a *narrativitás* került sorra. Az összehasonlítás eredményét az 1. táblázat mutatja be.

Test Statistics a,b

Narrativitás	
Kruskal-Wallis H	5,351
df	7
Asymp. Sig.	0,617

a. Kruskal Wallis Test

b. Grouping Variable: Vizsga

1. táblázat. Kruskal-Wallis-próba eredménye: narrativitás

Mint a táblázatból látszik, narrativitás szempontjából a vizsgákban használt szövegek között nem volt statisztikailag szignifikáns különbség ($p = 0,617$). Figyelemmel a korábban észlelt különbségekre, ez az eredmény némiképp meglepőnek tűnhet. Ne felejtjük azonban, hogy, amint az az *1. ábrán* is látható, a szövegek egy vizsgán belül is meglehetősen különböztek egymástól, s a jelek szerint ezek a differenciák mintegy kiegyenlítődték. Ez egyben azt is jelenti, hogy bár az egyes szövegek szintjén helyenként látványosak a különbségek, a vizsgaalkalmak szintjén a használt szövegek az áttekintett nyolc vizsgaidőszakban nem különböztek egymástól szignifikáns módon. Ez mindenképpen pozitívumnak tekinthető, hiszen a vizsga szintje adott, s így nyilvánvalóan az a cél, hogy a különböző vizsgaalkalmak azonos szintű kihívást jelentsenek a vizsgázóknak.

A következő szövegjellemző a *szintaktikai egyszerűség*. A vizsgák ezen szövegjellemző mentén történő összehasonlításának eredményét a *2. táblázat* mutatja be.

Test Statistics a,b

Szintaktikai egyszerűség	
Kruskal-Wallis H	4,745
df	7
Asymp. Sig.	0,691

a. Kruskal Wallis Test

b. Grouping Variable: Vizsga

2. táblázat. Kruskal-Wallis-próba eredménye: szintaktikai egyszerűség

Amint a táblázatból látszik, a vizsgákon használt szövegek között szintaktikai egyszerűség szempontjából sem volt szignifikáns különbség ($p = 0,691$). Ez az eredmény, a narrativitáshoz hasonlóan, arra utal, hogy bár a szövegek között ebben az esetben is voltak szembeszökő különbségek, a vizsgákon használt szövegcsoportok szintjén mindez már nem jelenik meg szignifikáns módon; a szövegek ezen paraméter mentén is kiegyenlítődték.

A vizsgák a harmadik szövegjellemző, a *szókonkrétság* alapján történő összehasonlításának eredménye a 3. táblázatban látható.

Test Statistics a,b

Szókonkrétság

Kruskal-Wallis H	6,857
df	7
Asymp. Sig.	0,444

a. Kruskal Wallis Test

b. Grouping Variable: Vizsga

3. táblázat. Kruskal-Wallis-próba eredménye: szókonkrétság

Az előző két szövegjellemzőhöz hasonlóan ebben az esetben sem derült fény szignifikáns különbségre ($p = 0,444$) a szövegek között. Ez ebben az esetben is arra utal, hogy a látványos különbségek a vizsgákban szereplő szövegcsoportok szintjén kiegyenlítődtek.

Vizsgálatunk következő szövegjellemzője a *referenciális kohézió*. Az összehasonlítás eredményét a 4. táblázat mutatja be.

Test Statistics a,b

Referenciális kohézió

Kruskal-Wallis H	6,088
df	7
Asymp. Sig.	0,530

a. Kruskal Wallis Test

b. Grouping Variable: Vizsga

4. táblázat. Kruskal-Wallis-próba eredménye: referenciális kohézió

Ebben az esetben is az látszik az eredményekből, hogy a szövegcsoportok között nincs szignifikáns különbség ($p = 0,530$). Tekintettel a 4. ábrán megfigyelhető eloszlásra, illetve arra a tényre, hogy ebből a szempontból eleve homogénebbnek látszottak a szövegek, kevésbé meglepő az eredmény. Ugyanakkor a kiegyenlítődési hatás vélhetően itt is hozzájárult a szignifikáns különbség hiányához.

Az ötödik szövegjellemző, a *mélykohézió* alapján történő összehasonlítás eredményét az 5. táblázat mutatja be.

Test Statistics a,b

Mélykohézió	
Kruskal-Wallis H	4,811
df	7
Asymp. Sig.	0,683

a. Kruskal Wallis Test

b. Grouping Variable: Vizsga

5. táblázat. Kruskal-Wallis-próba eredménye: mélykohézió

A korábban megfigyelt tendencia ebben az esetben is folytatódik. A vizsgákban szereplő szövegcsoportok között ezúttal sem mutatható ki szignifikáns különbség ($p = 0,683$). Megállapítható tehát, hogy a vizsgákat alkotó szövegcsoportok szintjén valamennyi TERA-szövegjellemző esetében kiegyenlítődték az egyes szövegek szintjén létező különbségek.

Vajon hasonlóan működött-e a Coh-Metrix *idegen nyelvi olvashatóság* szövegjellemzője esetében is a kiegyenlítődé? A 6. ábrán látható ennek az összehasonlításnak az eredménye.

Test Statistics a,b

Idegen nyelvi olvashatóság	
Kruskal-Wallis H	6,420
df	7
Asymp. Sig.	0,492

a. Kruskal Wallis Test

b. Grouping Variable: Vizsga

6. táblázat. Kruskal-Wallis-próba eredménye: idegen nyelvi olvashatóság

Mint a táblázatból látszik, ebben az esetben sem mutatható ki szignifikáns különbség ($p = 0,492$) a szövegcsoportok között. Ez egyben azt is jelenti, hogy valamennyi vizsgált szövegjellemző esetében ugyanazt a mintázatot lehet megfigyelni: az egyes szövegek szintjén léteznek látványos különbségek, a vizsgákban szereplő szövegcsoportok azonban nem különböznek egymástól szignifikáns módon.

Még egy kérdést azonban szükséges megvizsgálunk. A szövegek egyes szövegjellemzői mind egy-egy aspektusát tükrözik a vizsgált szövegeknek. A tényleges olvashatóság, illetve a szöveg nehézsége azonban vélhetően pontosabban megragadható, ha az egyes szövegjellemzőket kvantifikáló számokat egy kompozit mérőszámként

összesítjük, hiszen ez a szám tükrözi az egyes szövegjellemzők által képviselt minőségek összességét. Ez a megközelítés már más hasonló kutatás során is alkalmazásra került (lásd pl. Szabó 2014).

Vizsgáljuk meg tehát most azt, hogy van-e különbség a vizsgákon használt szövegcsoportok között az összesen hat szövegjellemzőből összeálló kompozit szövegnehézségi mutató tekintetében. Az összehasonlítás eredményét a 7. táblázat mutatja be.

Test Statistics a,b

Idegen nyelvi olvashatóság	
Kruskal-Wallis H	6,420
df	7
Asymp. Sig.	0,492

a. Kruskal Wallis Test

b. Grouping Variable: Vizsga

7. táblázat. Kruskal-Wallis-próba eredménye: kompozit szövegnehézség

A korábbi eredmények ismeretében immár nem meglepő módon azt látjuk, hogy a szövegcsoportok között ezen mutató alapján sem lehet szignifikáns különbséget ($p = 0,808$) kimutatni. Mindez arra utal, hogy a kutatás során elemzett szövegek a nyolc vizsgaalkalom kapcsán semmilyen vizsgált szövegmutató tükrében sem különböztek egymástól szignifikáns módon.

Következtetések

A kutatás eredményeinek ismeretében bizton állíthatjuk, hogy az angol nyelvi emelt szintű érettségi vizsga olvasáskomponensében alkalmazott szövegek nehézsége nagyfokú stabilitást mutat a kutatás során vizsgált időszakban. Ez a tény igen lényeges információ, ami azt sejteti, hogy a mérés konzisztenciája magas. A szövegcsoportok hasonló nehézsége azt valószínűsíti, hogy a vizsgázók vizsgaalkalomtól függetlenül hasonló kihívással kerülhettek szembe a vizsga olvasásértést mérő komponensében, s ez egyik fontos eleme a mérés megbízhatóságának. Szükséges azonban megállapítanunk, hogy a jelen kutatás alapján továbbra is csak részleges információkkal bírunk a vizsga tényleges nehézségéről, illetve minőségéről.

Egyrészt igen lényeges hangsúlyoznunk, hogy a vizsgálat csak a szövegekre terjedt ki, a feladatokra nem. Márpedig egy olvasásértést mérő vizsga tényleges nehézsége minden esetben a szövegnehézség és a feladatnehézség közös produktuma, sőt, még maga a vizsgázó is tényező a nehézség szempontjából (Castello 2008). Egy viszonylag nehéz szöveg esetében is fel lehet tenni könnyű kérdéseket, és az

is előfordulhat, még ha ez nem is kívánatos, hogy egy viszonylag könnyű szöveg alapján nehezen megválaszolható itemekkel találkozik a vizsgázó. Ugyanakkor a szövegnehézség konzisztenciája, ha nem is garantálja, de mindenképpen segíti a feladatok nehézségének stabilitását. Más szóval, ha a kutatás lényegi különbségeket mutatott volna ki a különböző vizsgákon alkalmazott szövegek között, kisebb lenne a valószínűsége annak, hogy a feladatok hasonló nehézségűek. Jelen esetben azonban inkább azt valószínűsíthetjük, hogy a hasonlóan nehéz szövegekhez hasonlóan nehéz feladatok társulhatnak.

További lényeges körülmény, hogy a kutatás csak a szövegek nehézségének stabilitását igazolta, azt azonban ezek után sem tudhatjuk bizonyosan, hogy ez a nehézségi szint megfelel-e a vizsga célszintjének. Ennek egyik oka a fentebb már jelzett koncepcionális különbség a szövegek és a feladatok szintje között. A másik ok azonban az, hogy míg a jelen kutatás során a szövegek objektíven mérhető, *kvantitatív* karakterisztikái alapján végeztük az elemzést, a vizsga szintbesorolása a KER, illetve a KER alapján kifejlesztett nyelvspecifikus háttéranyagok *kvalitatív* deskriptorain alapul. Amint pedig fentebb némely szövegjellemző kapcsán láthattuk is, a két megközelítés között nem könnyű egyértelmű kapcsolatot találni, bár ez a terület a jövő ígéretes kutatásainak adhat teret.

Végezetül megállapíthatjuk, hogy a szövegek nehézségének objektív mutatók alapján történő mérése a jövőben hasznos segítséget nyújthat a tesztfelkészítők számára, akár az érettségiről, akár más nagy tétellel bíró mérési kontextusról legyen szó. Ez a sokrétű alkalmazhatóság hozzájárulhat a jobb feladatok, a jobb tesztek megalkotásához, segítve ezzel mind a nyelvtanulást, mind a nyelvtanítás folyamatát.

IRODALOM

- Alderson, J. C. (2000): *Assessing reading*. Cambridge: Cambridge University Press.
- Bachman, L. F. – Lynch, B. K. – Mason, M. (1995): Investigating variability in tasks and rater judgements in a performance test of foreign language speaking. *Language Testing*, 12/2, 238–257.
- Brown, J. D. (1998): An EFL readability index. *JALT Journal*, 20/2, 7–36.
- Castello, E. (2008): *Text complexity and reading comprehension tests*. Bern: Peter Lang.
- Council of Europe (2018): *Common European Framework of Reference for languages: learning, teaching, assessment. Companion volume with new descriptors*. Strasbourg: Council of Europe.
- Crossley, S. A. – Allen, D. B. – McNamara, D. S. (2011): Text readability and intuitive simplification: A comparison of readability formulas. *Reading in a Foreign Language*, 23/1, 84–102.
- Crossley, S. A. – Greenfield, J. – McNamara, D. S. (2008): Assessing text readability using cognitively based indices. *TESOL Quarterly*, 42/3, 475–493.

- Einhorn Ágnes – Major Éva (2006): Az idegen nyelvek – vizsgafejlesztés nemzetközi kontextusban. In: Horváth Zsuzsanna – Lukács Judit (szerk.) *Új érettségi Magyarországon. Honnan, hová, hogyan? Egy folyamat állomásai*. Budapest: Országos Közoktatási Intézet, 127–137.
- Graesser, A. C. – McNamara, D. S. (2011): Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science*, 3, 371–398.
- Graesser, A. C. – McNamara, D. S. – Kulikowich, J. M. (2011): Coh-Metrix: Providing Multilevel Analyses of Text Characteristics. *Educational Researcher*, 40/5, 223–234.
- Halápi Magdolna – Hegedűs Kristóf (2013): *A 2012. május-júniusi érettségi feladatsorok megoldásainak itemszintű rögzítése, a feladatsorok és az egyes feladatok mérésmethodikai vizsgálata a TÁMOP-3.1.8-09/1-2010-0004 azonosító számú projekt keretében. Angol nyelv. Közép- és emelt szint*. https://www.oktatas.hu/pub_bin/dload/unios_projektek/tamop_318/meresmethodika/AngolNyelv.pdf (hozzáférés: 2019. 04. 26.)
- Halliday, M. A. K. (1978): *Language as social semiotic. The social interpretation of language and meaning*. London: Edward Arnold.
- Jackson, G. T. – Allen, L. K. – McNamara, D. S. (2017): Common Core TERA: Text Ease and Readability Assessor. In: Crossley, S. A. – McNamara, D. S. (eds.) *Adaptive Educational Technologies for Literacy Instruction*. New York: Routledge, 49–68.
- Jilly Viktor – Nagy Ildikó – Enyedi Ágnes – Wintsche Gergely – Morvay Zsuzsanna (2014): „Közép- és emelt szintű értékelési skálák összehasonlítása” tárgyú kutatás-fejlesztési projekt. *Összehasonlító elemzés angol nyelv vizsgatárgyból*. https://www.oktatas.hu/pub_bin/dload/unios_projektek/tamop318/ertekelesi_skalak_osszehasonlitasa/ertekelesi_skalak_angol_nyelv.pdf (hozzáférés: 2019. 04. 26.)
- Kecker, G. – Eckes, T. (2010): Putting the Manual to the test: the TestDaf-CEFR linking project. In: Martyniuk, W. (ed.), *Aligning tests with the CEFR: reflections on using the Council of Europe’s draft Manual*. Cambridge: Cambridge University Press, 50–79.
- Klare, G. R. (1974–1975): Assessing readability. *Reading Research Quarterly*, 10, 62–102.
- Krashen, S. D. (2004): *The power of reading: Insights from the research (2nd ed.)*. Portsmouth, NH: Heinemann.
- Major Éva (2006): Szövegek, feladatok, vizsga. Az angol nyelvi érettségi feladatsorok összeállításának folyamata. In: Horváth Zsuzsanna – Lukács Judit (szerk.) *Új érettségi Magyarországon. Honnan, hová, hogyan? Egy folyamat állomásai*. Országos Közoktatási Intézet., Budapest, 139–153.
- McNamara, D. S. – Graesser, A. C. – Cai, Z. – Kulikowich, J. M. (2011): *Coh-matrix easability components: Aligning text difficulty with theories of text comprehension*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- McNamara, D. S. – Louwse, M. M. – Cai, Z. – Graesser, A. (2013): *Coh-Metrix version 3.0*. <http://cohmetrix.com> (hozzáférés: 2019. április 30.)
- PTMIK (2002): *Közös európai referenciakeret. Nyelvtanulás, nyelvtanítás, értékelés*. Budapest-Pilisborosjenő: Pedagógustovábbképzési Módszertani és Információs Központ Kht.
- Szabó, G. (2014): Applying objective measures of text difficulty in the comparison of texts in reading comprehension tasks. In: Horváth, J. – Medgyes, P. (eds.) *Studies in Honour of Marianne Nikolov Pécs*: Lingua Franca Csoport, 385–398.

- Vígh Tibor (2008): Egy IRT-alapú nyelvi feladatbank létrehozásának módszertani kérdései. A német nyelvi érettségi vizsgafeladatok eredményei. *Magyar Pedagógia*, 108/1, 29–51.
- Vígh Tibor (2013): A középszintű angol és német nyelvi érettségin elért teljesítmények változásai 2007 és 2012 között. *Modern Nyelvoktatás*, 19/1–2, 17–35.

FORRÁSOK

- 100/1997. (VI. 13.) *Korm. rendelet az érettségi vizsga vizsgaszabályzatának kiadásáról.*
http://njt.hu/cgi_bin/njt_doc.cgi?docid=30517 (hozzáférés: 2019. 06.12.)
- https://nyak.oh.gov.hu/doc/statisztika.asp?strId=_43ny (hozzáférés: 2019. 06.12.)
- <https://www.ketszintu.hu/publicstat.php> (hozzáférés: 2019. 06.12.)

KISZELY ZOLTÁN

Értékelési skálák használata a beszédkésztség és íráskésztség vizsgákon

Rating scales in speaking and writing performance assessment in high-stakes examinations

While assessing examinees' speaking and writing performances in high-stakes examinations raters / examiners use rating scales. Rater effects like the halo effect, the severity effect, central tendency or inconsistency as construct-irrelevant factors should be minimized by language examination providers; therefore, it is of utmost importance to investigate similarities and differences in raters' interpretation of the rating scale. For this reason, the present study aims to compare the way in which the same group of examiners use the rating scales for the assessment of candidates' English B2 level performance in speaking and writing. The data were collected from the online retraining sessions organised for examiners at the BME Language Examination Centre in 2017 and 2018. First, the paper describes the most salient features of the rating scales for assessing speaking and writing performance. Second, it examines how stringently raters apply the different criteria of the two rating scales. Third, it discusses the similarities and differences in the use of the rating scales for evaluating candidates' production skills.

Keywords: productive skills, performance assessment, rating scales, rater effects, classical and modern test theory

Bevezetés

Minden Magyarországon akkreditált nyelvvizsga a vizsgázók négy nyelvi készségét méri és értékeli. Ezek közül a produktív készségek, vagyis a beszédkésztség és az íráskésztség értékelése olyan eljárás, amikor a vizsgázók kompetenciáját performanciájukon, azaz nyelvi teljesítményeiken keresztül vizsgálják (McNamara 1996). Az értékeléshez mérőeszközökre van szükség, azaz egy vagy több olyan feladatra, amelyek célja bizonyíthatóan az adott készség mérése, továbbá egy többnyire analitikus, azaz különböző értékelési szempontokat tartalmazó értékelési skálára, amelynek segítségével érvényes és megbízható módon lehet értékelni a teljesítményeket, és végül, de nem utolsósorban jól kiképzett értékelőkre, akik megfelelően használják az értékelési skálákat. A beszédkésztség és az íráskésztség ugyan eltérő konstruktumot takar, azaz nem ugyanazt méri, de mindkettő produktív készség, és az értékelés során használt eszközök és módszerek hasonlóak.

Felületesen gondolkodva lehetne azt állítani, hogy a skálahasználat egyértelműen annak a függvénye, hogy milyen nyelvi teljesítményt nyújtott a vizsgázó. Ha jól szerepelt, magas pontszámokat kap, ha gyengébben, akkor alacsonyabbakat. Ennek ideális esetben így is kellene lennie, de mivel humán értékelőkről van szó, nem csupán a vizsgázói nyelvi teljesítmény az egyetlen értékelést befolyásoló tényező, arra számos szubjektív elem is hat. Eckes (2009) jó néhány ilyen tényezőt sorol fel. Idetartozik például az értékelők szakmai háttere, a személyiségük és a leterheltségük, hogy csak néhányat említsünk. E szubjektív elemekkel terhelt értékelői hatásoknak (*rater effects*) a kiküszöbölése, illetve minimalizálása minden vizsgarendszer alapvető feladata, ugyanis az értékelői hatásoknak konstruktum-irreleváns (a mérés fókuszához nem tartozó) tényezőkként közvetlen következményük van az értékelés pontosságára vonatkozóan. A legfontosabb kérdés a skálahasználattal kapcsolatban tehát az, hogy a skála kritériumait megfelelő módon értelmezik-e a vizsgáztatók (Eckes 2019).

A fentiek miatt belátható, hogy az értékelés, azaz az értékelési skála használata olyan sokismeretlenes egyenlet, amelyet érdemes behatóbban is megvizsgálni. Mind a beszéd-készség, mind pedig az íráskészség tekintetében a szakirodalom már részletesen foglalkozott az értékelési skálák és a vizsgáztatók interakciójával (Kuiken–Vedder 2014, Wind–Peterson 2017), de általában ezek a kutatások a két készséget külön-külön, nem egymással összehasonlítva vizsgálták. Jelen tanulmány ezért arra tesz kísérletet, hogy a BME Nyelvvizsgaközpont 2017. és 2018. évi vizsgáztató-továbbképzése során összegyűlt adatok segítségével összehasonlítsa, hogyan használják ugyanazon angol nyelvi vizsgáztatók az íráskészség és a beszéd-készség méréséhez készült értékelési skálákat két-két B2 szintű, de különböző nyelvi színvonalú nyelvvizsgázói teljesítmény értékelése alapján. A tanulmány az értékelők, az értékelési skálák és a vizsgázók közötti interakcióval foglalkozik, de e három tényező közül elsősorban az értékelési skála használatára fókuszál, ebből a szempontból vizsgálja a kérdést. Annak ellenére, hogy az adatok vizsgáztató-továbbképzésekből származnak, a tanulmánynak nem célja annak bemutatása és elemzése, hogy a vizsgáztatók értékelése mennyire közelített a vizsgaközpont által felállított standardokhoz.

Az első részben rövid szakirodalmi áttekintést adunk az értékelési skálák használatával kapcsolatos korábbi kutatásokból. Ezt követően a kutatási kérdések és a választott módszertan leírása következik. Az elemzés során először a beszéd-készség és az íráskészség értékelésére alkalmazott skálák használatának modern tesztelméleti eszközökkel kimutatott jellemzőit vizsgáljuk, majd a skálák egyes kritériumaival kapcsolatos vizsgáztatói szigort helyezük a középpontba. Ezt követi annak vizsgálata, hogy milyen összefüggések és különbségek mutathatók ki a két skála használatában. A tanulmányban az *értékelő* és *vizsgáztató* szavakat, valamint az *értékelési kritérium* és az *értékelési szempont* kifejezéseket szinonimaként fogjuk használni.

Az eredményekből egyrészt hasznos tanulságok vonhatóak le a vizsgaközpont számára, illetve az eddigi esetleges megérzések konkrét adatokkal támaszthatók alá, másrészt pedig kirajzolódhat egy kép a mai nyelvtanári/nyelvvizsgáztatói szakma egy részének produktív készségeikkel kapcsolatos látásmódját és attitűdjét illetően.

Szakirodalmi áttekintés

Ahogy a bevezetőben említettük, jelen tanulmány az értékelési szempontok felől közelíti meg az értékelés problematikáját, azonban a kritériumok, azaz a skála használatára a vizsgáztatói értékelésekből lehet következtetéseket levonni, ezért a szakirodalmi áttekintés során az értékelőket és a skála használatát nem lehet szétválasztani.

Az értékelői hatásokkal sokan foglalkoztak, de talán a legátfogóbb áttekintést Myford–Wolfe (2003) tanulmánya adja. Az általuk említett hatások közül a legfontosabbak a következők. Holdudvarhatásról (*halo effect*) akkor beszélünk, ha az analitikus értékelési skála egyik kritériumára adott pontszám túlságosan befolyásolja a többi kritériumra adott pontszámot, például, ha a nyelvhelyességi szempont gyenge pontszámot kap, akkor a szókincs is és a tartalom is automatikusan kevés pontot kap. Az elfogultság az egyik kritériummal vagy feladattal szemben azt jelenti, hogy túlságosan szigorú vagy engedékeny a pontozás az értékelés egy aspektusával kapcsolatban. A középső pontszámra törekvés tendencia (*central tendency*) azt jelenti, hogy a kockázatok elkerülése végett a vizsgáztatók nem használják a skála szélső értékeit, ezért értékeléseikben többnyire a középső pontszámok szerepelnek, például egy ötfokozatú (1–5) skála esetén a 3 a középső pontszám. A következtelenség az a jelenség, amikor egy-egy kritériumra egészen más pontszámot ad egy értékelő, mint a többi. E négy értékelői hatás közül kettőre, a holdudvarhatásra és a középső pontszámra törekvés tendenciára az értékelési szempontok felől is lehet következtetni, ahogy azt később látni fogjuk.

Az értékelési skálák validálásának statisztikai eszközeiről Eckes (pl. 2009, 2015, 2019) több tanulmányában is világos összegzést nyújt. A kritériumok modern tesztelméleti eszközökkel kimutatható adataiból egyrészt megállapítható, hogy a skála szerkezete a pontozás szempontjából megfelelően működik-e, másrészt az is kimutatható, hogy az analitikus skálák egyes kritériumai milyen mértékben járulnak hozzá az adott készség méréséhez, azaz az adott készséget mérik-e, vagy esetleg valami mást.

Az egyes kritériumokkal szembeni szigorúság kérdésével szintén több tanulmány foglalkozott. McNamara (1990) egy ausztrál egészségügyi dolgozók számára kifejlesztett szaknyelvi teszt validálásakor két különböző időpontban is azt találta, hogy mind a beszéd-, mind az íráskészség teljesítmények értékelésekor a nyelvtan a domináns, azaz a legszigorúbban értékelt kritérium. A TestDaF, német mint idegen nyelvi vizsga során megírt íráskészség teljesítmények elemzésekor Eckes (2009) kimutatta, hogy a nyelvhelyesség (*linguistic realisation*) volt a legszigorúbban értékelt kritérium.

Corrigan (2007) olasz beszéd-készségvizsgán nyújtott nyelvi teljesítmények értékelése során szintén azt állapította meg, hogy a nyelvtani pontosság (*accuracy*) az a kritérium, amit az ítések a legszigorúbban értékelték, az interakciót és a folyékonyságot pedig a legenyhébben. Velük ellentétben Benke (2008) íráskészség-értékelők és -értékelési skála közötti interakciókat kutatva nem talált olyan kritériumot, amelyet konzekvensen szigorúbban vagy engedékenyebben értékelték volna a vizsgáztatók. Ez arra utal, hogy az értékelők szigora vagy engedékenysége nagyon feladatfüggő, és azt is megerősíti, hogy szituációfüggő az értékelők értékelőskála értelmezése (Lumley–McNamara 1995).

Tudomásunk szerint a szakirodalom nem foglalkozott közvetlenül azzal a kérdéssel, hogy ugyanaz az értékelői csoport a két produktív készség nyelvi teljesítményeit mennyire hasonlóan vagy különbözően értékeli, azaz az értékelési skálákat hogyan használja, ezért jelen tanulmány erre a kérdésre fókuszál.

Kutatási kérdések

A vizsgálat három kérdésre kereste a választ, amelyek a következők voltak:

1. Milyen jellemzőkkel rendelkeznek a BME B2 szintű beszéd-készség- és íráskészség-értékelési skálái a vizsgáztatói értékelések tükrében?
2. Feladatonként a beszéd-készség- és az íráskészség-teljesítményeket mely kritériumok mentén értékelték a legszigorúbban és a legenyhébben a vizsgáztatók 2017-ben és 2018-ban?
3. Milyen összefüggések és különbségek mutathatók ki a beszéd-készség és az íráskészség értékelésére használt skálák alkalmazása között?

Kutatási módszerek

Résztevők

A vizsgálat körébe 67, a BME általános nyelvvizsgarendszeréhez akkreditált angol nyelvi vizsgáztató értékeléseit vontuk be, akik 2017-ben és 2018-ban részt vettek mind beszéd-készség, mind pedig íráskészség vizsgarészekhez szervezett online vizsgáztató-továbbképzéseken.

A továbbképzés során értékelt szóbeli és írásbeli nyelvi teljesítményeket minden évben az angol nyelvi felelősök választották ki. A kiválasztáskor a beszéd-készség esetén a felvétel minősége mellett fontos szerepet játszott az is, hogy a két különböző évben különböző színvonalú teljesítmény legyen az értékelés tárgya, azaz ne hasonló pontszámú teljesítmények kerüljenek értékelésre.

Vizsgáztató-továbbképzés

A résztvevők 2017-ben és 2018-ban beszédkézség- és íráskészség-fókuszú online formátumú vizsgáztató-továbbképzésen vettek részt, amelynek során mindkét évben egy-egy vizsgázó beszédkézség- és íráskészség-teljesítményét értékelték a BME rendszerszerűen alkalmazott analitikus értékelési skálái segítségével.

A beszédkézség feladatlapon mindkét évben egy-egy B2 szintű vizsga hangfelvétele szerepelt, ami annyiban különbözött az éles vizsgától, hogy mivel audio anyagról volt szó, sem a vizsgázót, sem a vizsgáztatókat nem lehetett látni. A feladatlapon a vizsgán (azaz a hangzó anyagban) felhasznált feladatokat és az értékelési szempontokat is feltüntettük, így a hallgatás közben a feladatokat és a kritériumokat is egyszerre lehetett tanulmányozni. Az íráskészség feladatlapon mindkét évben egy-egy vizsgázó egy B2 szintű feladatsorhoz tartozó két feladatmegoldása szerepelt. A feladatlapon itt is feltüntettük a feladatléírásokat és az értékelési szempontokat. Értékelést mind a beszédkézség, mind az íráskészség esetén csupán egy alkalommal lehetett beküldeni. Mindkét feladatlap kitöltésére hozzávetőleg 25-25 percet irányoztunk elő, de szigorú időkorlátot nem iktattunk a rendszerbe (a vizsgáztató-továbbképzés további részleteiről lásd Kiszely 2018).

Feladatok és értékelési skálák

A BME általános nyelvvizsgarendszere a vizsgázók beszédkézségét B2 szinten három feladat segítségével méri. Az elsőben a vizsgázónak néhány kifejtő kérdés alapján kell a mindennapi életben előforduló témák széles körében saját személyével kapcsolatban beszélgetnie a vizsgáztatóval. A második részben a vizsgázó önállóan fejt ki gondolatait egy adott témáról, amelyhez képi stimulus tartozik. A harmadik részben a vizsgázónak szerepjátékot kell eljátszania a vizsgáztatóval, célnyelven leírt szituációs feladat alapján.

A feladatokat analitikus értékelési skála segítségével értékeljük, amely a következő öt kritériumot tartalmazza: *feladatmegoldás*, *kifejezőkézség*, *nyelvhelyesség*, *folyamatosság és koherencia*, valamint *kiejtés*. A *feladatmegoldás* kritérium a feladat végrehajtását, a beszédszándékok megvalósítását, a tartalmi relevanciát, a nyelvi funkciók megfelelő alkalmazását és a beszédértést jelenti. A *kifejezőkézség* a szókincs terjedelmére és alkalmazására, a stílus és a regiszter használatára utal. A *nyelvhelyesség* a morfológiai és szintaktikai elemek változatosságát és alkalmazását jelenti. A *folyamatosság és koherencia* a folyamatosságot és az információ szerkesztettségét, míg a *kiejtés* a hangképzést, a szó-, mondat- és beszédhangsúlyt takarja.

A *feladatmegoldást*, a *kifejezőkézséget* és a *nyelvhelyességet* a három feladat esetén külön-külön értékeljük egy 0-tól 5 pontig terjedő skálán. A *folyamatosság és koherencia* kritériumot az első és a harmadik feladat esetében együttesen használjuk szintén egy 0-tól 5 pontig tartó skálán, mert mindkét feladat interaktív jellegű dialógus, míg a

második feladat esetén külön értékeljük, mert ez egy monológ. A *kiejtés* kritériumot a három feladaton átívelően egyszer értékeljük szintén egy 0-tól 5 pontig tartó skálán. Az egyes szempontokra 0–5 pont adható, így összesen 60 nyerspontot lehet elérni.

A BME általános nyelvvizsgarendszere a vizsgázók íráskészségét B2 szinten két feladat segítségével méri. Az elsőben a vizsgázónak informális e-mailt kell írnia egy külföldi ismerősnek általános témákban 170–200 szóban négy megadott tartalmi pont alapján. A második részben internetes fórumhozzászólást kell írni 120–140 szóban, szintén négy szempont alapján.

A két feladatot az értékelők külön-külön értékelik analitikus értékelési skála segítségével, amely a következő négy kritériumot tartalmazza: *kommunikatív érték*, *kifejezőkészség*, *nyelvhelyesség* és *általános benyomás*. A *kommunikatív érték* kritérium a kommunikációs cél elérését, a tartalmi pontok kidolgozását, a kohéziót és a szöveg terjedelmét jelenti. Ez a kritérium gyakorlatilag megfeleltethető a beszéd-készségnél alkalmazott *feladatmegoldás* szempontnak. A *kifejezőkészség* és a *nyelvhelyesség* kritériumok ugyanazt jelentik, mint a beszéd-készség esetén, amelyekhez természetesen a helyesírás még hozzáadódik. Az *általános benyomás* kritérium alatt a szöveg természetességét, illetve az olvasóra gyakorolt hatását értjük. Az egyes kritériumokra 0–5 pont adható, így összesen 40 nyerspontot lehet szerezni.

Az elemzés eszközei

A kutatás a klasszikus és a modern (probabilisztikus) tesztelmélet alapján készült programok segítségével kereste a kérdésekre a választ. A klasszikus tesztelméleten alapuló leíró statisztikai adatokat az Excel táblázatkezelő, az összefüggés- és különbözőség-vizsgálatok adatait pedig az SPSS-programcsomag segítségével számoltuk ki.

A modern tesztelmélet alapján kimutatható eredményeket a Facets programmal (Linacre 2018) állítottuk elő. Ez utóbbi mind az értékelőkről, mind a skáláról, mind pedig a vizsgázókról olyan információkat képes nyújtani, amire a klasszikus tesztelmélet nem képes (Bond–Fox 2015). Facetsnek az értékelést befolyásoló tényezőket nevezzük; ilyen például a vizsgázói nyelvtudás, az értékelői szigor, az értékelési skála kritériumai, a téma nehézsége és maga a feladat is, hogy csak néhányat említsünk a sok közül. A Facets program elvi alapját jelentő többtényezős Rasch-modell (*Many-Facet Rasch Measurement*, rövidítése: MFRM) képes megállapítani a vizsgázók vizsgán elért konkrét eredményét, a vizsgázók becsült képességszintjét, az értékelési szempontok megfigyelt és becsült nehézségét és az értékelők megfigyelt és becsült szigorát azzal a céllal, hogy semlegesítse az értékelői és egyéb hatásokat, azaz a konstuktum-irreleváns tényezőket a vizsgázók teljesítményének értékelésében. Az MFRM ugyanarra a skálára helyezi az értékelői szigorúságot, az itemek (kritériumok) nehézségét és a vizsgázók képességszintjét (Bond–Fox 2015). Más szavakkal kifejezve tehát az egyes kritériumokra adott pontok alapján a többtényezős Rasch-modell egy

valószínűségi becslést, egy modellt ad arra vonatkozóan, hogy egy bizonyos értékelő egy bizonyos vizsgázói teljesítmény egy bizonyos kritériumára milyen valószínűséggel ad egy bizonyos pontszámot, és ezt a becslést összehasonlítja a tényleges pontszámokkal (McNamara 1996). (A Rasch-modellről magyarul lásd még pl. Molnár 2013.)

Eredmények és elemzés

Az alábbiakban a kutatási kérdések sorrendjében közöljük és elemezzük az adatokat.

1. Milyen jellemzőkkel rendelkeznek a BME B2 szintű beszédképesség és az írásképesség értékelési skálái a vizsgázatói értékelések tükrében?

Az 1. és a 2. táblázat a Facets program segítségével készült. Ezekon a két év adatai képségenként és kritériumonként egyben láthatóak, például az 1. táblázat első sorában a 2. feladat nyelvhelyesség (*nyelvhelyesség2*) kritériumának 2017-es és 2018-as adatai összesítve szerepelnek.

Az első oszlopban a kritériumra adott összpontszám szerepel. A második oszlop az értékelések számát jelenti, ami azért 134, mert a 67 vizsgázató az adott kritériumot kétszer értékelte, egyszer az egyik évben, a következő évben pedig még egyszer ($67 \times 2 = 134$). A harmadik oszlop az adott szempontra adott pontszámok átlagát mutatja, míg a negyedik oszlop a program által elvárt (fair) átlageredményt tartalmazza. Ez az a számsor, amit a többszörös Rasch-modell alapján a Facets megállapít, kizárva, vagy legalábbis csökkentve a nem kívánatos értékelői hatásokat. Az ötödik oszlop a logitban meghatározott nehézségi értéket mutatja: minél nagyobb ez a szám, annál nehezebb az item, azaz annál szigorúbban értékelték azt a kritériumot. Az infit mnsq és az outfit mnsq értékek az úgynevezett illeszkedési (fit) statisztikák, amelyeknek értékelési skálák esetén Wright–Linacre (1994) szerint a 0,6–1,4 sávban kell mozogniuk. Amennyiben ez a helyzet, akkor az értékelők az elvártaknak megfelelően használják a kritériumokat. Az infit és outfit zstd értékek, amelyek statisztikai hipotézis-tesztként funkcionálnak, -2 és $+2$ között megfelelőek (Linacre 2003). A legutolsó oszlopban a kritériumok találhatóak egymás alatt nehézségi sorrendbe állítva, azaz például a beszédképesség esetén a *nyelvhelyesség2* a legszigorúbban, a *feladatmegoldás1* pedig a legenyhébben értékelt kritérium.

A táblázatok mindkét késziségnél azt mutatják, hogy a megfigyelt pontszám minden egyes kritérium esetén, a beszédképéségnél nagyobb, az írásképéségnél kisebb mértékben, magasabb, mint a fair pontszám, tehát enyhébben értékelték a vizsgázatók a modell által elvártnál. Ez alól a beszédképéségnél a *feladatmegoldás1*, írásképéségnél pedig az *általános benyomás1* a kivétel, de csupán elhanyagolható mértékben.

A beszédképesség esetén a legszigorúbban a *nyelvhelyesség2*, a *folyékonyág2* és a *nyelvhelyesség3*, legenyhébben pedig a *feladatmegoldás1*, a *feladatmegoldás3*, valamint a *kifejezőképesség1* kritériumokat értékelték. Figyelemre méltó, hogy az 1. feladatot,

azaz az ahhoz tartozó kritériumokat enyhébben értékelték, mint a 2. és 3.-hoz tartozókat, aminek több oka lehet (átlagok: 1: 3,73; 2: 3,29; 3: 3,55). Az egyik az, hogy valóban itt nyújtották a legjobb teljesítményt a vizsgázók, hiszen talán ez a legkönnyebben teljesíthető feladat, lévén a vizsgázó személyéhez kapcsolódóan folyik a társalgás, míg a 2. és a 3. feladat nagyobb kihívás elé állítja a vizsgázókat. A másik lehetséges ok pedig pszichológiai természetű a vizsgáztatók oldaláról: nem kizárt, hogy az első feladatot még enyhébben értéklik, mivel ez a kezdete a vizsgának, amikor még a vizsgázók kissé feszültebbek, ezért ekkor még nem fog olyan erősen a ceruza. Megjegyzendő azonban, hogy az átlagok közötti különbség egyik esetben sem volt szignifikáns, tehát messzemenő következtetést jelen adatokból levonni nem lehet (Wilcoxon próba: 1–2: $Z = -1,604$ $p = 0,109$; 1–3: $Z = -0,272$ $p = 0,785$; 2–3: $Z = -1,826$ $p = 0,068$).

Az íráskészség esetén a legszigorúbban a *kommunikatív érték2*, a *nyelvhelyesség2* és a *nyelvhelyesség1*, legenyhébben pedig az *általános benyomás1*, a *kommunikatív érték1* és a *kifejezőkészség2* kritériumokat értékelték. Érdekes, hogy az 1. feladatot, azaz az ahhoz tartozó kritériumokat enyhébben értékelték, mint a 2.-hoz tartozókat (átlagok: 1: 3,66; 2: 3,46). Ennek az lehet az oka, hogy annak ellenére, hogy ez utóbbi feladattípus (internetes fórumhozzászólás) már 2016 novembere óta részét képezi a vizsgafeladatoknak, az értékelők még mindig nem tudnak megfelelő módon különbséget tenni az 1. és 2. feladattal kapcsolatos elvárások között, azaz az első feladatra vonatkozó elvárásait rávetítik a másodikra, de ott nem kapják meg ugyanazt a kidolgozottságot, hiszen ez egy más műfajú és rövidebb írásmű, mint az első, ezért szigorúbban értékelnek. Természetesen az is lehetséges, hogy a vizsgázók valóban jobb teljesítményt nyújtottak ezen szempont mentén az 1. feladatban, mint a másodikban. Az átlagok közötti különbség azonban ebben az esetben sem volt szignifikáns, tehát az eredmény a véletlen műve is lehet (Wilcoxon próba: $Z = -1,095$ $p = 0,273$).

Az illeszkedési (infit és outfit) statisztikák egy-egy kivétellel az elvárt 0,6–1,4 értékek között mozognak mindkét készség szinte minden kritériuma esetén, ami azt jelenti, hogy készségenként az összes kritérium egységesen hozzájárul a készség egydimenziósságának bizonyításához, azaz a készségekhez tartozó kritériumok ugyanannak a készségnek egy-egy részképessége, ugyanannak a készségnek az alkomponensei (Eckes 2009). Az egyik kivétel a beszéd-készségskálán található *kiejtés* kritérium, ahol az illeszkedési statisztikák túl alacsonyok (0,47 és 0,42) csakúgy, mint a zstd értékek, ami, bár nem teljesen egyértelműen, utalhat arra, hogy ezen kritérium esetén csoportszinten egy jellegzetes értékelői hatásról, a szakirodalmi áttekintésben említett középső pontszámra törekvés tendenciáról beszéljünk. Utóbbira a többi kritérium esetén nincs utalás az adatok alapján. Az íráskészség esetén a kivétel a *nyelvhelyesség1* kritérium, ahol mind az infit, mind az outfit értékek kismértékben ugyan, de magasnak bizonyultak (1,43 és 1,43), és a zstd értékek itt sem megfelelőek, ami elvben azt jelentené, hogy ennél a kritériumnál tapasztalható a legnagyobb variabilitás, azaz

kissé zűrzavaros az értékelés. Ennek azonban esetünkben nincs jelentősége a kritériumhasználatot tekintve, ugyanis a magas fit értékek oka bizonyára az, hogy ez az a kritérium, amelynél a legnagyobb a különbség a 2017-ben és a 2018-ban adott pontszámok között (lásd 7. és 8. táblázat: 2,7015 vs 4,1791) és emiatt tűnik úgy, hogy túlságosan széttartanak az értékelők értékelései.

A táblázatok alsó soraiban található szeparációs megbízhatósági index mindkét esetben magas, a maximális 1-hez közelít (0,97 és 0,93), ami azt jelenti, hogy az értékelők megbízhatóan képesek különbséget tenni a kritériumok között, azaz csoport-szintű holdudvarhatásról nem beszélhetünk (Myford–Wolfe 2004).

Össz. erd.	Összes kitölt.	Megf. átlag	Fair (M) átlag	- Logit	Mérési hiba	Infit MnSq ZStd	Outfit MnSq ZStd	Becs. Diszk.	Korreláció Ténysl. Elv.	No.Kritérium
428	134	3.19	3.12	1.31	.17	.89 -1.0	.92 -.6	1.13	.77 .78	6 Nyelvhely.2
430	134	3.21	3.13	1.25	.17	.92 -.6	.91 -.8	1.08	.79 .78	7 Foly.2
439	134	3.28	3.19	.98	.18	1.07 .6	1.07 .6	.94	.83 .79	10 Nyelvhely.3
452	134	3.37	3.28	.58	.18	1.28 2.0	1.30 2.1	.70	.70 .79	4 Feladtm.2
453	134	3.38	3.29	.54	.18	1.14 1.1	1.17 1.2	.82	.71 .79	5 Kif.készség.2
454	134	3.39	3.30	.51	.18	1.04 .3	1.06 .4	.95	.77 .79	3 Nyelvhely.1
458	134	3.42	3.33	.39	.18	1.15 1.1	1.17 1.2	.85	.79 .80	9 Kif.készség.3
472	134	3.52	3.45	-.07	.18	.47 -4.8	.42 -5.2	1.48	.87 .80	12 Kiejtés
494	134	3.69	3.66	-.79	.18	.98 -.1	.96 -.2	1.03	.86 .81	11 Foly.1-3
497	134	3.71	3.68	-.89	.18	.86 -1.0	.87 -.9	1.12	.87 .81	2 Kif.készség.1
508	134	3.79	3.78	-1.25	.18	1.11 .9	1.17 1.2	.83	.81 .81	8 Feladtm.3
548	134	4.09	4.10	-2.56	.18	.90 -.8	.82 -1.2	1.15	.83 .81	1 Feladtm.1
469.4	134.0	3.50	3.44	.00	.18	.98 -.2	.99 -.2		.80	Mean (Count: 12)
34.4	.0	.26	.29	1.11	.00	.20 1.7	.22 1.8		.05	S.D. (Population)
35.9	.0	.27	.30	1.16	.00	.20 1.8	.23 1.9		.06	S.D. (Sample)

Model, Populn: RMSE .18 Adj (True) S.D. 1.10 Szeparáció 6.15 Rétegek 8.53 Megbízhatóság .97
 Model, Sample: RMSE .18 Adj (True) S.D. 1.15 Szeparáció 6.43 Rétegek 8.90 Megbízhatóság .98
 Model, Fixed (all same) chi-square: 457.8 d.f.: 11 szignifikancia (valószínűség): .00
 Model, Random (normal) chi-square: 10.7 d.f.: 10 szignifikancia (valószínűség): .38

1. táblázat. Kritériumok elemzése: beszédkészség 2017-2018

Össz. ered.	Összes kitölt.	Megf. átlag	Fair (M) átlag	- Logit	Mérési hiba	Infit MnSq ZStd	Outfit MnSq ZStd	Becs. Diszk.	Korreláció Ténysl. Elv.	No.Kritérium
429	134	3.20	3.18	1.14	.16	1.21 1.6	1.21 1.6	.80	.47 .62	5 Komm.érték2
460	134	3.43	3.41	.40	.15	.87 -1.0	.86 -1.1	1.12	.63 .63	7 Nyelvhely.2
461	134	3.44	3.41	.37	.15	1.43 3.0	1.43 3.0	.57	.69 .63	3 Nyelvhely.1
474	134	3.54	3.52	.07	.15	.89 -.9	.88 -.9	1.12	.69 .64	8 Alt.benyomás2
489	134	3.65	3.64	-.28	.15	.71 -2.6	.71 -2.6	1.31	.73 .64	2 Kif.készség1
492	134	3.67	3.66	-.35	.15	.88 -1.0	.88 -.9	1.11	.59 .64	6 Kif.készség2
498	134	3.72	3.71	-.49	.15	1.14 1.1	1.16 1.3	.82	.47 .64	1 Komm.érték1
513	134	3.83	3.83	-.84	.15	.86 -1.2	.84 -1.4	1.21	.78 .64	4 Alt.benyomás1
477.0	134.0	3.56	3.54	.00	.15	1.00 -.1	1.00 -.1		.63	Mean (Count: 8)
24.9	.0	.19	.19	.59	.00	.22 1.8	.23 1.8		.11	S.D. (Population)
26.6	.0	.20	.21	.63	.00	.24 1.9	.24 1.9		.11	S.D. (Sample)

Model, Populn: RMSE .15 Adj (True) S.D. .57 Szeparáció 3.70 Rétegek 5.26 Megbízhatóság.93
 Model, Sample: RMSE .15 Adj (True) S.D. .61 Szeparáció 3.97 Rétegek 5.63 Megbízhatóság.94
 Model, Fixed (all same) chi-square: 116.6 d.f.: 7 szignifikancia (valószínűség): .00
 Model, Random (normal) chi-square: 6.6 d.f.: 6 szignifikancia (valószínűség): .36

2. táblázat. Kritériumok elemzése: íráskészség 2017-2018

2. Feladatonként a beszéd-készség és az íráskészség teljesítményeket mely kritériumok mentén értékelték a legszigorúbban és a legenyhébben a vizsgáztatók 2017-ben és 2018-ban?

A második kutatási kérdésre a klasszikus tesztelmélet alapján kerestük a választ, ugyanis a kérdés megválaszolásához modern tesztelméleti módszerekre nem volt szükség. A vizsgálatot az Excel táblázatkezelő és az SPSS programcsomag segítségével végeztük.

A 3., 4., 7. és 8. táblázatokban az első sorban az év mellett az egyes értékelési szempontok szerepelnek rövidítve (a rövidítések jegyzékét lásd a függelékben). A rövidítések melletti szám az adott feladat sorszámára utal, tehát például a *Kif.2* a 2., azaz az önálló témakifejtés feladat *kifejezőkészség* kritériumát jelenti. A második sor a feladatot értékelő vizsgáztatók számát jelenti (67). A többi sorban a leíró statisztika legfontosabb mutatói találhatók.

A beszéd-készség táblázatokban (3. és 4.) szereplő átlagokból az látszik, hogy amennyiben feladatonként tekintjük a számokat, akkor a két év során a hat feladatból öt alkalommal a *nyelvhelyesség* szempontot értékelték a legszigorúbban a vizsgáztatók, ami azért érdekes mert a két vizsgáztatói teljesítmény színvonala nagyban különbözött egymástól (2017: 37 vs. 2018: 50 nyerspont a megszerezhető 60-ból), ám ez mégsem változtatott a kritériumok megítélésén. Az elvégzett páros Wilcoxon-próbák (amelynek során a *nyelvhelyesség* és az adott feladathoz tartozó többi kritérium átlagai közötti különbség szignifikanciaszintjét vizsgáltuk meg) azt mutatták, hogy a *nyelvhelyesség* szigorúsága a 13 esetből nyolcszor valós különbséget takar ($p < 0,05$), azaz az eredmény nem a véletlen műve (5. és 6. táblázat).

2017	Fel.1	Kif.1	Ny.1	Fel.2	Kif.2	Ny.2	Foly.2	Fel.3	Kif.3	Ny.3	Foly.1-3	Kiej.	Összes
N	67	67	67	67	67	67	67	67	67	67	67	67	67
Átlag	3,462	2,955	2,836	2,880	2,910	2,657	2,642	3,134	2,761	2,582	2,955	2,97	34,75
Módusz	3	3	3	3	3	3	3	3	3	3	3	3	33
Medián	3	3	3	3	3	3	3	3	3	3	3	3	34
Szórás	0,502	0,405	0,539	0,564	0,570	0,509	0,513	0,6	0,525	0,497	0,474	0,171	3,221

3. táblázat. Leíró statisztika: beszéd-készség 2017

2018	Fel.1	Kif.1	Ny.1	Fel.2	Kif.2	Ny.2	Foly.2	Fel.3	Kif.3	Ny.3	Foly.1-3	Kiej.	Összes
N	67	67	67	67	67	67	67	67	67	67	67	67	67
Átlag	4,716	4,462	3,940	3,865	3,850	3,731	3,776	4,447	4,074	3,970	4,417	4,074	49,32
Módusz	5	4	4	4	4	4	4	4	4	4	4	4	48
Medián	5	4	4	4	4	4	4	4	4	4	4	4	49
Szórás	0,454	0,531	0,518	0,625	0,500	0,479	0,572	0,530	0,531	0,626	0,554	0,470	4,020

4. táblázat. Leíró statisztika: beszéd-készség 2018

2017	Ny.1 – Fel.1	Ny.1 – Kif.1	Kif.3 – Ny.3	Fel.3 – Ny.3	Foly.13 – Ny.3
Z	-6,410 ^a	-1,569 ^b	-2,191 ^b	-5,778 ^b	-4,642 ^b
Asymp. Sig. (2-tailed)	,000	,117	,028	,000	,000

5. táblázat. Beszédkészség különbözőségvizsgálat 2017: Wilcoxon-próba

2018	Ny.1 – Fel.1	Kif.1 – Ny.1	Kif.2 – Ny.2	Fel.2 – Ny.2	Ny.2 – Foly.2	Kif.3 – Ny.3	Fel.3 – Ny.3	Foly.13 – Ny.3
Z	-6,761 ^a	-5,840 ^b	-1,569 ^b	-1,567 ^b	-,577 ^a	-1,400 ^b	-5,154 ^b	-4,727 ^b
Asymp. Sig. (2-tailed)	,000	,000	,117	,117	,564	,162	,000	,000

6. táblázat. Beszédkészség különbözőségvizsgálat 2018: Wilcoxon-próba

Az íráskészség táblázatokban (7. és 8.) szereplő átlagokból kiolvasható, hogy amennyiben feladatonként tekintjük a számokat, akkor a két év során a négy feladtból három alkalommal a *kommunikatív érték* szempontot értékelték a legszigorúbban a vizsgáztatók, ami azért figyelemre méltó, mert a két vizsgázói teljesítmény színvonala nagyban különbözött egymástól (2017: 28 vs. 2018: 34 pont a megszerezhető 40-ből), ám ez mégsem változtatott a kritériumok megítélésén. A *nyelvhelyesség* kritériumot csupán a 2017-es 1. feladat esetén értékelték a legszigorúbban. Megjegyzendő, hogy ebben a feladatban éppen a *kommunikatív érték* lett a legenyhébben értékelt szempont, szöges ellentétben a másik három feladattal. Az itt is elvégzett Wilcoxon-próbák, amelynek során a *kommunikatív érték* és az adott feladathoz tartozó többi kritérium átlagai közötti különbség szignifikanciaszintjét vizsgáltuk meg, azt mutatták, hogy a *kommunikatív érték* szigorúsága a kilenc esetből nyolcszor valós különbséget takar ($p < 0,05$), azaz az eredmény nem a véletlen műve (9. és 10. táblázat).

2017	Komm.ért.1	Kif.1	Nyelvh.1	Ált.b.1	Komm.ért.2	Kif.2	Nyelvh.2	Ált.b.2	Összes
N	67	67	67	67	67	67	67	67	67
Átlag	3,4925	3,1493	2,7015	3,2239	2,9851	3,4776	3,1045	3,2537	25,3881
Módusz	4	3	3	3	3	4	3	3	23
Medián	4	3	3	3	3	4	3	3	25
Szórás	0,5871	0,4353	0,4928	0,4546	0,6852	0,6120	0,5261	0,6116	2,4738

7. táblázat. Leíró statisztika: íráskészség 2017

2018	Komm.ért.1	Kif.1	Nyelvh.1	Ált.b.1	Komm.ért.2	Kif.2	Nyelvh.2	Ált.b.2	Összes
N	67	67	67	67	67	67	67	67	67
Átlag	3,9403	4,1493	4,1791	4,4328	3,4179	3,8657	3,7612	3,8209	31,5672
Módusz	4	4	4	5	3	4	4	4	30
Medián	4	4	4	4	3	4	4	4	32
Szórás	0,6715	0,5295	0,5484	0,6086	0,6068	0,6251	0,6534	0,7370	3,0213

8. táblázat. Leíró statisztika: íráskészség 2018

2017	Kif.2 – Komm.ért.2	Ny.2 – Komm.ért.2	Ált.b.2 – Komm.ért.2
Z	-4,814 ^a	-1,241 ^a	-3,000 ^a
Asymp. Sig. (2-tailed)	,000	,215	,003

9. táblázat: Íráskészség különbözőségvizsgálat 2017: Wilcoxon-próba

2018	Kif.1 – Komm.ért.1	Ny.1 – Komm.ért.1	Ált.b.1 – Komm.ért.1	Kif.2 – Komm.ért.2	Ny.2 – Komm.ért.2	Ált.b.2 – Komm.ért.2
Z	-2,556 ^a	-2,317 ^a	-4,907 ^a	-4,328 ^a	-3,450 ^a	-3,430 ^a
Asymp. Sig. (2-tailed)	,011	,021	,000	,000	,001	,001

10. táblázat: Íráskészség különbözőségvizsgálat 2018: Wilcoxon-próba

Az egyik legnagyobb különbség tehát a beszéd-készség és az íráskészség feladatok értékelése között az, hogy a beszéd-készség esetén, hasonlóan McNamara (1990) és Corrigan (2007) eredményeihez, többségben a *nyelvhelyesség* a legszigorúbban értékelt kritérium, az íráskészség esetén pedig a *kommunikatív érték* a domináns, ami tulajdonképpen a beszéd-készség skála *feladatmegoldás* kritériumának feleltethető meg. Ennek a magyarázata az lehet, hogy a vizsgáztatók azokon a területeken szigorúbbak, amelyeket a legjobban ismernek, mert itt veszik észre legkönnyebben a hibákat (Mike Linacre, e-mail-kommunikáció, 2019), ami a beszéd-készség esetén a *nyelvhelyesség*. Érdekes viszont, hogy az íráskészség értékelésekor, ugyanezen vizsgáztatók esetén, ez a magyarázat nem állja meg a helyét, aminek talán az az oka, hogy annyira részletesen kidolgozott az íráskészség-értékelési útmutató a tartalmi pontok kifejtettségét illetően, ami a *kommunikatív érték* kritérium megítélésének elsődleges összetevője, hogy a vizsgáztatók még szigorúbban, illetve még szabálykövetőbben használják ezt a skálát, mint az általuk jól ismert *nyelvhelyesség* skálát. A jelenség magyarázata lehet továbbá az is, hogy az íráskészség esetén a tartalom szempontjából magasabb elvárásokat támasztanak az értékelők a vizsgázókkal szemben, mint a beszéd-készség esetén.

3. Milyen összefüggések és különbségek mutathatók ki a beszéd-készség és az íráskészség értékelésére használt skálák alkalmazása között?

A harmadik kérdésre, a másodikhoz hasonlóan, a klasszikus tesztelmélet alapján kerestük a választ, az SPSS programcsomag felhasználásával. A korrelációelemzéssel azt kívántuk megvizsgálni, hogy a vizsgáztatók mennyire azonos tendencia szerint értékelik a beszéd-készség- és az íráskészség-teljesítményeket. Mivel a tesztpontszámok nem intervallum, hanem rangsor skálán helyezkednek el, ezért a Spearman-féle rangkorrelációs együtthatót számoltuk ki, amelynek az a lényege, hogy nem az adott pontszámok, hanem a sorban elfoglalt hely szerint állapítja meg az összefüggést két adatsor között.

A beszéd- és az íráskészség értékelések korrelációs együtthatója negatív előjelű ($r = -0,074$, $p = 0,554$), ami azt a tendenciát rajzolta ki, hogy a vizsgáztatók alapvetően ellentétesen ítélték meg a beszéd- és írásteljesítményeket, azaz aki az egyiket szigorúbban értékelte, az a másikat engedékenyebben. Ez az együttható azonban nem szignifikáns, ami azt jelenti, hogy az eredmény a véletlen műve is lehet, tehát további vizsgálatok szükségesek annak megállapítására, hogy ez a jelenség milyen mértékben valós. A vizsgált csoportról azonban megerősítette azt a korábbi elemzések által is előrejelzett képet, hogy különbözően használják a két skálát, mások a prioritások egyik és másik esetben.

A beszéd- és az íráskészség adatsorral kapcsolatban különbözőségvizsgálatot is végeztünk azzal a céllal, hogy kiderüljön, hogy a két adatsor, azaz a teljesítményekre adott értékelések átlagai (százalékban kifejezve: beszéd- és íráskészség: 70,21%; íráskészség: 71,36%) milyen mértékben térnek el egymástól. A páros t-próba az átlagok között nem mutatott ki szignifikáns különbséget, azaz az értékelők átlageredményei nagyon hasonlóak voltak egymáshoz ($t = -1,276$, $df = 66$, $p = 0,206$).

Az összefüggés- és a különbözőségvizsgálatok tehát nagyon érdekes eredményeket hoztak: a vizsgáztatók ugyan ellentétes módon értékelték a beszéd- és íráskészség vizsgáztatói teljesítményeket, az átlageredményeket tekintve azonban nagyon hasonlóan értékelték. Vizsgálatunk célját tekintve ez azt jelenti, hogy az értékelők a skálák egyes kritériumait különböző módon, a teljes skálát viszont nagyon hasonló módon használták. Megjegyzendő azonban, hogy mivel az összefüggés és különbözőségvizsgálatok számítása is nyerspontokkal történt, ezért a különbségek nem csupán a skálahasználatra, hanem a vizsgáztatói teljesítményekre is vonatkozhatnak.

Összefoglalás

Jelen tanulmány arra tett kísérletet, hogy a BME Nyelvvizsgaközpont 2017. és 2018. évi vizsgáztató-továbbképzése során összegyűlt adatok segítségével összehasonlítsa, hogyan használják ugyanazon angol nyelvi vizsgáztatók az íráskészség és a beszéd- és íráskészség méréséhez készült értékelési skálákat két-két B2 szintű, de különböző nyelvi színvonalú nyelvvizsgáztatói teljesítmény értékelése során.

Az eredmények azt mutatták, hogy kevés kivétellel minden kritérium esetén, a beszéd- és íráskészségnél nagyobb, az íráskészségnél kisebb mértékben, a vizsgáztatók engedékenyebben értékelték a modern tesztelméleti alapon kialakított modell elvárásainál. Mind a beszéd- és íráskészség, mind pedig az íráskészség esetén az első feladatot, azaz az ahhoz tartozó kritériumokat, enyhébben értékelték, mint a többi feladathoz tartozókat, bár ez az adat még további megerősítést igényel. Sikerült kimutatni, hogy a skálák egyes kritériumai az adott készség alkotórészei. Összességében az értékelők megbízhatóan tudtak különbséget tenni az értékelési szempontok között, azaz csoportszinten holdudvarhatásról vagy közepes pontszámra törekvés tendenciáról nem beszélhetünk.

Feladatonként tekintve a kritériumokkal szembeni szigorúság kérdésében érdekes megállapítást tehattünk. A beszéd-készség esetén a *nyelv-helyesség*, az íráskészség esetén pedig a *kommunikatív érték* bizonyult a legszigorúbban értékelt kritériumnak. A *nyelv-helyesség* primátusát azzal magyaráztuk, hogy a vizsgáztatók azokon a területeken szigorúbbak, amelyeket a legjobban ismernek, az íráskészség esetén azonban a *kommunikatív érték* dominanciája felülírta ezt a szabályt, amit az értékelési útmutató rendkívüli részletezettségével, illetve a tartalommal kapcsolatos elvárások különbözőségével magyaráztunk. További kutatások szükségesek annak feltárására, hogy ez a jelenség általánosítható-e. Érdekes azt is megvizsgálni, hogy más nyelvek esetében hasonló eredmények születnének-e.

A vizsgálatok során az is kiderült, hogy a vizsgált értékelői csoport ugyan ellentétes módon értékelte a beszéd-készség- és az íráskészség-vizsgálói teljesítményeket, azaz aki az egyiket szigorúbban értékelte, az a másikat engedékenyebben, az átlageredményeket tekintve azonban nagyon hasonlóan. Vizsgálatunk fókuszára nézve ez azt jelenti, hogy az értékelők a skálák egyes kritériumait különböző módon, a teljes skálát viszont nagyon hasonló módon használták.

A kutatás eredményei fontos információkat nyújthatnak a vizsgaközpont számára. A beszéd-készségre koncentráló vizsgáztató-továbbképzéseken az első feladat a másik kettőhöz képest történő enyhébb megítélésére érdemes fókuszálni. Ki lehet térni továbbá a *feladatmegoldás* és a *nyelv-helyesség* szempontok megítélésére is, az egyiknél ugyanis a túlzott engedékenység, a másikon pedig az esetleges túlzott szigor jelenthet problémát. Elképzelhető az is, hogy az értékelési kritériumokon kell további finomításokat végezni ezen két kritérium esetén.

Az íráskészség esetén a jövőbeni továbbképzéseken azt érdemes megvizsgálni, hogy miért van különbség az első és a második feladat megítélésében, hiszen kiderült, hogy a vizsgázók teljesítményét a viszonylag új második feladaton szigorúbban ítélik meg a vizsgáztatók, mint az első feladat esetében. Ezen készség esetén is elképzelhető, hogy az értékelési skálát felülvizsgálat alá kell vetni, mert lehetséges, hogy nem egyformán alkalmas mindkét feladat értékelésére.

Az eredmények értelmezésekor két körülményt nem szabad elfelejteni. Az egyik, hogy az elemzett értékelést nem valódi vizsgán, hanem egy továbbképzés keretében végezték a résztvevők, amelyek tétje nem hasonlítható egymáshoz. A másik pedig az, hogy a vizsgálat online válaszadáson alapult, ami elképzelhető, hogy gondot okozott a papír-toll alapú vizsgáztatáshoz szokott vizsgáztatók számára (Myford–Wolfe 2003). A vizsgálatokat a fent említett irányok mellett vizsgáztatókkal készült interjúkkal és hangos gondolkodtatásos (*think-aloud protocol*) adatfelvételi eljárással is érdemes ki-gészíteni.

IRODALOM

- Benke Eszter (2008): Rater and rating scale interaction in the validation of the assessment of writing performance. In: Majoros Pál (szerk.) *BGF Tudományos Évkönyv 2007: Reformok útján*. Budapest: Budapesti Gazdasági Főiskola, 327–335.
- Bond, T. – Fox, C. (2015): *Applying the Rasch model: Fundamental measurement in the human sciences*. Routledge: New York and London.
- Corrigan, M. (2007): *Seminar to calibrate examples of spoken performance: Report on the analysis of rating data*. Strasbourg, France: Council of Europe/Language Policy Division.
- Eckes, T. (2015): *Introduction to many-facet Rasch measurement: Analyzing and evaluating rater-mediated assessments*. Frankfurt am Main: Peter Lang.
- Eckes, T. (2019): Many-facet Rasch measurement: Implications for rater-mediated language assessment. In: Aryadoust, V. – Raquel, M. (eds.) *Quantitative data analysis for language assessment: Volume I: Fundamental techniques*. London and New York: Routledge, 153–175.
- Eckes, T. (2009): Many-facet Rasch measurement. In: Takala, S. (ed.) *Reference supplement to the manual for relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (Section H)*. Strasbourg, France: Council of Europe/Language Policy Division.
- Kiszely Zoltán (2018): Egy online vizsgáztató-továbbképzés tapasztalatai. *Modern Nyelvoktatás*, 24/1, 3–18.
- Kuiken, F. – Vedder, I. (2014): Raters' decisions, rating procedures and rating scales. *Language Testing*, 31/3, 279–284.
- Linacre, J. M. (2003): Rasch Power Analysis: Size vs. Significance: Standardized Chi-Square Fit Statistic. *Rasch Measurement Transactions*, 17/1, 918, <https://www.rasch.org/rmt/rmt171n.htm>
- Linacre, J. M. (2018): *A user's guide to FACETS: Rasch-model computer programs. Program manual*. Chicago, IL: Winsteps.com. Letöltve www.winsteps.com/facets.htm
- Lumley, T. – McNamara, T. F. (1995): Rater characteristics and rater bias: Implications for training. *Language Testing*, 12/1, 54–71.
- McNamara, T. (1990): Item response theory and the validation of an ESP test for health professionals. *Language Testing*, 7/1, 52–75.
- McNamara, T. (1996): *Measuring second language performance*. Longman: Harlow.
- Molnár Gyöngyvér (2013): *A Rasch-modell alkalmazási lehetőségei az empirikus kutatások gyakorlatában*. Gondolat: Budapest.
- Myford, C. M. – Wolfe, E. W. (2003): Detecting and measuring rater effects using many-facet Rasch measurement: Part I. *Journal of Applied Measurement*, 4, 386–422.
- Myford, C. M. – Wolfe, E. W. (2004): Detecting and measuring rater effects using many-facet Rasch measurement: Part II. *Journal of Applied Measurement*, 5, 189–227.
- Wind, S. A. – Peterson, M. E. (2017): A systematic review of methods for evaluating rating quality in language assessment. *Language Testing*, 35/2, 161–192.
- Wright, B. D. – Linacre, J. M. (1994): Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8/3, 369–370. <https://rasch.org/rmt/rmt83b.htm>

FÜGGELÉK

Rövidítések jegyzéke:

Fel.1	Feladatmegoldás 1	Foly.1-3	Folyamatosság és koherencia 1, 3
Kif.1	Kifejezőkészség 1	Kiej.	Kiejtés
Ny.1	Nyelvhelyesség 1	Komm.ért.1	Kommunikatív érték 1
Fel.2	Feladatmegoldás 2	Kif.1	Kifejezőkészség 1
Kif.2	Kifejezőkészség 2	Nyelvh.1	Nyelvhelyesség 1
Ny.2	Nyelvhelyesség 2	Ált.b.1	Általános benyomás 1
Foly.2	Folyamatosság és koherencia 2	Komm.ért.2	Kommunikatív érték 2
Fel.3	Feladatmegoldás 3	Kif.12	Kifejezőkészség 2
Kif.3	Kifejezőkészség 3	Nyelvh.2	Nyelvhelyesség 2
Ny.3	Nyelvhelyesség 3	Ált.b.2	Általános benyomás 2

CSÉPES ILDIKÓ

Hogyan fejleszthető a tanárok nyelvtudásmérési és értékelési műveltsége? Az Erasmus+ TALE projekt tanulságai

How to facilitate language teachers' assessment literacy? Lessons from the Erasmus+ TALE Project

The attitudes, knowledge and skills related to efficient assessment practices have been recently summarized under the umbrella term "Language Assessment Literacy" (LAL). There has been considerable research in the field of LAL with a focus on theoretical conceptualisations, needs analyses for teachers, etc. However, the needs of teachers and learners have not yet been established simultaneously with a view to designing open-access training material for teachers. With funding from the European Union, the project 'Teachers' Assessment Literacy Enhancement (TALE)' set out to create an online course in LAL. In the first phase of the project, a needs analysis survey was carried out involving 852 English language teachers and 1788 language learners in Cyprus, Germany, Greece and Hungary. Respondents were asked about their assessment practices, confidence in using assessments, and LAL training needs. The results revealed that alternative, student-centred forms of assessment are employed much less frequently than traditional forms of assessment.

Keywords: language assessment literacy, needs analysis, teaching, feedback, training needs

Bevezetés

Az elmúlt két évtizedben a nyelvtudásmérési szakma egyre nagyobb figyelmet fordított az osztálytermi értékelés kutatására (Rea–Dickins 2001, Davison–Leung 2009, Poehner 2009, Hill–McNamara 2012, Tzagari–Vogt 2014), és ennek kapcsán arra is, hogy meghatározza, a nyelvtanároknak milyen elméleti és gyakorlati tudásra van szükségük a mérés és értékelés területén. Ennek a tudás és képesség területnek az elnevezéseként manapság a nyelvtudásmérési műveltség (*Language Assessment Literacy – LAL*) kifejezést használják, melyet Inbar-Lourie (2008: 385) vezetett be.

Az angoltanárképzés kontextusában a nyelvtudásmérési műveltség képzési igényeit többen is vizsgálták már az elmúlt 15 évben (pl. Hasselgreen–Carlsen–Helness 2004, Kvasova–Kvavitska 2014, Vogt–Tzagari 2014), és a kapott eredmények azt mutatták, hogy a válaszadók döntő többsége szükségét érzi annak, hogy nyelvtudásmérési műveltségét fejleszteni tudja. Az igényfelmérések azonban főleg a tanárok meglátásaira szorítkoztak, és a tanulók, akik az osztálytermi értékelésnek szintén fontos résztvevői,

eddig kimaradtak a vizsgálatokból. Kutatásunk újszerűsége abban rejlik, hogy mindkét érintett meglátásait egyszerre vizsgálja, és így a kapott eredmények összevethetőek.

A kutatás az Európai Unió támogatásával (Erasmus+, Innovációt és bevált gyakorlatok megosztását célzó együttműködések, Stratégiai partnerség a közoktatásban, KA2) valósult meg a *Teachers' Assessment Literacy Enhancement (TALE)* [Tanárok nyelvtudásmérési műveltségének a fejlesztése] elnevezésű projekt keretében. A résztvevők az Egyesült Királyság, Ciprus, Görögország, Magyarország, Németország és Norvégia mérési és értékelési, valamint tanárképzésben jártas szakemberei voltak.

Kutatási kérdések

Jelen tanulmányunk arról ad átfogó képet, hogy négy európai országban (Ciprus, Görögország, Magyarország, Németország) a megkérdezett, többségében angolt tanító tanár és angolul tanuló diák milyen nyelvtudásmérési szükségletekről számolt be. Célkitűzésünk volt továbbá, hogy feltérképezzük, mi jellemzi az osztálytermi gyakorlatot a nyelvtudásmérésre és az értékelésre vonatkozóan. A következő főbb kutatási kérdésekre kerestük a választ:

1. Milyen az osztálytermi értékelés jelenlegi gyakorlata?
 - Mely nyelvi készségeket/elemeket értékelik a tanárok?
 - Milyen visszajelzési formákat használnak a tanárok?
2. Milyen a tanárok értékelési profilja?
 - Mely területeken érzik magukat a tanárok magabiztosnak, és mely területeken nem?
 - Mely osztálytermi értékeléshez kapcsolódó területen szeretnék magukat (tovább)képezni?
3. A tanulók meglátásai hogyan viszonyulnak a tanárok állításaihoz?

A következőkben röviden áttekintjük a pedagógiai értékeléshez kapcsolódó tanári kompetencia alapkövetelményeit, valamint néhány, az előbbiekre épülő nyelvtudásmérési műveltségértelmezést. Ezt követően bemutatjuk a TALE-projekt keretében végzett kutatás vizsgálati módszerét, az adatközlőket és a kutatási kérdéseinkre kapott válaszokat.

Szakirodalmi áttekintés

A pedagógiai mérés angol nyelvű szakirodalmában már az 1970-es és 80-as években körvonalazták, hogy a tanároknak milyen készségekkel és tudáselemekkel kell rendelkezniük az értékelés területén (Schafer 1991, Stiggins 1991). Azonban az első átfogó leírás az elvárt követelményekről csak a kilencvenes években jelent meg *Standards for Teacher Competence in Educational Assessment of Students* [A tanári kompetencia követelményei a pedagógiai értékelésben] címmel a National Council on Measurement in Education és National Education Association (1990) kiadványaként. A követelmények

listája hét főbb területet jelölt ki, melyeket a tanárképzési programoknak ajánlott volt a képzés során fejleszteniük. A pedagógiai értékelés elemei a következők voltak: a tanároknak képesnek kell lenniük arra, hogy

1. kiválasszák a megfelelő értékelési módszereket, melyek összhangban állnak a tanítási célkitűzésekkel,
2. megtervezzék és létrehozzák a megfelelő értékelési módszereket,
3. megfelelően alkalmazzák az akár máshonnan átvett, akár saját maguk által kifejlesztett értékelési eszközt, és annak eredményét megfelelően értelmezzék a saját tanításukban,
4. az értékelés során kapott eredményeket fel tudják használni a diákok egyéni fejlődésének, a tanítás megtervezésének, a kurrikulum fejlesztésének és az iskola jobbításának érdekében,
5. érvényes, teljesítményértékelésen alapuló osztályzási formát alakítsanak ki,
6. az értékelés során kapott eredményeket megfelelően tudják közvetíteni a diákok, a szülők és más külső fél számára,
7. felismerjék az értékelés azon formáját vagy felhasználását, mely nem etikus, jogsértő vagy egyéb okoknál fogva helytelen.

Az elmúlt évtizedekben a nyelvtudásmérési műveltség mint a pedagógiai értékeléshez kapcsolódó értékelési műveltség egyik sajátos területe, egyre erőteljesebben jelent meg a szakirodalomban. Fulcher (2012) a tartalmi elemeket a mérési szakma történelmi, politikai és etikai kérdéseinek az ismeretével egészítette ki, azaz hogy az idegen nyelvi méréshez szükséges ismereteket, készségeket, valamint fogalmakat és elveket tágabb kontextusban kell értelmezni ahhoz, hogy a mérésnek a társadalomra vagy az egyénre gyakorolt hatását megértsük. O’Loughlin pedig a kritikus szemlélet kívánalmát építette be a nyelvtudásmérési műveltség definíciójába: „A tesztfejlesztéshez kapcsolódó azon készségek, melyek elengedhetetlenek ahhoz, hogy a kapott teszteredményeket megfelelően értelmezni és hasznosítani tudjuk, valamint egy teszt minőségének megítélése során kritikusan vizsgáljuk az értékelésnek a társadalomban betöltött szerepét és funkcióját”¹ (2013: 363). Vogt–Tzagari (2014) O’Loughlinhoz hasonlóan a kritikus szemlélet fontosságát emelik ki definíciójukban, különösen az értékelési eljárás mint folyamat és annak eredményét illetően.

A nyelvtudásmérési műveltség fogalmi és tartalmi meghatározásának gazdagodásához még egy jelentős szemléletváltást kell megemlítenünk. A nyelvtanulók eredményességének megítélése már nemcsak a teljesítmények összevethető mérésén alapszik (*assessment of learning*, azaz ’a tanulás értékelése’), hanem olyan értékelési formán is alapulhat, mely a tanulók előrehaladását alternatív értékelési módszerek segítségével közvetlenül,

¹ Az idézet a szerző saját fordítása

egyénre szabottan támogatják. Az értékelésnek ezt a fajtáját gyakran nevezik 'a tanulás érdekében történő értékelésnek' (*assessment for learning*, Black–Wiliam 1998). A tanulói portfólió, a tanulói önértékelés és tanuló társ általi értékelés ilyen típusú értékelési formáknak tekinthetők. Earl és Katz (2006) megkülönbözteti még az 'értékelés mint tanulás' fogalmát (*assessment as learning*), mely során a tanulók értékelésre vonatkozó metakognitív tudatosságát fejleszthetjük.

A TALE-projekt

Mivel jelen kutatásunk szorosan kapcsolódik a TALE-projekthez, ezért fontosnak tartjuk a projekt rövid ismertetését. A projekt résztvevőinek legfőbb célkitűzése az idegen nyelvi mérés és értékelés hatékonyságának növelése volt, melyhez egy szabad hozzáférésű, interaktív, online képzési platformot hoztak létre. A hároméves futamidejű projekt (2015. szeptember–2018. augusztus) három főbb fázisból állt:

1. Szükségletelemzés angoltanárok/tanárjelöltek és tanulók részvételével
2. Online tréning: kurzustervezés és kipróbálás
3. *A Nyelvi tesztelés és értékelés kézikönyvének összeállítása (Handbook of Language Testing and Assessment)*

A TALE-projekt arra vállalkozott, hogy

- hatékony nemzetközi együttműködést hozzon létre tanárok, tanárjelöltek és tanulók között, hogy a nyelvtanulás sikerességén javítani lehessen;
- kevésbé alkalmazott értékelési kompetenciák (alternatív, tanulóközpontú értékelési módszerek alkalmazása) fejlesztésére is sort kerítsen;
- a nyelvtudásmérés és értékelés szakmaiságát növelje Európában (és azon kívül);
- a nyelvtudásmérés és értékelés területén a jó gyakorlatot és innovációt megossza.

A következőkben az angoltanárokra/tanárjelöltekre és tanulókra vonatkozó szükségletelemzés eredményeiről számolunk be.

A kutatás módszere

A tanári kérdőív

A kérdőív ahhoz kívánt segítséget nyújtani, hogy a feltárt valós igények alapján jöhesse létre az online képzés a TALE-projekt keretein belül. A projekt első évében a kérdőív tartalmi kidolgozásakor figyelembe vettük a fentebb említett *Standards for Teacher Competence in Educational Assessment of Students* (1990), továbbá a *Joint Committee on Standards for Education* (2015) útmutatásait és a Közös európai referenciakeret (2002) nyelvtudásmérésre vonatkozó egyes ajánlásait is. A kutatási minta a leendő és már gyakorló nyelvtanárokat egyaránt magában foglalta, mivel a TALE

online képzés is a lehető legszélesebb célközönségnek készült. A képzés moduláris jellegénél fogva hasznos tudást kívánt nyújtani a tapasztalt és kevésbé tapasztalt résztvevőknek egyaránt. A kérdőív három főbb területet ölelt fel:

- A válaszadók háttere: nem, életkor, tanítási tapasztalat, végzettség, tanított nyelv(ek) és korábbi képzés a nyelvtudásmérés és értékelés területén (a válaszadás formája: a megadott lehetséges válaszokból kellett bejelölni a megfelelőt).
- A jelenlegi osztálytermi gyakorlat feltérképezése az értékelésre vonatkozóan: milyen készségeket és nyelvi elemeket értékelnek; milyen tanulóknak szánt visszajelzést használnak; milyen értékelési formákat és módszereket alkalmaznak, és milyen gyakran használják azokat (a válaszadás formája: a megadott listából kellett bejelölni a releváns opciókat, illetve egy 4 fokozatú Likert-skálán kellett megjelölni a kiválasztott értékelési módszerek gyakoriságát).
- Mennyire magabiztosak a tanárok az idegen nyelvi tesztelés és értékelés egyes területeinek alkalmazásában (a válaszadás formája: egy 4 fokozatú skálán kellett bejelölni a magabiztosságot a megadott lista elemeire vonatkozóan), és mely területeken szeretnének továbbképzésben részt venni (a válaszadás formája: be kellett jelölni a kiválasztott listaelemeket).

Az angol nyelvű kérdőívet kismintás kipróbálás után a TALE-projekt résztvevői véglegesítették (a teljes kérdőív elérhető itt: <http://taleproject.eu>) A leíró kutatás eredményei a válaszadók osztálytermi gyakorlatát tükrözik tanári szemszögből.

A tanulói kérdőív

A tanulóknak szánt kérdőívben igyekeztünk a tanári kérdőívben szereplő kérdéseket a diákok szemszögéből is vizsgálni. Ennek megfelelően a kérdések az osztálytermi értékelés elemeire, azok formáira és gyakoriságára irányultak. Kíváncsiak voltunk arra is, hogy milyen formában kapnak visszajelzést a diákok az értékelés során. Mivel a nyelvtanítás és értékelés során fontos figyelembe venni a tanulói igényeket, ezért a diákoknak arról is nyilatkozniuk kellett, hogy szerintük mely értékelési formák segíti őket leginkább abban, hogy az angoltudásuk fejlődhessen. A tanulói kérdőívet kismintás kipróbálás után a TALE-projekt résztvevői véglegesítették, ezt minden vizsgált országban a diákok anyanyelvére lefordítottunk, hogy a válaszadást az esetleges idegen nyelvi korlátok ne befolyásolják. A leíró kutatás eredményei a válaszadók osztálytermi gyakorlatát tükrözik tanulói szemszögből (a teljes angol nyelvű kérdőív elérhető itt: <http://taleproject.eu>).

Adatfelvétel és elemzés

A tanári kérdőívek kitöltése önkéntes volt, és az adatfelvétel online felületen történt egy meghatározott időintervallum (kb. 1 hónap) alatt 2016 tavaszán. A tanulók bevonása a felmérésbe kiskorúak esetében írásos szülői beleegyezés mellett történt. A tanulói

adatfelvételt papíralapú kérdőívek segítségével ugyanabban az időszámban, de tanórai keretek között végeztük. A tanulói válaszokat manuálisan vittük gépre az Excel programban. Az összesített adatokat leíró és következtető statisztikai módszerekkel elemeztük, melyhez az SPSS 20. verzióját használtuk.

A következőkben a tanári és tanulói kérdőíves válaszok alapján formálódó legfontosabb tendenciákat mutatjuk be az összes válaszadóra vonatkozóan, illetve a magyarországi helyzetet külön is jellemezzük.

A kérdőívek kitöltőinek általános jellemzése

A 852 tanári kérdőívet kitöltő angoltanár/tanárjelölt 87%-a nő volt, mely feltételezésünk szerint jól reprezentálhatja a tanári pályán a jelenlegi helyzetet, azaz hogy a tanárok döntő többsége nő. A válaszadók közel fele ciprusi (404 fő/47%), a többi résztvevő három ország arányosabban képviseltette magát: Görögország (91 fő/11%), Magyarország (230 fő/27%), Németország (127 fő/15%). Az válaszadók közötti aránytalanság annak tudható be, hogy a vizsgált országokban igencsak eltérő volt a részvételi hajlandóság. A válaszadók tanítási tapasztalatát tekintve nagyjából arányosan vannak képviselve az eltérő korosztályok és az eltérő tanítási tapasztalattal rendelkező válaszadók (lásd *1. táblázat*).

Életkor	összes válaszadó	magyarországi válaszadó
25 alatti	18%	26%
26–35	16%	30%
36–45	19%	27%
46–55	29%	15%
56+	18%	2%
Tanítási tapasztalat		
0 év (tanárjelölt)	22%	33%
1–5 év	12%	18%
6–10 év	13,5%	14%
11–15 év	9,5%	8%
15+	43%	27%

1. táblázat. A válaszadók életkora és tanítási tapasztalata

Az összes válaszadó tanár döntő többsége (98%) angol nyelvet tanít elsősorban, bár a magyarországi válaszadók 21%-a másik idegen nyelvet is tanít (döntően német nyelvet). A nyelvtudásmérésben való jártasság feltérképezéséhez a válaszadóknak arra is kellett válaszolniuk, hogy korábban vettek-e részt (tovább)képzésben az idegen nyelvi értékelést és mérést illetően. A négy országban megkérdezettek 63% azt állította, hogy

részt vett valamilyen képzésben az adott szakterületen (a képzés jellegét nem definiáltuk a kérdőívben). Magyarországi viszonylatban ez az adat még kedvezőbb képet mutat, hiszen a válaszadók 69%-a állította, hogy képzésben részesült. A képzés eredményessége annak fényében lesz értelmezhető, hogy milyen képet kapunk a tanárok jelenlegi értékelési gyakorlatáról és magabiztosságáról.

A tanulói kérdőívet összesen 1788 tanuló töltötte ki a négy részt vevő országban. A nemek aránya tekintetében arányosnak mondható a megoszlás, hiszen 50–50% a fiúk és a lányok aránya. A tanulók életkorát tekintve 10 és 20 év közöttiek voltak a válaszadók: a döntő többség (47%) a 13–17 év közötti korosztályt képviselte, míg a 10–12 év közöttiek aránya 21%, a 16–18 év közöttieké pedig 28% volt. A négy országot hasonló arányban képviselték a tanulók, mint a tanárok: Ciprus adta a tanulók 51%-át, míg a többi ország egyenként 16 vagy 17%-ot (az aránytalanság ismételten a részvételi hajlandóságnak tudható be).

A Magyarországon megkérdezett diákok összetétele némiképp eltérő volt, mert itt a válaszadók 75% a 13–17 évesek köréből került ki, 13 év alatti válaszadó egyáltalán nem volt. Ennek az volt az oka, hogy csak középiskolákban történt adatfelvétel, viszont a kiválasztott iskolákkal igyekeztünk minél szélesebb középiskolai tanulói réteget elérni: volt kis-, illetve nagyvárosban található szakközépiskola és gimnázium, elitnek tekinthető és átlagos középiskola is.

Eredmények

Az osztálytermi értékelés jelenlegi gyakorlata

A tanári kérdőívben arra voltunk kíváncsiak, hogy milyen készségeket és nyelvi elemeket értékelnek a tanárok. A válaszadóknak listán felsorolt elemekről kellett nyilatkozniuk, hogy azokat az osztálytermi munka során értékelik-e. Hat lehetséges területből lehetett megjelölni azokat, amelyek relevánsak voltak a válaszadó számára (lásd a 2. táblázatban), és megnevezhettek egy olyan területet (Egyéb), amelyet még fontosnak tartottak, de nem szerepelt a felsorolásban.

Az 2. táblázat szerint Magyarországon, ahogy az összes részt vevő országban általában, a megkérdezettek a legkevésbé a hallott szöveg értését és a beszédkésztséget értékelik. Az íráskészséget nem számítva a szókinccs és nyelvtan értékelése megelőzi a nyelvi készségek értékelését a négy országot összességében tekintve. Ez a tendencia feltehetően összefüggésben áll azzal, hogy az írásban teljesített feladatok értékelése vonzóbb a tanárok számára látszólagos objektivitásuk miatt, vagy a kész tesztekhez való könnyű hozzáférés miatt, illetve a beszédértés és beszédkésztség értékelése bizonyultabbnak tűnhet a számukra.

A tanulók válaszai többnyire hasonló képet mutatnak (3. táblázat), mint a tanároké a nyelvi készségek/elemek értékelését illetően. A nyelvórákon minden bizonnyal a hallott

Tanárok: Mely nyelvi készségeket/elemeket értékelik?				
	Összes válaszadó N	Összes válaszadó %	Magyarország N	Magyarország %
Íráskészség	795	93,31	210	91,30
Szókincs	783	91,90	206	89,57
Nyelvtan	768	90,14	215	93,48
Olvasott szöveg értése	755	88,62	211	91,74
Beszéd-készség	738	86,62	201	87,39
Hallott szöveg értése	706	82,86	178	77,39
Egyéb (pl. házi feladat, projektmunka, prezentálási készség)	59	6,92	0	0,00

2. táblázat. Milyen nyelvi készségeket/elemeket értékelnek az angol nyelvi órán a tanárok

szöveg értése a legkevésbé értékelt nyelvi készség minden vizsgált országban, hiszen itt mindkét csoport válasza megegyeznek. Ami eltér azonban a tanárok és tanulók meglátása szerint, az a receptív készségek %-os előfordulási aránya: a tanárok szerint kedvezőbb, mint a tanulók szerint. A magyar és a többi vizsgált ország tanulóinak a meglátása különbözik abban a tekintetben, hogy a beszéd-készség tűnik Magyarországon a második legkevésbé értékelt készségnek, míg a többi országban ez a második leginkább értékelt készség.

Tanárok: Mely nyelvi készségeket/elemeket értékelik?				
	Összes válaszadó N	Összes válaszadó %	Magyarország N	Magyarország %
Íráskészség	795	93,31	210	91,30
Szókincs	783	91,90	206	89,57
Nyelvtan	768	90,14	215	93,48
Olvasott szöveg értése	755	88,62	211	91,74
Beszéd-készség	738	86,62	201	87,39
Hallott szöveg értése	706	82,86	178	77,39
Egyéb (pl. házi feladat, projektmunka, prezentálási készség)	59	6,92	0	0,00

3. táblázat. Mely nyelvi készségeket/elemeket értékelik a tanárok a tanulók szerint

A kérdőívekben arra is hangsúlyt fektettünk, hogy adatot nyerjünk a tanulóknak szánt visszajelzési/visszacatolási formákról, melyek megjelenhetnek minősítő, szummatív értékelésként (pl. érdemjegy) vagy formatív elemként (pl. rövid szöveges megjegyzés), de mindkét esetben befolyással lehetnek a tanulók előmenetelére. A tanároknak egy adott listából kellett kiválasztani azokat a visszajelzési formákat, melyeket a nyelvórákon használnak (4 megadott visszajelzési formából tetszőleges számút megjelölhettek, és egy általuk megnevezett, a listán nem szereplő formát megnevezhettek, lásd a 4. táblázatban). A tanulói kérdőív pedig arra kérdezett rá, hogy a diákok milyen visszajelzési módszereket tapasztalnak az órákon (a lehetséges formák megegyeztek a tanári kérdőívben szereplő opciókkal). A 4. táblázatban látható eredmények szerint a tanárok érdemjegyet és rövid szöveges megjegyzést használnak, a részletes értékelés vagy a diáknak szánt egyéni útmutatás (hogy tudna jobban fejlődni) sokkal kevésbé használt formatív visszacsatolási formák. A részletes visszajelzés elengedhetetlennek tűnik ahhoz, hogy a diák értse, hogy min kell változtatnia ahhoz, hogy előbbre tudjon lépni. A helyzet megértéséhez sajnos csak feltételezésekre hagyatkozhatunk: lehetséges, hogy a válaszadók nem kaptak arra irányuló képzést, hogy a részletes visszajelzést kezelni tudják, és hogy értsék a tanulás folyamatában betöltött szerepét. Sok tanár a visszacsatoláson mindössze annyit ért, hogy a tanuló hibáit ki kell javítani (Sheen–Ellis 2011), jöllehet a pozitív visszajelzés (a tanuló erősségeinek elismerése) és a továbbfejlődéshez adott konkrét útmutatás/jó tanács is a formatív értékelés részét képezik.

Tanárok: Milyen visszajelzési formákat használnak?				
	Összes válaszadó N	Összes válaszadó %	Magyarország N	Magyarország %
Érdemjegy (pl. százalék, osztályzat)	741	86,97	214	93,04
Rövid szöveges megjegyzés (pl. „Szép munka!”)	691	81,10	180	78,26
Részletes értékelés a szóbeli és/vagy írásbeli teljesítményről	507	59,51	122	53,04
Megjegyzés, útmutatás arra vonatkozóan, hogy hogyan tudna a diák fejlődni / jobb eredményt elérni	591	69,37	148	64,35
Egyéb	40	4,69	9	3,91

4. táblázat. Milyen visszajelzési formákat használnak a tanárok

A diákok perspektívájából nézve hasonló kép bontakozik ki (lásd 5. táblázat), bár a részletes értékelés gyakorisága Magyarországon nemcsak a tanárok meglátásához képest alacsonyabb, hanem az összes vizsgált ország tanulóihoz képest is. Úgy tűnik, hogy a vizsgált négy országból nálunk van legkevésbé hagyománya a formatív értékelés eme formájának. A népszerűtlenség oka az is lehet, hogy ez az értékelési forma kétségtelenül munkaigényes, hiszen sok odafigyelést, időt, tapasztalatot és a szituációnak megfelelő nyelvi megfogalmazást kíván a tanár részéről. Az ilyen típusú visszajelzés csak akkor lesz hasznos, ha egyénre szabott, támogató jellegű és a diák számára is érthető a megfogalmazott célok elérése érdekében, és persze a diák követi is a tanári útmutatást.

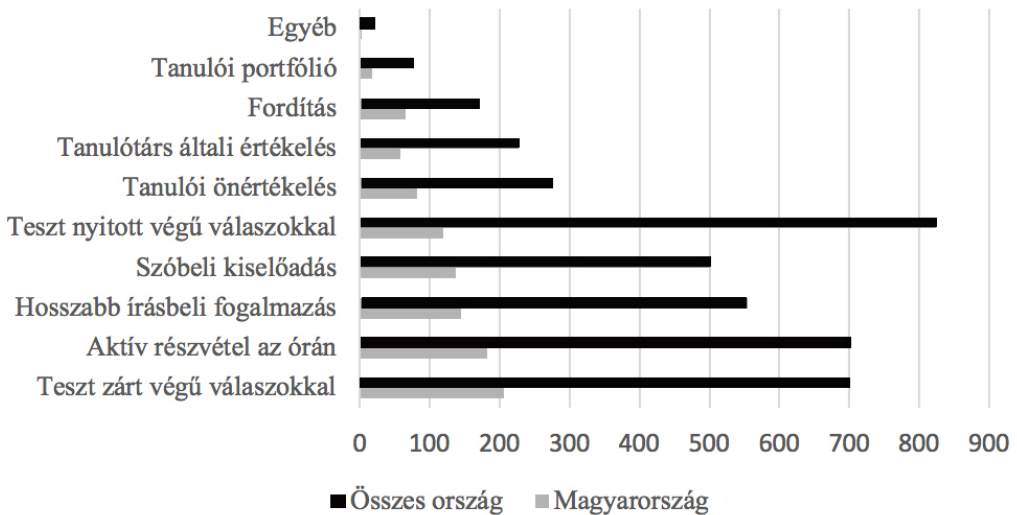
Tanulók: Milyen visszajelzési formákat használnak a tanárok?				
	Összes válaszadó N	Összes válaszadó %	Magyarország N	Magyarország %
Érdemjegy (pl. százalék, osztályzat)	1587	88,75	279	93,00
Rövid szöveges megjegyzés (pl. „Szép munka!”)	1208	67,56	178	59,33
Részletes értékelés a szóbeli és/vagy írásbeli teljesítményről	751	42,00	71	23,67
Megjegyzés, útmutatás arra vonatkozóan, hogy hogyan tudna a diák fejlődni / jobb eredményt elérni	1026	57,38	170	56,67
Egyéb	0	0,00	0	0,00

5. táblázat. Milyen visszajelzési formákat használnak a tanárok a tanulók szerint

Az osztálytermi értékelés jelenlegi gyakorlatának feltárásához végül azt is igyekeztünk megvizsgálni, hogy a különböző értékelési módszereket milyen gyakran alkalmazzák a tanárok, egy négyfokozatú skálán (0 [soha] – 3 [nagyon gyakran]) bejelölve a rájuk jellemző gyakoriságot. A válaszadók önbevallás alapján megjelölt gyakorisági adatait, a „nagyon gyakran” és „gyakran” válaszokat összevontuk a fokozott átláthatóság kedvéért, melynek alapján az a tendencia figyelhető meg, hogy a tanulóközponitú értékelési formák alig jelennek meg az osztályteremben a vizsgált országokban, hiszen a tanulói portfólió és a tanulótárs általi értékelés tűnik a legkevésbé használt értékelési módszernek, melyeket alig előz meg gyakoriságban a tanulói önértékelés (lásd 1. ábra). A leggyakoribb értékelési formák tanárközponitúak, azaz

hagyományosnak mondhatók: zárt vagy nyitott végű kérdéseket tartalmazó tesztek, aktív részvétel az órán, hosszabb írásbeli fogalmazás és szóbeli kiselőadás. A tanulóközpontú értékelés háttérbe szorulásának tényén talán egyedül az javít, hogy a fordítás mint értékelési forma is kevésbé gyakran használt a tanárok állítása szerint. A legnépszerűbb értékelési forma az összes válaszadó tanár véleménye alapján a nyitott végű kérdéseket tartalmazó teszt, a megkérdezett magyarországi válaszadó tanárok viszont zárt kérdéseket tartalmazó teszteket alkalmaznak a leggyakrabban.

A négy vizsgált ország között szignifikáns különbségek tapasztalhatók az egyes értékelési formák bevallott gyakorisága között (lásd 6. táblázat), kivéve egy értékelési formát, a *Tanulótárs általi értékelést* (egyszempontos varianciaanalízis, $F = 1,53$, $p = 0,205$), ahol nincs szignifikáns különbség a négy ország válaszadó tanárai között a bevallott gyakorisági adatok alapján. Ebből arra következtethetünk, hogy ez az értékelési forma mind a négy vizsgált országban azonos gyakoriságú és egyaránt kevésbé használt. Jóllehet ennek okát érdemes lenne megfejteni, a rendelkezésre álló adatok alapján ezt jelenleg nem tudjuk feltérképezni.



1. ábra. A különböző értékelési módszerek gyakoriság szerinti listája a magyarországi és az összes válaszadó tanár (230 vs. 852 fő)

Mivel a tanulók is véleményezték, hogy milyen gyakran használja a tanár a különböző értékelési formákat, így árnyaltabb képet kapunk a valóságról a két válaszadó csoport véleményének összevetésével. A 7. táblázatban látható az átlag gyakoriság (0 [soha] – 3 [nagyon gyakran]) a tanárok és a diákok állítása alapján az egyes országokra külön-külön lebontva.

<i>Értékelési módszerek</i>	F	p
Teszt nyitott végű válaszokkal	42,477	,000
Tanulói portfólió	4,184	,006
Tanulótárs általi értékelés	1,530	,205
Teszt zártvégű válaszokkal	21,392	,000
Tanulói önértékelés	8,806	,000
Hosszabb írásbeli fogalmazás	29,147	,000
Aktív részvétel az órán	8,524	,000
Fordítás	34,201	,000
Szóbeli kiselőadás	6,898	,000

6. táblázat. Egyszempontos varianciaanalízis eredménye a különböző értékelési módszerek országonként bevallott gyakoriságára vonatkozóan (sig < 0,05)

	<i>Ciprus</i>	<i>Német-ország</i>	<i>Görög-ország</i>	<i>Magyar-ország</i>	<i>Összes ország</i>
<i>Értékelési forma</i>	Tanárok				
Szóbeli kiselőadás	1,85	1,48	1,93	1,72	1,76
Teszt nyitott végű válaszokkal	1,85	0,96	1,79	1,57	1,63
Tanulói portfólió	0,62	0,56	0,70	0,43	0,56
Tanulótárs általi értékelés	1,09	0,95	1,19	1,04	1,06
Teszt zárt végű válaszokkal	2,24	1,67	2,32	2,29	2,18
Tanulói önértékelés	1,23	0,93	1,53	1,29	1,23
Hosszabb írásbeli fogalmazás	1,96	1,18	1,87	1,74	1,77
Aktív részvétel az órán	2,45	2,12	2,43	2,16	2,32
Fordítás	0,64	1,70	1,48	1,12	0,85
<i>Értékelési forma</i>	Tanulók				
Szóbeli kiselőadás	0,87	1,21	1,52	0,95	1,05
Teszt nyitott végű válaszokkal	1,67	1,43	1,73	1,64	1,64
Tanulói portfólió	0,47	0,83	1,40	0,23	0,64
Tanulótárs általi értékelés	1,01	1,26	1,16	1,01	1,07
Teszt zárt végű válaszokkal	2,23	2,17	2,38	1,82	2,17
Tanulói önértékelés	1,26	1,12	1,68	0,86	1,24
Hosszabb írásbeli fogalmazás	1,66	1,90	1,68	1,84	1,73
Aktív részvétel az órán	2,20	1,55	2,31	1,48	1,99
Fordítás	1,61	1,58	1,72	1,96	1,68

7. táblázat. Különböző értékelési formák országonkénti gyakorisági átlaga a válaszadó tanárok és tanulók szerint

A 7. táblázatban található adatok alapján összevetettük az összes válaszadó tanár és diák állításait, és csak három esetben találtunk szignifikáns különbséget a tanárok és tanulók meglátásai között: a szóbeli kiselőadás, az aktív részvétel az órán és a fordítás mint értékelési formák gyakoriságában mutatkozott szignifikáns különbség a független mintás t-próba alapján (lásd 8. táblázat). Az aktív részvétel az órán mint értékelési forma valószínűleg informálisan és ezért kevésbé észrevehető módon zajlik a tanulók számára, így a gyakoriság megítélésében mutatkozó különbség indokolható. A szóbeli kiselőadás és fordítás gyakoriságának megítélésében jelentkező különbség azonban nehezen magyarázható: mindkét forma direkt teljesítményt hív elő, és jól beazonosítható módon zajlik. Különösen a fordítás megítélése elgondolkoztató, mivel az kétszer olyan gyakori előfordulású a tanulók szerint, mint a tanárok megítélése alapján.

<i>Értékelési módszerek</i>	t	p
Teszt nyitott végű válaszokkal	-,164	,870
Tanulói portfólió	-1,932	,053
Tanulótárs általi értékelés	-,297	,767
Teszt zártvégű válaszokkal	,033	,974
Tanulói önértékelés	-,383	,702
Hosszabb írásbeli fogalmazás	1,052	,293
Aktív részvétel az órán	8,440	,000
Fordítás	19,849	,000
Szóbeli kiselőadás	20,480	,000

8. táblázat. Független mintás t-próba eredménye a különböző értékelési módszerek gyakoriságára vonatkozóan (a tanárok és tanulók megítélésének összevetése alapján, sig < 0,05)

A tanulói kérdőív arra is kitért, hogy a fentebb felsorolt értékelési formák közül melyeket tekintik a leghasznosabbnak a tanulók az angoltudásuk fejlesztése szempontjából (lásd 9. táblázat). A tanulóknak a megadott listán szereplő értékelési formákról kellett nyilatkozniuk egy 4 pontos Likert-skálán: ha használják azokat az órán, mennyire segítik őket a nyelvtanulásban (0 = nem sokat segít, 3 = nagyon hasznos). A magyarországi válaszadó diákok meglátásai némiképp eltérnek az összes ország adatai alapján megmutatkozó trendtől. Míg az összes válaszadó diák szerint az aktív részvétel jelenti a legnagyobb segítséget, addig ugyanez az értékelési forma a magyar diákok szerint csak a 4. leghasznosabb forma. Egyetértés mutatkozik azonban az összes válaszadó részéről annak tekintetében, hogy a hosszabb írásbeli fogalmazás, a fordítás és a zártvégű

kérdéseket tartalmazó tesztek hasznos értékelési formák a tanulás támogatásában. Ez az eredmény azért meglepő, mert a tanárok a fordítást az alig használt értékelési formák közé sorolták. Az adatok alapján csak spekulálni lehet az okokat illetően: a tanulók válaszai egyrészt tükrözhetik a valós osztálytermi gyakorlatot, mely ez esetben ellentmond a tanári állításoknak, másrészt azt jeleníthetik meg, hogy a tanulók mit tekintenek ideális értékelési formának. Az előbbi ellentmondás mellett igen nagy összhang fedezhető fel a tanulóközpontú értékelési formákat illetően: a tanulói önértékelés, a tanulótárs általi értékelés és a tanulói portfólió sem a tanárok, sem a diákok által nem preferált értékelési formák. Ez az eredmény azt jelzi, hogy ezen a területen feltétlenül szemléletváltásra lenne szükség, hiszen a tanulás érdekében történő értékelés (*assessment for learning*) kifejezetten ezekre az értékelési formákra támaszkodik.

<i>Értékelési módszerek</i>	Összes ország a megítélések átlaga	Magyarország a megítélések átlaga
Teszt nyitott végű válaszokkal	1,71	1,51
Tanulói portfólió	0,80	0,58
Tanulótárs általi értékelés	1,19	0,98
Teszt zártvégű válaszokkal	1,98	1,79
Tanulói önértékelés	1,54	1,09
Hosszabb írásbeli fogalmazás	1,98	2,19
Aktív részvétel az órán	2,10	1,66
Fordítás	1,95	2,10
Szóbeli kiselőadás	1,43	1,34

9. táblázat. A tanulók megítélései a különböző értékelési formákat illetően (4 pontos Likert-skála: 0 = nem sokat segít, 3 = nagyon hasznos)

A tanárok értékelési profilja és továbbképzési szükségletei

A tanári kérdőív központi része arra irányult, hogy a válaszadók nyelvtudásmérési műveltségét érintően megvizsgálja, hogy mely területek alkalmazásában magabiztosak, és mely területeken igénylik a (tovább)képzést. Értelemszerűen a kapott válaszok jelentős mértékben befolyásolták a TALE online képzés tartalmi összeállítását a tervezés fázisában, hiszen arra törekedtünk, hogy minél nagyobb mértékben megfeleljünk a képzési igényeknek.

A kérdőívben 20 értékelési témakörrel kellett nyilatkozniuk a válaszadóknak: egyrészt arról, hogy mennyire érzik magukat magabiztosnak az adott témakörben, másrészt hogy szeretnének-e (tovább)képzésben részesülni az adott területet érintően. A magabiztosságot egy Likert típusú, 4 pontos skálán kellett bejelölni, ahol 1 = egyáltalán nem vagyok magabiztos, 2 = kicsit vagyok magabiztos, 3 = kellően magabiztos vagyok, 4 = nagyon

magabiztos vagyok. A továbbképzési igényeket pedig egy külön oszlopban jelölhették meg a válaszadók ugyanahhoz a listához tartozó értékelési témakörökre vonatkozóan. A 10. táblázatban találhatóak az értékelési témakörök, valamint az összes ország és a magyarországi válaszadó tanárok magabiztossági átlaga.

<i>Értékelési/mérési témakörök</i>	Összes ország magabiztossági átlag	Magyarország magabiztossági átlag
1. Az értékelés különböző céljainak beazonosítása	1,94	2,93
2. A megfelelő értékelési formák/technikák kiválasztása	2,04	3,03
3. Osztálytermi haladást mérő tesztek összeállítása	2,04	2,96
4. Hallott szöveg értésének mérése	1,88	2,78
5. Beszédképesség értékelése	2,05	3,05
6. Olvasott szöveg értésének mérése	2,19	3,12
7. Írásképesség értékelése	2,11	2,99
8. A négy nyelvi készségek integrált mérése	1,81	2,60
9. Tanulói önértékelés használata	1,41	2,36
10. Tanulótárs általi értékelés használata	1,33	2,16
11. Tanulói portfólió használata	,86	1,62
12. Speciális tanulási nehézséggel rendelkező tanulók értékelése (pl. diszlexia)	,70	1,67
13. A tanulók felkészítése külső vizsgákra (pl. érettségi, nemzetközi nyelvvizsgák)	1,85	3,03
14. Az értékelés szerepe egyéni tanulókra vonatkozó döntések meghozatalában	1,86	2,87
15. Az értékelés szerepe a tanított anyag megtervezésében	2,01	3,02
16. Annak felismerése, hogy a nyelvi tesztelés/mérés hogyan befolyásolja a tanítást	1,78	2,74
17. A Közös európai referenciakeret (KER) ismeretek alkalmazása a tanulók értékeléséhez	1,40	2,53
18. Az értékelés eredményének értelmezése a tanulók számára	2,08	3,15
19. Az értékelés eredményének értelmezése szülők és mások számára	2,03	3,02
20. A nem megfelelő értékelési módszerek felismerése	1,67	2,69

10. táblázat. Az összes ország és a magyarországi válaszadó tanárok magabiztossági átlagai az egyes értékelési/mérési témakörökre vonatkozóan

<i>Igényfelmérés az értékelési/mérési témakörök iránt</i>	Magyarország		Összes ország	
	N	%	N	%
Speciális tanulási nehézséggel rendelkező tanulók értékelése (pl. diszlexia)	173	75,22%	566	66,43%
Tanulói portfólió használata	148	64,35%	433	50,82%
Tanulótárs általi értékelés használata	124	53,91%	339	39,79%
Tanulói önértékelés használata	117	50,87%	346	40,61%
A nem megfelelő értékelési módszerek felismerése	104	45,22%	299	35,09%
Hallott szöveg értésének mérése	98	42,61%	279	32,75%
A Közös európai referenciakeret (KER) ismeretek alkalmazása a tanulók értékeléséhez	91	39,57%	340	39,91%
A négy nyelvi készség integrált mérése	90	39,13%	357	30,16%
A megfelelő értékelési formák/technikák kiválasztása	89	38,70%	282	33,1%
Beszédképesség értékelése	85	36,96%	240	28,17%
Osztálytermi haladást mérő tesztek összeállítása	83	36,09%	237	27,82%
Írásképesség értékelése	71	30,87%	209	24,53%
A tanulók felkészítése külső vizsgákra (pl. érettségi, nemzetközi nyelvvizsgák)	69	30,00%	210	24,65%
Az értékelés különböző céljainak beazonosítása	67	29,13%	239	28,05%
Annak felismerése, hogy a nyelvi tesztelés/mérés miként befolyásolja a tanítást	65	28,26%	198	23,24%
Az értékelés szerepe a tanított anyag megtervezésében	56	24,35%	174	20,42%
Olvasott szöveg értésének mérése	55	23,91%	157	18,43%
Az értékelés szerepe egyéni tanulókra vonatkozó döntések meghozatalában	46	20,00%	159	18,66%
Az értékelés eredményének értelmezése szülők és mások számára	43	18,70%	150	17,61%
Az értékelés eredményének értelmezése a tanulók számára	36	15,65%	123	14,44%

11. táblázat. A válaszadó tanárok képzés iránti szükségletei

Az összes ország és a magyarországi válaszadók összevetésekor jól látható, hogy az összes válaszadó tanár mind a négy nyelvi készség mérésének tekintetében elég magabiztos (a magabiztossági átlagok 1,88 és 3,12 között mozognak). Az összes válaszadó tanár a legkevésbé magabiztos a speciális tanulási nehézséggel rendelkező diákok értékelésében, a tanulói portfólió, a tanuló társ általi értékelés, a tanulói önértékelés használatában és a KER-ismeretek alkalmazásában (a magabiztossági átlagok 0,7 és 1,33 között mozognak). A magyar adatok is hasonló tendenciákat tükröznek, mint az összes részt vevő országban. Továbbá megállapítható a független mintás t-próba eredménye alapján, hogy szignifikáns eltérés mutatkozik a magyarországi és az összes ország válaszadó tanárainak magabiztosságában az összes értékelési/mérési témakör tekintetében ($p = 0,000$, $\text{sig} < 0,05$).

A vizsgálat arra is kiterjedt, hogy a korábban kapott képzések vajon miként befolyásolták a válaszadók magabiztosságát. Nem meglepő módon minden területen szignifikánsan magabiztosabbnak mutatkozott minden válaszadó ($\text{sig} < 0,05$), aki korábban (tovább)képzésben részesült a nyelvi értékelés és mérés területén, kivéve a tanulói portfólió használatát, melyre vonatkozóan nem mutatkozott szignifikáns különbség a már (tovább)képzésben részesült, illetve nem részesült válaszadók között (egyszempontos varianciaanalízis, $F = 2,360$ $p = 0,125$, $\text{sig} < 0,05$).

Az idegen nyelvi értékelésre és mérésre vonatkozó tanári magabiztosság vizsgálata mellett a válaszadók képzés iránti szükségleteit is feltérképeztük a képzési terület megjelölésével (lásd *II. táblázat*). A várakozásoknak megfelelően a válaszadók maguk is azokon a területeken szeretnének továbbképzésben részesülni, amelyekben a magabiztosságuk is többnyire alacsony. Idetartozik a speciális tanulási nehézséggel rendelkező tanulók értékelése, a tanulói portfólió használata, a tanuló társ általi értékelés és tanulói önértékelés használata. Az igen heterogén válaszadói populáció visszajelzései azt is tükrözik, hogy egy olyan javasolt képzési terület sincs, amelyre vonatkozóan nincs senkinek szüksége tréningre.

Összefoglalás

A tanulmányban a TALE-projekt keretében végzett kutatásról számoltunk be, mely alapul szolgált egy online, szabad hozzáférésű, innovatív tanár-továbbképzési platform létrehozásához azzal a céllal, hogy a nyelvtanárok nyelvtudásmérési műveltségét megerősítse. Kutatásunk egyik újszerű eleme az volt, hogy nemcsak tanári meglátásokat gyűjtöttük az osztálytermi értékelés gyakorlatát illetően, hanem a nyelvtanulók oldaláról is megvizsgáltuk ugyanazokat a kérdéseket, valamint négy európai ország viszonylatában folytattuk le a vizsgálatainkat.

Az első kutatási kérdésünk az osztálytermi értékelés jelenlegi gyakorlatának feltérképezésére irányult. A vizsgálat eredményei alapján kibontakozó kép azt mutatja, hogy a hagyományos, tanárközpontú értékelési formák dominálnak az angol nyelvi

órán, a tanulói, tanulótárs általi értékelésnek, illetve a tanulói portfóliónak kevesebb szerep jut. Ezt egységesen ítélték meg a tanárok és a tanulók is, és a jelenség egyformán jellemző mind a négy vizsgált országban (Ciprus, Görögország, Németország, Magyarország). Amennyiben az élethosszig tartó tanulás támogatása, illetve a tanulói autonómia erősítése az oktatáspolitikai célkitűzései közé tartoznak, elengedhetetlen, hogy az osztálytermi munka során a tanulást közvetlenül támogató értékelési formák nagyobb hangsúlyt kapjanak. Ehhez mind a tanárok, mind a tanulók szemléletváltására szükség van, mely irányába az első lépést a tanárok továbbképzéseken tehetik meg.

A második kutatási kérdésünk a tanárok értékelési profiljára (magabiztosságukra és képzési igényeikre) vonatkozott. A válaszadó angoltanárok magabiztossága a nyelvtudásmérési műveltségüket tekintve elég vegyes képet mutat, kiemelkedő magabiztossággal rendelkező alcsoport egyik régióban sem fordult elő. A kapott eredmények egyértelműen alátámasztják azt a feltételezést, hogy nagy szükség és egyben igény van (tovább)képzésre ezen a területen. Az alternatív, tanulóközpontú értékelési módszerek (önértékelés, tanulótárs általi értékelés, tanulói portfólió) és a tanulási nehézséggel rendelkező diákok értékelése iránt mutatkozott a legnagyobb igény. Hasonló eredményre már korábbi kutatások is jutottak (Kvasova–Kavytska 2014, Vogt–Tszagari 2014).

A harmadik kutatási kérdésünkre, azaz az osztálytermi értékelési gyakorlat árnyaltabb megítéléséhez a tanári és tanulói vélemények összevetésével kerestük a választ. A vizsgálat során a tanulói és tanári vélemények többnyire összecsengést mutattak, kivéve a fordítás gyakoriságát és súlyát tekintve. A tanárok állítása szerint viszonylag ritkán használják ezt a módszert, azonban a tanulók szerint a gyakrabban alkalmazott fordítás kívánatos lenne számukra, mert megítélésük szerint az nagyban segítené őket a tanulásban. A tanulók nyelvtudásmérési műveltségét mindenképpen indokolt fejleszteni ezen a téren, hiszen feltételezhető, hogy nem ismerik kellően az olyan tanulóközpontú értékelési módszereket, mint a tanulói portfólió vagy tanulótárs általi értékelés. Ha az órákon csak kevésbé alkalmazzák őket a tanárok, a tanulóknak esélyük sincs arra, hogy a nyelvtudásuk fejlesztését támogató alternatív értékelési formákat is megismerhessenek.

Az osztálytermi gyakorlat vizsgálatát illetően további fontos eredmény, hogy a tanárok és a tanulók egyaránt úgy ítélik meg, hogy az értékelés során adott visszajelzés/visszacsatolás leggyakoribb formái az érdemjegy és rövid értékelő megjegyzés. Ez azt sugalmazza, hogy a nyelvórákon a formatív értékelés megítélése miatt a tanárok kevésbé informatív és előremutató formákat használnak, részletes útmutatót és értékelést csak nagyon kevés tanuló kap.

A kutatás korlátai között feltétlenül meg kell említeni, hogy a teljes válaszadói minta a négy vizsgált országban nem volt kiegyenlített, a ciprusi tanárok és tanulók igencsak magas létszámarányban képviselték magukat, mivel ebben az országban volt a legerősebb a részvételi hajlandóság. Ez a tény a kapott eredményeket befolyásolhatta.

Továbbá nem minden ország válaszadói mintája kiegyensúlyozott a különböző paraméterek tekintetében (pl. kor, tanítási tapasztalat). A kérdőíves felmérés korlátait is el kell ismernünk: egyes terminológiák értelmezése eltérést mutathat, különösen a tanulók vonatkozásában. Továbbá a válaszadók állításai önbevalláson alapulnak, és ezért valóságtartalmukat nem lehet ellenőrizni. A kutatás utóbbi korlátjának az ellensúlyozása okán az osztálytermi gyakorlatot nemcsak egy, hanem két érintett (tanárok és tanulók) szemszögéből is vizsgáltuk egyidejűleg, és ez lehetővé tette számos kérdés összehasonlító elemzését. A tanulói vélemények több ponton megerősítették a tanári állításokat, de volt, ahol ellentmondást eredményezett az összevetés, például a fordítás mint értékelési módszer használatát illetően. A kutatási kérdések vizsgálatának további lehetséges irányai közé sorolható az osztálytermi értékelési módszerek részletesebb vizsgálata tanárok által készített tesztek elemzésével, vagy az osztálytermi értékelés órai megfigyelésével.

A TALE-projekt által létrehozott online, nyílt hozzáférésű, önálló tanulásra szánt képzési anyagok igen széles közönséget céloznak meg, mely magában foglalja a tanárjelölteket, a kezdő és tapasztalt tanárokat, az angoltanárokat, valamint az angolul tudó, de más idegen nyelvet tanító tanárokat is. A képzés ingyenesen hozzáférhető a világ bármely pontjáról, és ezzel a lehetőséggel élnek is sokan. A honlap látogatottságát illetően megállapítható, hogy egyre nő a felhasználók száma. 2019. január 28-ig összesen 703 regisztrált felhasználója és 2,127 látogatója volt a TALE projekt honlapjának. A Kézikönyv oldala pedig már 7976 megtekintést számlált 111 felhasználótól. Meglátásunk szerint a TALE-projekt igen nagy hatást gyakorolt a nyelvtudásmérés területére. Több rangos szakmai egyesület (pl. ILTA, EALTA) a hasznos forrásanyagok listájára tette az online kurzus elérhetőségét. 2019 májusában a TALE-projekt elnyerte az „Innováció az értékelésben” díjat, melyet a British Council oszt ki minden évben (Assessment Research Awards and Grants – Kutatási ösztöndíjak és elismerések az értékelés területén).

IRODALOM

- Black, P. – Wiliam, D. (1998): Inside the black box: Raising standards through classroom assessment, *Phi Delta Kappan International*. www.pdkintl.org/kappan/kbla9810.htm
- Davison, C. – Leung, C. (2009): Current Issues in English Language Teacher-Based Assessment. *TESOL Quarterly*, 43/3, 393-415.
- Earl, L. – Katz, S. (2006): *Rethinking classroom assessment with purpose in mind*. Winnipeg, Manitoba: Western Northern Canadian Protocol. <https://digitalcollection.gov.mb.ca/awweb/pdfopener?smd=1&did=12503&md=1>
- Fulcher, G. (2012): Assessment Literacy for the Language Classroom. *Language Assessment Quarterly*, 9/2, 113-132.

- Hasselgreen, A. – Carlsen, C. – Helness, H. (2004): *European Survey of Language Testing and Assessment Needs. Part 1: General findings*. Gothenburg, Sweden: European Association for Language Testing and Assessment. <http://www.ealta.eu.org/documents/resources/survey-report-pt1.pdf>
- Hill, K. – McNamara, T. (2012): Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29/3, 395–420.
- Inbar-Lourie, O. (2008): Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25/3, 385–402.
- Joint Committee on Standards for Education (JCSEE) (2015): *Classroom assessment standards for pre-K12 teachers*. Kindle edition. <http://www.jcsee.org/the-classroom-assessment-standards-new-standards>
- Közös Európai Referenciakeret. Nyelvtanulás, nyelvtanítás, értékelés. 2002. Budapest: Pedagógustovábbképzési Módszertani és Információs Központ Kht.
- Kvasova, O. – Kavytska, T. (2014): The Assessment Competence of University Foreign Language Teachers: A Ukrainian Perspective. *CercleS*, 4/1, 159–177.
- O’Loughlin, K. (2013): Developing the assessment literacy of university proficiency test users. *Language Testing*, 30/3, 363–380.
- Poehner, M. E. (2009): Group Dynamic Assessment: Mediation for the L2 Classroom. *TESOL Quarterly*, 43/3, 471–491.
- Rea-Dickins, P. (2001): Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, 18/4, 429–462.
- Schafer, W. D. (1991): Essential Assessment Skills in Professional Education of Teachers. *Educational Measurement: Issues and Practice*, 10/1, 3–6.
- Sheen, Y. – Ellis, R. (2011): Corrective Feedback in Language Teaching. In: Hinkel E. (szerk.) *Handbook of Research in Second Teaching and Learning*. Vol. II. New York/London: Routledge. 593–610.
- Standards for Teacher Competence in Educational Assessment of Students* (1990): American Federation of Teachers, National Council on Measurement in Education, National Education Association. <https://eric.ed.gov/?id=ED323186>
- Stiggins, R. J. (1991): Relevant Classroom Assessment Training for Teachers. *Educational Measurement: Issues and Practice*, 10/1, 7–12.
- Vogt, K. – Tsagari, D. (2014): Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11/4, 374–402.

Hogyan értelmezik a tanárszakosok a tanulói értékelés fogalmát?

How do teacher trainees interpret student assessment?

An important objective of teacher education is familiarizing teacher trainees with a differentiated assessment approach and training teachers who are committed to using assessment that facilitates learning. However, research often confirms that previous experiences as learners are still rigid, therefore traditional interpretation of student assessment with summative purposes is determinate. Less is known about the development of trainees' assessment literacy in the framework of teacher education. The purpose of this paper is to discuss the results of research conducted in a compulsory course dealing with pedagogical assessment. This study examines the characteristics of, and the changes in, teacher trainees' interpretation of student assessment at the beginning and at the end of the semester as well as comparing the complexity of trainees' interpretations with their self-evaluation of perceived changes. Findings indicate that teacher trainees tend to have more complex interpretations, but they still mostly define student assessment as feedback for only students about their subject knowledge. Trainees can appreciate the differentiation of their views and knowledge of assessment processes and goals. However, they need more opportunities and support in learning to use student assessment.

Keywords: teacher trainees, student assessment, assessment literacy, content analysis, action research

Bevezetés

A pedagógusmunka minőségének fejlesztése az elmúlt időszakban kiemelt oktatáspolitikai prioritás, mivel az oktatási rendszerek eredményességének egyik meghatározó tényezője a pedagógusok szakmai tudásának és képességeinek színvonala (Guerriero 2017). Ezért szükséges a tanárjelöltek kompetenciáinak fejlesztése, amelyek közé tartozik a pedagógiai folyamatok és a tanulók értékelése. E folyamatban azonban kihívást jelent, hogy a tanulói értékelés értelmezése a modern tanári szerepértelmezéssel változott: a segítő, támogató magatartás már nem a tanulók összegző minősítését teszi a legfőbb értékelési feladattá. A hazai tanárképzés egyik fontos célja ezért a differenciált, a formatív értékelést is magában foglaló értékelési szemlélet kialakítása. A tanárjelöltek a képzésük során találkozhatnak azokkal az értékeléssel kapcsolatos ismeretekkel és magatartással, amelyek a tradicionális mellett a modern (tutorjellegű) tanárszerepre vonatkoznak, ám ezek a saját tapasztalataikkal nem feltétlenül állnak összhangban. A megfelelő fejlesztési lehetőségek azonosítása, a szemléletváltás megvalósítása és

a tudományosan igazolt értékelési módoknak az oktatási folyamatba való integrálása éppen ezért fontos (Kotschy 2016). Kevés hazai kutatást végeztek azonban arról, hogy a tanárrá válás folyamata során miként alakulnak a tanárjelöltek értékelésről alkotott nézetei, viszonylag kevés empirikus kutatás foglalkozik az értékelési tudás elsajátítási folyamatának nyomon követésével, illetve azzal, hogy ezt a folyamatot a formális képzés keretében milyen tanulási környezet kialakításával lehet leginkább támogatni.

A tanulmányunkban egy akciókutatás eredményeit mutatjuk be. Ennek célja, hogy a tanárképzésben részt vevők körében, pedagógiai értékeléssel foglalkozó gyakorlati kurzus keretében kövessük nyomon a tanárjelöltek értékeléssel kapcsolatos gondolkodását, a fejlődésük támogatása érdekében pedig aktív tanulásra alkalmas környezetet alakítsunk ki. Az elsajátítási folyamat feltárása és a kurzus eredményességének értékelése céljából a tanárszakosok értékeléssel kapcsolatos fogalomértelmezését vizsgáltuk két mérési ponton, a félév elején és a végén; valamint elemeztük a félév végi önértékelésüket abból a szempontból, hogy ők miben érzékelték változást. E tanulmányban két fő célt tűzünk ki. Egyrészt bemutatjuk a fogalomértelmezéseken végzett tartalom-elemzés módszereit és megbízhatóságát. Másrészt célunk a tanárjelöltek értékelésértelmezésének jellemzése a félév elején és ehhez képest a változás bemutatása, valamint a meghatározás komplexitásának összehasonlítása az önértékelésükkel. Mindezek alapján fogalmazzuk meg a következtetéseket az oktatói célok és a tanárjelöltek megnyilatkozásainak összevetésével.

Először a vizsgálatot közvetlenül megalapozó szakirodalmi háttérrel tárgyaljuk. Ezt követően jellemezzük az akciókutatásunk módszereit: a kurzusok főbb jellemzőit, a vizsgálat eszközeit és fázisait. Majd bemutatjuk a tartomelemzés és a fogalom meghatározásokon végzett statisztikai elemzések eredményeit és azok értelmezését. Végül összegezzük a főbb eredményeket, és olyan következtetéseket fogalmazzuk meg, amelyek segíthetik a képzés hatékonyságának növelését, a fejlesztési irányvonalak kijelölését.

A pedagógiai értékelés és a tanárjelöltek értékelési kompetenciájának jellemzői és vizsgálata

A vizsgálatot az alábbi szakirodalmi előzményre építettük. Ezek közé tartoznak egyrészt a pedagógiai értékelés jellemzői, másrészt az értékelési kompetencia fogalmának értelmezése és összetevőinek meghatározása. Harmadrészt pedig a tanárképzésben részt vevők értékeléssel kapcsolatban elvárt tudása és az abban bekövetkező fejlődést feltáró vizsgálatok tanulságai. A továbbiakban ezeket vázoljuk röviden.

A pedagógiai értékelés meghatározó jelentőségű a tanítás-tanulás rendszerében, hiszen ennek során szervezett és differenciált módon gyűjthetünk pedagógiai információkat, amelyek alapján az oktatási rendszer egyes szereplői számára adhatunk visszajelzéseket (Báthory 2000). Scriven (1967) munkája nyomán hagyományosan az értékelés három

célját különböztetjük meg: a tanítási-tanulási folyamat elején a helyzetfeltáró céllal végezhető diagnosztikus értékelést, a folyamatot segítő-támogató formatív értékelést, illetve a tanulási szakaszok végén a minősítést betöltő szummatív értékelést. E célok mentén pedagógiai információk gyűjthetők, amelyek főként a tanulók tudásában bekövetkező változásokra vonatkoznak a kognitív (ismeret- és képességjellegű tudás), az affektív (pl. motívumok, attitűdök, szokások) és a pszichomotoros (pl. finom- és nagymozgások elvégzésére vonatkozó) tudásterületeken (lásd Bloom 1956, Anderson–Krahtwohl 2000).

Az értékelés többféle módszerrel végezhető, céltudatos, tervszerű folyamat. Az értékelés módszerei közé tartoznak a teljesítményeket számszerűsítő mennyiségi értékelési eljárások (pl. becslés, mérés, tesztelés), illetve a minőségi értékelés, amelynek segítségével komplex módon, kötöttségek nélkül jellemezhető a teljesítmény (Csapó 2005). Lényeges módszertani kérdés, hogy mihez viszonyítva történjen a teljesítmények értékelése. Háromféle viszonyítás lehetséges: a kritériumorientált értékeléssel a követelményekhez, a normaorientált értékeléssel a referenciacsoport többi tagjához lehet viszonyítani, illetve a korábbi teljesítményszinthez is. Ezek a viszonyítási módok gyakran kerülhetnek ellentmondásba egymással, amelynek feloldásához szükséges a transzparens, a különböző jóságmutatókat érvényesítő értékelés alkalmazása (lásd Csapó 2004). Mindezek mentén három kritériumot különböztethetünk meg: az objektivitás akkor érvényesül, ha az értékelés kizárólag a vizsgált teljesítménytől függ, és nem befolyásolja az adatfelvevő és az értékelő személye; a validitás akkor teljesül, ha az értékelés módszere összhangban van az előre tisztázott értékelési célokkal; a reliabilitás esetén pedig az értékelés pontosságát vesszük figyelembe. Az értékelés további szempontja, hogy ki végzi az értékelést (a pedagógus vagy előre megértett szempontok segítségével a tanuló vagy más külső szakember), a gyűjtött információkat az oktatási rendszer mely szintjére vonatkoztatják, és azokat ki használja fel (lásd Golnhofner 2003). Mindezen módszertani kérdéseket alapvetően a célok tükrében lehet és kell meghatározni.

A pedagógiai értékelésre vonatkozó hazai irodalomból az értékelés céljai, típusai, funkciói jól ismertek (lásd Vidákovich 1990, Csíkos 2002). Az elmúlt évtizedben a fejlesztő értékelés jellemzése meghatározó (Brassói–Hunya–Vass 2005, Bognár 2006, Vass 2006, Lénárd–Rapos 2009), valamint számos kutatási projektben dolgoztak ki és alkalmaztak olyan módszereket és eszközöket, amelyek a tanítási-tanulási folyamatra fókuszáló értékelést támogatják (lásd Lénárd–Rapos 2006, Kerber 2011, Csapó–Zsolnai 2015). A tanárképzésben is hangsúlyos elemmé vált a pedagógiai értékelés és azon belül a fejlesztő értékelés megismertetése, differenciált értékelési szemlélet kialakítása (lásd Knausz 2008, Bárdossy–Dudás 2009, Molnár–Víg Tibor 2015).

A tanárok, illetve a tanárjelöltek szakmai fejlődését és tudását vizsgáló nemzetközi kutatásokban az *assessment literacy* kifejezés használatával főként az értékelés három fő célja mentén, a szummatív és a formatív, esetleg a diagnosztikus pedagógiai értékelési helyzetek megkülönböztetését tanulmányozzák (Kanjee–Mthembu 2015,

Volante–Fazio 2007), s e vonatkozásban vizsgálják az értékelési ismeretek, attitűdök és készségek működését. E terület a tanár képességét jelöli, amely révén alkalmazza az értékelési tudását értékelési feladatainak tervezésében, elvégzésében és továbbfejlesztésében (Fulmer–Lee–Tan 2015: 479). Ez tág értelmezési keretet jelent, hiszen nem csupán a tanulói és az osztályszintű értékelésről van szó, hanem ez magában foglalja az intézményi, az országos és a nemzetközi vizsgálatok eredményeinek értelmezését is a saját tanítási és osztálytermi környezetben. Így tehát ebben az értelmezésben a rendszerszemlélet érvényesítése is megjelenik, vagyis az értékelés kérdéseinek összefüggésekben való kezelése (lásd Xu–Brown 2016).

A pedagógiai folyamatok és a tanulók értékelésével kapcsolatos tanári kompetencia-területen elsajátítandó tudás jellemzésére, az értékelési komponensek meghatározására nemzetközi kontextusban már dolgoztak ki értékelési rendszereket (Kahl–Hofman–Bryant 2013), amelyek három fő területre, 1. az értékelés céljainak a megkülönböztetésére, 2. az értékelés minőségének megítélésére, valamint 3. az eredmények megfelelő felhasználására vonatkoznak. A pedagógussá válás és a szakmai fejlődés kutatásalapú standardjai hazánkban is ismertek (Kotschy 2011), amelyek a tanárképzés kimeneti követelményeiben (8/2013 I. 20. EMMI rendelet) is megjelennek. A formális képzés keretében elsajátítandó, a tanárjelöltek értékelési tudását leíró komponensekhez egyrészt a gyakorlatra bocsátás feltételeként az értékelés funkcióira, folyamatára, formáira, módszereire, a mérésmetodika alapjaira vonatkozó ismeretek és a tanulást támogató értékelés iránti elköteleződést mutató attitűdök tartoznak, másrészt a záróvizsgázott pedagógus szintjét jellemző komponensek, amelyek a tanulókra, a pedagógusra és a pedagógiai folyamatra vonatkozó értékelést jellemzik.

A tanárjelöltek értékelési kompetenciájának fejlődését, a megfelelő fejlesztést támogató oktatási környezet jellemzőit a nemzetközi kutatásokban többféle kutatási módszer együttes alkalmazásával tárják fel. Jellemzők az akciókutatások, amelyek keretében egyrészt az értékeléssel kapcsolatos ismeret- és képességjellegű tudást, másrészt az attitűdöket és az önértékelést vizsgálják (lásd Lukin et al. 2005, Volante–Fazio 2007, Levey–Vered–Alhija 2015). E vizsgálatok fő tanulságai szerint az értékelési kurzusoknak a tanárjelöltek fejlődésére általában kicsi a hatásuk, jellemzően a szummatív értékeléssel kapcsolatos, diákként szerzett tapasztalatok a meghatározók, a jelöltek az értékelést főként e területre szűkítve értelmezik, a formatív értékelés és egyéb alternatív értékelési módszerek kevésbé meghatározók. E tanulságok összhangban vannak az értékelési kompetenciára vonatkozó hazai empirikus kutatások eredményeivel (lásd Kovácsné Duró 2004, 2007), amelyek alátámasztják, hogy a tanárjelöltek értékeléssel kapcsolatos nézeteit a tradíciók, a diákként átélt tapasztalatok és minták nagymértékben meghatározzák, míg a tanárképzésben tanultak kevésbé.

Az akció kutatásunk jellemzői és módszerei

Ezen előzményekre építve az akció kutatásunkat egy magyarországi egyetem tanárképző programjának pedagógia-pszichológia moduljában szereplő pedagógiai értékeléssel kapcsolatos kötelező gyakorlat keretében végeztük 2018 tavaszán. A kurzus résztvevőinek a kutatást megelőző félévben, egy előadás keretében már volt alkalmuk megismerkedni a pedagógiai értékelés elméleti alapjaival. Ezért e kurzus alapvető célja, hogy a tanárjelölteket megismertesse az oktatási rendszer különböző szintjein felhasznált értékelési módszerekkel és eszközökkel, ezek jóságmutatóival és funkciójával, és képessé tegye őket az iskolai gyakorlatban alkalmazott értékelési eljárások önálló megválasztására, megítélésére és elemzésére, valamint önálló mérőeszközök fejlesztésére.

E cél elérése érdekében a kurzuson számos, főként formatív értékelési helyzetbe hoztuk a tanárjelölteket: a félévközi beadandó feladatok megoldása során többször kellett reflektálniuk a diákként szerzett értékeléssel kapcsolatos tapasztalataikra és tudásukra, a feladatokról az oktatótól kaptak visszajelzéseket, valamint az órán a formatív értékelés számos formáját alkalmaztuk ön- és társértékeléssel. A kurzus keretében a tanárszakosok önállóan fejlesztettek egy tantárgyi tudásszintmérő tesztet, amelyet két társuk előre megadott és megértett szempontok mentén értékelt, valamint azt a tanárjelölteknek az oktatói visszajelzések alapján is tovább kellett fejleszteniük.

A tanárjelöltek értékeléssel kapcsolatos gondolkodásának megismerése, valamint a tudásváltozásuk megítélésének vizsgálata céljából egy papíralapú kérdőívet készítettünk, amelyet a tanárszakosok a kurzus legelső és legutolsó óráján, közvetlenül a gyakorlati jegyük megismerése előtt töltöttek ki. Ennek része volt az alábbi kérdés: „Kérem, határozza meg, hogy a saját maga számára mit jelent a »tanulók értékelésének« fogalma, mit jelent a diákokat értékelni!” A kérdést azért így fogalmaztuk meg, mert nem az volt a szándékunk, hogy a tanárszakosok a képzésük során elsajátítandó definícióra gondoljanak, hanem az, hogy szabadon fogalmazhassák meg a válaszaikat. Ezáltal célunk volt annak vizsgálata, milyen értékelési szempontokra helyezik a hangsúlyt a meghatározásaikban a félév elején és végén, hogy így a gondolkodásukat és annak változását megismerjük. Mindezek alapján az értékelési kompetencia összetevői közül főként az ismereteket vizsgáltuk, de célunk volt az attitűdök elemzése is, ezért kértünk önértékelést: a félév végi kérdőívben egy nyílt végű kérdésre adott válaszban azt kellett a jelölteknek összegezniük, hogy megítélésük szerint miben és milyen mértékben fejlődtek, miben változott a tudásuk a kurzus hatására. Ezenkívül a félév végén két, ötfokú Likert-skálán megítték, hogy a kurzus kezdetén, majd visszatekintve, a végén mennyire látják komplex tevékenységnek az értékelést.

A kérdőívet öt, pedagógiai értékeléssel foglalkozó csoportban töltötték ki a tanárszakosok. A kurzus résztvevői főként másodéves, osztatlan tanárképzésben részt vevő, valamint kisebb részben osztatlan képzésben részt vevő tanárjelöltek voltak. Ezekben a

csoportokban két oktató tartotta a kurzust folyamatosan egyeztetve, azonos módszerekkel, feladatokkal és elvek alapján. A kérdőívet az osztott tanárképzés egy szabadon választott kurzusán is kitöltötték a tanárjelöltek, ennek célja értékelési helyzetek reflektív elemzése volt. A kérdőív anonim volt, a két időpontban történt adatfelvétel összekapcsolását a hallgatók kódja biztosította. A kérdőívet a félév elején 80-an, a végén 63-an töltötték ki, 59-en mindkét adatfelvételben részt vettek. A minta összetétele szakok szerint heterogén.

Az elemzés két fázisban történt. Először a tanárjelöltek által írt meghatározásokon végeztünk tartalomelemzést. Ezt követően azt elemeztük, hogy 1. mi jellemzi a tanárszakosok értékelésértelmezését a félév elején, 2. milyen változás azonosítható ehhez képest a félév végére, valamint 3. milyen hasonlóságok és különbségek vannak a meghatározás komplexitása és a tanárszakosok önértékelése között.

A tartalomelemzés kategóriái, szempontrendszere és megbízhatósága

A tanárjelöltek értékeléssel kapcsolatos meghatározásainak vizsgálatához egy differenciált, többféle területet, kategóriát és szempontot figyelembe vevő rendszert dolgoztunk ki. A tartalomelemzés során alkalmazott szempontrendszert az *1. táblázat* tartalmazza. A továbbiakban ezt jellemezzük részletesebben, majd bemutatjuk ennek alkalmazását és megbízhatóságát.

A tartalomelemzés eszközt a szakirodalmi háttérre alapozva, főként deduktív úton alakítottuk ki. A pedagógiai értékelés jellemzői alapján jelöltük ki az értékelési területeket és azokon belül a kategóriákat, illetve a tanárjelöltek értékelési kompetenciáinak összetevői szerint határoztuk meg a szempontokat. Ennek megfelelően négy értékelési területet definiáltunk, amelyek a *Miért? Mit? Hogyan?* és *Ki?* kérdőszavak köré rendezhetők. A *Miért?* esetén azt vizsgáltuk, hogy a tanárjelölt milyen célt fogalmaz meg, milyen szerepet szán az értékelésnek, a tanítási-tanulási folyamat melyik szakaszára vonatkoztatja. A *Mit?* kérdés esetén a besorolás szempontja az volt, hogy milyen tartalmat ad meg a tanárjelölt, melyik tudásterületre vonatkoztatja az értékelést. A *Hogyan?* vizsgálatához négy kategóriát vettünk figyelembe: aszerint történt a besorolás, hogy a tanárjelölt milyen módszert nevez meg, miként jellemzi az értékelési folyamatot, mihez viszonyítva történhet az értékelés, illetve a besorolás szempontja volt az is, hogy a tanárjelölt szerint milyen szempontokat, jószágmutatókat kell érvényesíteni az értékelés folyamatában. A *Ki?* kérdés esetén azt vizsgáltuk, hogy a meghatározás alapján ki végzi az értékelést, a tanárjelölt kire vonatkoztatja az értékelés eredményét, és a gyűjtött információkat ki hasznosítja. E szemponton belül az értékelés vonatkoztatása esetén nemcsak a személyre, hanem a tanítási-tanulási folyamat egyes elemeire (célokra, tartalomra és magára a folyamatra) vonatkozó értékelést is megkülönböztettük. A tanárszakosok válaszainak feldolgozása során szükségesnek bizonyult a szempontrendszer kiegészítése olyan megfogalmazásokkal, amelyekben a jelöltek csak utalnak valamilyen

Értékelési terület	Értékelési kategória	Szempont
Miért?	1. Célok	A tanárjelölt ...
		1. <i>utal</i> valamilyen célra, de nem határozza meg pontosan (pl. keveri az értékelési funkciókat, nem nevez meg érvényes célt).
		2. megnevezi a <i>diagnosztikus</i> értékelést, és/vagy példát ír rá.
		3. megnevezi a <i>formatív</i> értékelést, és/vagy példát ír rá.
Mit?	2. Tartalom	4. megnevezi a <i>szummatív</i> értékelést, és/vagy példát ír rá.
		5. a <i>tanuló személyét</i> nevezi meg.
		6. a <i>témát/tananyagot/teljesítményt</i> nevezi meg.
		7. a <i>tanuló magatartását/szorgalmát</i> nevezi meg.
		8. általánosan értelmezve a <i>tanuló tudását</i> nevezi meg.
		9. specifikusabban értelmezve a <i>tanuló előzetes tudását</i> nevezi meg.
		10. <i>ismeretet</i> nevez meg.
		11. <i>képességet/készséget</i> nevez meg.
		12. <i>affektív területet</i> (pl. motivációt, attitűdöt) nevez meg.
		13. a <i>tanuló munkáját</i> vagy valamilyen <i>produktumát</i> nevezi meg.
Hogyan?	3. Módszer	14. <i>utal</i> valamilyen értékelési módszerre, de nem pontosan (pl. nem nevezi meg konkrétan, vagy a példa nem sorolható egyik csoportba sem).
		15. <i>minőségi</i> értékeléshez kapcsolódó példát ír (pl. szöveges értékelés).
		16. <i>menyiségi</i> értékeléshez kapcsolódó példát ír (pl. tesztek, mérések).
	4. Jellemző	17. megnevezi az értékelés <i>szervezett/célorientált/tervezett</i> jellegét.
		18. <i>viszajelzéseként</i> értelmezi az értékelést.
	5. Viszonyítás	19. <i>utal</i> valamilyen, nem pontosan meghatározott viszonyítási lehetőségre.
		20. a <i>személyre szabott</i> értékelést nevezi meg, és/vagy példát ír rá.
		21. a <i>normaorientált</i> értékelést nevezi meg, és/vagy példát ír rá.
		22. a <i>kritériumorientált</i> értékelést nevezi meg, és/vagy példát ír rá.
		23. <i>utal</i> valamilyen, az értékelésre vonatkozó kritériumra.
		24. az <i>objektivitást</i> nevezi meg, és/vagy példát ír rá.
	6. Jóságmutatók	25. a <i>reliabilitást</i> nevezi meg, és/vagy példát ír rá.
26. a <i>validitást</i> nevezi meg, és/vagy példát ír rá.		
27. nyelviileg <i>utal</i> valamilyen személyre, de expliciten nem nevezi meg.		
28. a <i>tanárt</i> nevezi meg.		
7. Értékelő személye	29. a <i>diákot</i> nevezi meg.	
	30. <i>külső szereplőt</i> (pl. intézményi vagy országos szintet) nevez meg.	
	31. nyelviileg <i>utal</i> arra, hogy az értékelés irányul valakire.	
	32. az értékelést a <i>célokra</i> vonatkoztatja.	
	33. az értékelést a <i>tartalomra</i> (pl. tananyag, tudás) vonatkoztatja.	
	34. az értékelést a tanítási és/vagy a tanulási <i>folymatra</i> vonatkoztatja.	
	35. az értékelést a <i>tanárra</i> vonatkoztatja.	
	36. az értékelést a <i>diákra</i> vonatkoztatja.	
	37. az értékelést a diák <i>önértékelésére</i> vonatkoztatja.	
	38. az értékelést a <i>társértékelésre</i> vonatkoztatja.	
8. Értékelés vonatkoztatása	39. <i>utal</i> arra, hogy az információkat valaki felhasználhatja.	
	40. a <i>tanulót</i> nevezi meg az információ hasznosítójaként.	
	41. a <i>tanárt</i> nevezi meg az információ hasznosítójaként.	
	42. a <i>szülő</i> t nevezi meg az információ hasznosítójaként.	
	43. az <i>intézményt</i> nevezi meg az információ hasznosítójaként.	
	44. a <i>szakembereket</i> és/vagy az <i>országos szintet</i> nevezi meg az információ hasznosítójaként.	

1. táblázat. A tartalomelemzés kategóriái és szempontjai

értékelési célra, módszerre, az értékelő személyére. Így tehát a szempontokat induktív módon is alakítottuk, melynek eredményeként jött létre az összesen 9 kategóriát és 44 szempontot tartalmazó rendszer.

A kialakított szempontrendszer mentén azonosítottuk a fogalommeghatározásokban szereplő információkat, és ezeket egyenként soroltuk be a megfelelő kategóriákba és azokon belüli szempontokba. Minden besorolást 1-es kóddal jelöltünk az adattáblában. Ez tette lehetővé az egyes szempontok és kategóriák, illetve a tanárjelöltek fogalomértelmezésének számszerű vizsgálatát. A kilenc kategória elemzéséhez mindegyik esetében összevont mutatókat hoztunk létre, amelyekben azt összegeztük, hogy a tanárjelöltek meghatározásai az adott kategóriából hány szempontot tartalmaznak. A tanárjelöltek fogalommeghatározásának globális jellemzéseként egy komplexitás-indexet is létrehoztunk, amely azt fejezi ki, hogy a meghatározás hány szempontot tartalmaz. A továbbiakban értéketnek nevezzük azokat a számszerűsített eredményeket, amelyeket a besorolás eredményeként és az összevont mutatók képzésekor létrehoztunk.

A továbbiakban két példán szemléltetjük a tartalomelemzés kategóriáit, szempontjait és a besorolás módját. A 2. táblázat példaként egy tanárjelölt félév elején és végén megfogalmazott meghatározását tartalmazza. Ennek segítségével azt mutatjuk be, hogy a közlésben szereplő információkat melyik kategóriába és szempontba soroltuk. A harmadik oszlopban lévő szám az 1. táblázat alapján tartalmazza az egyes szempontok sorszámát. Az első példában a tanárjelölt célként a szummatív értékelést nevezi meg, és a tartalom esetén három szempontot, illetve a minőségi értékelést és az értékelést végző személyt. E meghatározás tehát összesen 6 elemet tartalmaz. A második példában a célok hiányoznak, a tartalom esetén a meghatározás írója két elemet említ, például hoz a mennyiségi és a minőségi értékelésre, valamint utal az értékelés jellemzőjére. Az értékelést a tanulóra vonatkoztatja és az eredmények felhasználója szerint több szereplőt is megemlít. Összesen 11 elemet tartalmaz ez a meghatározás.

Példa	Kategória	Szempont száma	Komplexitás
„A tanulók értékelése az, amikor a tanár valamilyen területen minősíti, véleményezi a teljesítményüket, tudásukat, képességüket.”	Értékelő személye	28.	6
	Célok és módszer	4., 15.	
	Tartalom	6., 8., 11.	
„A tanulókat többféle szempontból lehet értékelni (pl. előzetes tudás, tananyag elsajátítása) számszerűsítve vagy anélkül, hogy a tanuló, szülők, tanárok, intézmények és egyéb szakemberek képet kapjanak a felmérni kívánt dologról.”	Tartalom	6., 9.	11
	Módszer	15., 16.	
	Eredmények felhasználója	40., 41., 42., 43. 44.	
	Jellemző	18.	
	Értékelés vonatkoztatása	31.	

2. táblázat. Példák a szempontrendszer alapján történő besorolásra

Mivel a szempontrendszerben több kategóriát és szempontot kellett egyszerre figyelembe venni, ezért a közös értelmezést követően, egymástól függetlenül mindketten besoroltuk a tanárjelöltek meghatározásaiban szereplő információkat a megadott szempontokba, így ebben az értelemben végeztünk kettős értékelést. Ezt követően egyezésvizsgálattal ellenőriztük a szempontrendszer működését. E vizsgálatot a 9 kategória és a komplexitásindex mentén létrehozott összevont mutatók szerint végeztük. A 3. táblázat tartalmazza mindkét mérési pontban a kategóriák és a komplexitásindex alapján a két értékelő besorolása közötti korrelációs együtthatókat és az egyes kategóriák kettős értékelésének megbízhatóságát.

Megjegyzés: minden korrelációs együttható szignifikáns $p < 0,001$ szinten.

Értékelési kategória	Szempontok száma	Korrelációs együttható (r-érték)		Megbízhatóság (Cronbach-alfa)	
		Félév elején	Félév végén	Félév elején	Félév végén
Célok	4	0,64	0,87	0,77	0,93
Tartalom	9	0,80	0,89	0,89	0,94
Módszer	3	0,63	0,87	0,78	0,93
Jellemző	2	0,83	0,88	0,91	0,93
Viszonyítás	4	0,86	0,77	0,92	0,86
Jóságmutatók	4	0,65	0,93	0,78	0,96
Értékelő személye	4	0,74	0,66	0,85	0,79
Értékelés vonatkoztatása	8	0,53	0,35	0,64	0,47
Eredmények felhasználója	6	0,63	0,82	0,77	0,90
Komplexitás		0,81	0,89	0,89	0,95

3. táblázat. Az értékelők közti egyezésvizsgálat eredménye az első és a második mérési ponton

A korrelációs együttható -1 és 1 közötti érték, az értékek közti kapcsolat irányát és mértékét jellemzi. E mutató alapján azt várjuk, hogy a két értékelő között szoros és pozitív irányú a kapcsolat. Mindkét adatfelvételi ponton a komplexitásindex és a 9 kategória szerint is szoros egyezést találtunk ($p < 0,001$), így e szempont teljesült, az egyes kategóriák között a korreláció erősségében azonban vannak eltérések ($0,35 \leq r \leq 0,89$). Az egyes kategóriák kettős értékelésének megbízhatóságát mutatja a Cronbach-alfa érték, amely egy 0 és 1 közötti mutató, és minél közelebb van 1-hez, annál pontosabban történt az adott kategória értékelése. E mutató alapján az egyes kategóriák értékelésének megbízhatósága elfogadható, a Cronbach-alfa értékek jellemzően 0,8 fölöttiek.

A szempontok többsége a kettős értékelés során az elvártan megfelelően működött, az értékelés vonatkoztatását tekintve azonban alacsonyabb a korrelációs együttható mindkét mérési ponton, valamint e kategória esetében alacsony a megbízhatóság is. Ez részben adódik a szempontrendszer komplexitásából, hiszen e kategóriát nyolc

szempontból, differenciáltan kellett értékelni, amelyben nemcsak a személyre, hanem a tanítási-tanulási folyamat egyes elemeire vonatkozóan is el kellett végezni a besorolást. Másrészt pedig e kategóriába főként olyan közléseket kellett besorolni, amelyekben a tanárjelöltek inkább csak nyelviileg utaltak arra, hogy az értékelés kire vonatkozik, ami eredményezett eltéréseket a két értékelő között. Harmadrészt pedig e kategórián belül voltak olyan szempontok is (pl. a célokra, az ön- és társértékelésre vonatkozó értékelés), amelyekbe történő besorolás alig volt jellemző, ami befolyásolhatja a megbízhatóságot. Mindezek alapján szükségesnek bizonyult a kérdéses esetek kezelése, amelynek érdekében a legnagyobb eltérést mutató besorolásokat egyeztetettük, illetve a két értékelő értékleit átlagoltuk, majd ezek alapján számoltuk ki és elemeztük az eredményeket.

Az értékelésről alkotott meghatározások a két mérési ponton

Elemzésünkben először az értékelésértelmezést vizsgáltuk meg a félév eleji és végi mérés eredményei alapján. Célunk volt a félév elején írt meghatározások és az ehhez viszonyított változás főbb tendenciáinak jellemzése. E cél elérése érdekében meghatároztuk, hogy a tanárjelöltek válaszában jelentősebb része melyik szempontokba sorolható a két mérési ponton, azaz a félév elején és végén.

A 4. táblázat az említések relatív (százalékos) gyakoriságát tartalmazza, amelyben az eredmények további értelmezéséhez megadtuk azt is, hogy az adott szempont melyik kategóriába tartozik. Eredményeink szerint a tanárjelöltek értékelésértelmezésének kétharmada 11 szemponttal írható le mindkét mérési ponton.

A félév elején a tanárszakosok szűken értelmezték a fogalmat: többségében a tanulóra irányuló, többnyire pontatlanul meghatározott cél érdekében, a tantárgyi teljesítményt vagy tudást vizsgáló visszajelzésként értelmezték, amelyről a tanár mennyiségi értékelése alapján a tanuló szerez információt, ami a tanítási-tanulási folyamatra vonatkoztatható. A félév végén ehhez nagyon hasonló meghatározást adtak: most is főként a tanulóra irányuló visszajelzésként értelmezték, amit formatív vagy szummatív értékelési célból a tanár végez a tanuló tudását vagy tantárgyi teljesítményét mennyiségi értékeléssel vizsgálva, annak érdekében, hogy a tanuló tájékoztatást szerezzen a tanítási-tanulási folyamatról.

Az értékelésről alkotott meghatározások változása

Megvizsgáltuk azt is, hogy kategóriák szerint miként jellemezhető a tanárjelöltek értékelésértelmezése, mire helyezik a hangsúlyt és mire kevésbé, valamint összehasonlítottuk a változást a két adatfelvétel során.

Értékelési kategória	Szempont A tanárjelölt ...	Adatfelvétel	
		Félév elején	Félév végén
Célok	<i>utal</i> valamilyen célra, de nem határozza meg pontosan.	7	4
	megnevezi a <i>formatív</i> értékelést, és/vagy példát ír rá.	3	7
	megnevezi a <i>szummatív</i> értékelést, és/vagy példát ír rá.	3	4
Tartalom	a <i>témát/tananyagot/teljesítményt</i> nevezi meg.	7	4
	általánosan értelmezve a <i>tanuló tudását</i> nevezi meg.	7	6
Módszer	<i>menyiségi</i> értékeléshez kapcsolódó példát ír (pl. tesztek, mérések).	6	5
Jellemző	<i>visszajelzéseként</i> értelmezi az értékelést.	7	8
Értékelő személye	nyelviileg <i>utal a tanárra</i> , vagy megnevezi a <i>tanárt</i> .	6	6
Értékelés vonatkoztatása	nyelviileg <i>utal a diákra</i> , vagy az értékelést a <i>diákra</i> vonatkoztatja.	16	14
	az értékelést a tanítási és/vagy a tanulási <i>folyamatra</i> vonatkoztatja.	4	5
Eredmények felhasználója	a <i>tanulót</i> nevezi meg az információ hasznosítójaként.	6	4
Összesen		72	67

4. táblázat. A leggyakoribb szempontok relatív gyakorisága (%) az első és a második mérési ponton

Az 5. táblázatban a kategóriák szerint csökkenő sorrendben adtuk meg a közlések arányát. Ez alapján láthatjuk, hogy nincs jelentős elmozdulás a két mérési pont között. A közlések többsége három kategóriába sorolható, amelyek az értékelés vonatkoztatása, az értékelés tartalma és célja. Az értékelés módszere, az információ hasznosítója, az értékelés jellemzője és az értékelést végző személy szempontjából nagyon hasonló a félév elején és végén az arány. E kategóriákhoz képest az értékelés viszonyítása és a jóságmutatók megnevezése egyik mérési ponton sem volt hangsúlyos.

Értékelési kategória	Adatfelvétel	
	Félév elején	Félév végén
Értékelés vonatkoztatása	24	22
Tartalom	21	21
Célok	13	16
Módszer	10	11
Eredmények felhasználója	10	9
Jellemző	8	9
Értékelő személye	6	6
Viszonyítás	5	4
Jóságmutatók	3	2

5. táblázat. Az értékelésre vonatkozó kategóriákba történő besorolás relatív gyakorisága (%) az első és a második mérési ponton

Mindkét adatfelvétel során a legnagyobb hangsúly azon van, hogy kire irányul az értékelés. Ahogyan a 4. táblázatban már láthattuk, a tanárszakosok főleg a tanulót jelölték meg (a félév elején a közlések 16, a végén 14%-ában), de az ön- és társértékelés támogatása egyáltalán nem jelenik meg a meghatározásaikban, a célokra, a tartalomra, a tanárra kismértékben (csupán a közlések 1-2%-ában) gondolnak. A tanuló személyére vonatkozó értékelésben nagyon kismértékű (1%-nyi) csökkenés figyelhető meg a félév végi közlésekben.

Hangsúlyos a tanárjelöltek által írt meghatározásokban az értékelés tartalma: a félév elején és végén a közléseik 21%-ában ez jelenik meg, ugyanakkor a két mérési ponton némi átrendeződés megfigyelhető: a félév végén a tananyag és a tanuló tudásának általános megnevezése kisebb arányú a félév elejéhez képest (lásd 4. táblázat), és hangsúlyosabbá válik az előzetes és a képesség jellegű tudás, valamint a tanuló produktumának értékelése (ez a félév végén a közlések 2-3%-ában jellemző); ugyanakkor az affektív terület (pl. a motiváció, tantárgyi attitűdök, tanulói szokások) értékelése egyáltalán nem jelenik meg a fogalom meghatározásokban.

Az értékelés céljainál jelentősebb változás van, e kategória néhány elemében a két adatfelvétel között jelentős eltérés figyelhető meg. A 6. táblázat azon tanárjelöltek esetén mutatja e szempont szerinti besorolások átlagait, akik mindkét adatfelvételben részt vettek (N = 59). A páros t-próba alapján, a félév végére szignifikánsan csökkent a pontatlan célmeghatározás: míg a félév elején a minta 42%-ában volt jellemző, addig ez 21%-ra csökkent. A diagnosztikus és a formatív értékelés említése gyakoribb a második adatfelvételnél, tehát a tanárjelöltek a félév végén jobban kiemelték ezt a két célt, 12 és 40%-ra volt jellemző ennek említése. A szummatív értékelés esetén nem volt szignifikáns a különbség a két adatfelvétel között, vagyis hasonló arányban említették ezt az értékelési funkciót (lásd még 4. táblázat).

Szempont	Említések átlaga (%)		Különbség a félév elejéhez képest	
	Félév elején	Félév végén	t	p
A tanárjelölt ...				
utal valamilyen célra, de nem határozza meg pontosan.	42	21	3,09	0,003
megnevezi a <i>diagnosztikus</i> értékelést, és/vagy példát ír rá.	3	12	2,65	0,010
megnevezi a <i>formatív</i> értékelést, és/vagy példát ír rá.	19	40	3,09	0,003
megnevezi a <i>szummatív</i> értékelést, és/vagy példát ír rá.	14	25	1,82	0,074

6. táblázat. Különbségek az értékelési célokhoz tartozó szempontok említésének átlagában a két mérési ponton (N = 59).

Az értékelés módszerei közül mindkét mérési ponton a mennyiségi értékelés hangsúlyos; a félév végén megjelenő 1%-os csökkenés (4. táblázat) mellett a minőségi értékelés a félév végén a közlések 3%-ra volt jellemző, ami azt jelzi, hogy a tanárjelöltek

bár nem neveznek meg több módszert a félév végén, de a meghatározásaikban megjelenik, hogy a dolgozatokon, osztályzatokon kívül más értékelési módszerek is rendelkezésre állhatnak.

A tanárjelöltek általában a tanulót nevezik meg az értékelés során gyűjtött információ hasznosítójaként (4. táblázat). Ugyanakkor a félév végén az első méréshez képest szignifikánsan nagyobb azon tanárszakosok aránya, akik kiemelik az eredmények felhasználását és visszacsatolását az intézmények számára ($t = 2,1$; $p = 0,04$). A többi szempont esetében nem találunk jelentős eltérést a két adatfelvétel között.

Az értékelés jellemzőjeként főként a visszajelzést emelik ki a tanárszakosok. Az értékelés céltudatos, tervszerű jellegét a félév elején egyáltalán nem említik, a félév végén is csak kis arányban (a közlések 1%-ában).

A jelöltek közléseiben tipikusan a tanár az értékelést végző személy; ebben nincs elmozdulás a két mérési ponton. A tanuló vagy külső szakemberek által végezhető értékelés lehetősége egyáltalán nem jelenik meg a válaszaikban.

Az értékelés viszonyítása szempontjából elsősorban a tanuló korábbi teljesítményéhez való viszonyítást emelik ki a tanárjelöltek. A norma- és kritériumorientált értékelés lehetősége alig (csupán a közlések 1-2%-ában) jelenik meg a meghatározásaikban.

A jószágmutatók kapcsán az objektivitás kritériumát emelik ki főleg úgy, hogy ne legyen szubjektív az értékelés. A félév végén a közlések 1%-ában megjelenik a célokkal összhangban a validitás, de a tanulói értékelés megbízhatóságát jelölő reliabilitás egyáltalán nem.

Az értékelésről alkotott meghatározások komplexitásának jellemzői és a változás önértékelése

Az elemzés utolsó fázisában megvizsgáltuk az összevont mutatóként képzett komplexitásindex szerinti különbséget, amit összevetettünk a tanárjelöltek által érzékelt tudásváltozással. A tanárszakosok fogalomértelmezésének komplexitásában nem találtunk szignifikáns különbséget a két adatfelvétel között ($t = 0,17$; $p = 0,87$), aminek háttérében az áll, hogy a minta 49%-ában az első adatfelvételkor volt a meghatározás komplexebb, és csupán a minta harmadánál volt a második komplexebb (7. táblázat). Mindkét részmintában három szempont szerint azonosítottunk szignifikáns eltérést ($p < 0,05$) a két adatfelvétel során adott megfogalmazás között. A továbbiakban ezeket az eltéréseket jellemezzük. Azon tanárszakosok körében, akiknél az első meghatározás volt komplexebb, azt találtuk, hogy a félév elején 50%-ukra volt jellemző, hogy utaltak valamilyen célra, 95%-uk az értékelést a tanulóra vonatkoztatta, és a tanulót nevezte meg az információ hasznosítójaként (ezt a szempontot e rész minta 41%-a említette). E csoport fogalomértelmezésének változását azt jelzi, hogy ezeken a területeken a félév végére szignifikáns csökkenés volt tapasztalható, vagyis ezeket a szempontokat kisebb arányban említették. Azon tanárszakosok, akiknek a második fogalom meghatározása

komplexebb, az jellemző, hogy a formatív értékelést a félév elején e rész minta 18%-ában, a végén már 50%-ban említették; a minőségi értékelés megnevezése is hangsúlyosabbá vált a félév végén (ezt e rész minta 37%-a említette, míg a félév elején ez az arány 11% volt). Ezenkívül hangsúlyosabbá vált az utalás valamilyen, nem pontosan meghatározott viszonyítási lehetőségre (a félév elején ez a rész minta 3%-ára, a végén 24%-ra jellemző).

Eltérés	Elemzés	Arány (%)
Az első meghatározás komplexebb	29	49
Nincs különbség a két meghatározásban	11	19
A második meghatározás komplexebb	19	32

7. táblázat. A komplexitásindexben lévő különbségek eloszlása

A félév végi kérdőívben azt kértük a tanárjelöltektől, hogy két, ötfokú Likert-skálán ítélik meg, hogy az értékelést a második adatfelvételkor mennyire tartják komplexnek, és visszagondolva, a félév elején mennyire tartották komplexnek (8. táblázat). Eredményeink szerint szignifikáns különbség van az értékelés komplexitásának múltbéli és jelenlegi megítélése között, vagyis a félév végén komplexebb tevékenységnek ítélik meg az értékelést a félév elejéhez képest: míg a visszatekintésre vonatkozó kérdőív-tételnél az ötfokú skálán közepes az átlag, a félév végére jelentősebb változásról nyilatkoznak a tanárjelöltek; a szórások is eltérnek, a minta homogénebb a második kérdőív-tételre adott válaszok esetében. Ugyanerre a következtetésre jutunk, ha a 7. táblázatban lévő három rész minta szerint vizsgáljuk az értékelés komplexitásának megítélését. Mindez azt mutatja, hogy bár a tanárszakosok érzékelik az értékelési szemléletük differenciálódását, de ez nem jelenik meg a fogalomértelmezésük komplexitásában.

Kérdőív-tétel	Átlag	Szórás	t	p
Visszatekintve, a félév elején mennyire látta komplex tevékenységnek az értékelést?	3,03	1,02	9,61	< 0,001
Most, a félév végén mennyire látja komplex tevékenységnek az értékelést?	4,64	0,64		

8. táblázat. Az értékelés komplexitásának megítélése

A félév végén kitöltött kérdőívben egy nyílt végű kérdésben azt kértük a válaszadóktól, hogy összegezzék, miben látják a változást (9. táblázat). A tanárjelöltek válaszaikat inductív módon tartalomelemzést végeztünk, és azokat két fő kategóriába soroltuk. Ezek egyrészt az értékeléssel kapcsolatos részterületekre és az érzékelt fejlődés

egyes aspektusainak megnevezésére vonatkoznak. A válaszok kétharmada alapján a tanárjelöltek leginkább az értékelési folyamat megismerésében (24%), a fogalmi gazdagodásban és pontosságban (14%), a tanári szempont érvényesítésében (10%), az értékeléssel kapcsolatos tudás és gondolkodás fejlődésében (9%), illetve az elmélet és a gyakorlat összekapcsolásában, és az értékeléssel szemben támasztott kritériumok fontosságának megismerésében (8%) érzékelték eltérést. Nagyrészt tehát inkább ismeretjellegű tudásváltozásról számoltak be, ami magyarázhatja, hogy a meghatározásaik komplexitása csak néhány területen azonosítható.

Kategória	Szempont	Arány (%) az összes említéshez képest
Értékeléssel kapcsolatos tudásterületek	értékelés céljainak, funkcióinak megismerése	7
	értékeléssel szemben támasztott kritériumok, jóságmutatók fontosságának megismerése	8
	értékelési módszerek megismerése	3
	értékelési eszközök megválasztása, megítélése, elemzése, eredmények értelmezése	4
	értékelési folyamat lépéseinek megismerése	28
Érzékelt fejlődés az értékelés területén	értékelés komplexitásának, összetettségének megismerése	2
	értékeléssel kapcsolatos tudás, gondolkodás általános fejlődése	9
	több fogalom, ismeret, információ, fogalmi pontosság	14
	elmélet-gyakorlat kapcsolása, gyakorlatiasabb tudás, és annak alkalmazása	8
	értékelési tudás rendszerezése, mélyítése	3
	tanári nézőpont érvényesítése, későbbi pedagógiai munka támogatása	10
	értékelés, visszacsatolás fontosságának felismerése, változás a hozzáállásban	3

9. táblázat. A tanárjelöltek által érzékelt tudásváltozás területei az említések relatív gyakorisága alapján (%)

Összegzés

Akciókutatásunk célja az volt, hogy egy pedagógiai értékeléssel kapcsolatos gyakorlati kurzus keretében kövessük nyomon, hogy a tanárszakosok miként értelmezik a tanulói értékelést a félév elején és végén, mely elemekre helyezik a hangsúlyt, és hogyan változik a meghatározásuk, valamint ennek komplexitása mennyire van összhangban az önértékelésükkel. E célok elérése érdekében egy, az értékelési kompetenciát vizsgáló papíralapú kérdőív részeként kértük a tanulói értékelés meghatározását a félév elején és végén, amelyen tartalomelemzést végeztünk. Ezt követően vizsgáltuk a kidolgozott szempontrendszer megbízhatóságát, és a tanárjelöltek értékelési fogalomértelmezésének változását az önértékelésükkel összehasonlítva.

A tartalomelemzés részeként egy szempontrendszer alapján két, független értékelő sorolta be a tanárszakosok meghatározásában szereplő közléseket tartalmi kategóriákba és a hozzájuk tartozó szempontokba. Az értékelésre vonatkozó kategóriák szerint szoros egyezést találtunk a besorolások között, a kettős értékelés megbízhatónak bizonyult.

Az akciókutatásunk céljaira reflektálva megállapítható, hogy a tanárjelöltek az értékelést a félév elején leginkább a tanulóra irányuló, a tantárgyi teljesítményt/tudást vizsgáló visszajelzésként, pontatlan célmeghatározással értelmezték, amelyről a tanár mennyiségi értékelése alapján a tanuló szerez információt. A tanárszakosok értékelésről való gondolkodásában van némi elmozdulás ettől a szűk fogalomértelmezéstől a komplexebb felé, de ez csak néhány területen, főként az értékelési céloknál tapintható ki. Az önértékeléssel való összevetés alapján megállapítható, hogy a jelöltek érzékelik az értékelési szemléletük differenciálódását az értékelés célját és folyamatát tekintve, valamint a rendszerszemlélet érvényesülése is megjelenik abból a szempontból, hogy a félév végén többen említették az eredmények felhasználását és visszacsatolását az intézmények számára.

Az akciókutatásunk eredményei összhangban vannak a szakirodalmi háttérben röviden vázolt vizsgálatok főbb tanulságaival: a tanárjelöltek alapvetően szűken értelmezik az értékelést, a meghatározásaik hangsúlyaiban jelentősen érződnek a közoktatásban átélt, diákszerepből adódó tapasztalataik, ezekhez képest a tanári szempont érvényesítése háttérbe szorul. Az értékelési szemlélet differenciálódásában az értékelési kurzus hatása jellemzően kicsi, bár ennek céljai szerint tapasztalhatók változások a tanárszakosok megnyilatkozásában, az oktatói alapelvek visszatükröződtek az értékelésértelmezésekben: a kurzuson nagy hangsúly került a különböző értékelési célok differenciált értelmezésére, az értékelési folyamat lépéseinek megismertetésére, a tesztkészítéssel ennek gyakorlására, a tesztek társértékelésével a formatív értékelés alkalmazására. Ezek a szempontok a félév végén több esetben megjelentek a tanárjelöltek megnyilatkozásaiban. A két mérési ponton végzett vizsgálat alapján azonban azt tapasztaltuk, hogy a változás lassú és az értékelési szemlélet nehezen differenciálható, de a tanárszakosok érzékelik és érvényesnek fogadják el ezt a szándékot.

Mindezek alapján elemzésünk fontos tanulsága, hogy a differenciált értékelési szemlélet kialakításához a tanárképzés során további, célirányos támogatásra, és a fejlesztés hatékonyságának nyomon követésére van szükség. Eredményeink a tanárjelöltek, a pedagógusok és a pedagógiai értékelést oktatók számára lehetnek jelentősek. Az értékelési kompetencia, s ilyen módon az értékelés fogalmának széles spektrumú értelmezése a leendő tanárok hatékonyabb működése szempontjából fontos. Ez a tudás és képesség feltételezi annak megértését, hogy az értékelés funkcióinak és a létrejött információk felhasználási lehetőségeinek háttérében tudatos és reflektív döntések húzódnak.

IRODALOM

- Anderson, L. W. – Krathwohl, D. R. (2000, szerk.): *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. London: Pearson.
- Bárdossy Ildikó – Dudás Margit (2009): *A tanulás tervezése és értékelése*. Budapest: Edukáció Társadalmi Szolgáltató Nonprofit Kft.
- Báthory Zoltán (2000): *Tanulók, iskolák – különbségek: Egy differenciális tanuláselmélet vázlat*. Budapest: OKKER Oktatási Kiadó.
- Bloom, B. (1956): *Taxonomy of educational objectives*. David McKay Company: New York.
- Bognár Mária (2006): A fejlesztő értékelés osztálytermi gyakorlata. *Új Pedagógiai Szemle*, 56/3, 19–26, <http://epa.oszk.hu/00000/00035/00101/2006-03-ko-Bognar-Fejlesztő.htm>
- Brassói Sándor – Hunya Márta – Vass Vilmos (2005): A fejlesztő értékelés: az iskolai tanulás minőségének javítása. *Új Pedagógiai Szemle*, 55/7–8, 4–17, <http://epa.oszk.hu/00000/00035/00094/2005-07-ta-Tobbek-Fejlesztő.html>
- Csapó Benő (2004): Tudásszintmérő tesztek. In: Falus Iván (szerk.) *Bevezetés a pedagógiai kutatás módszereibe*. Budapest: PSZMP – Keraban Kiadó, 277–316.
- Csapó Benő (2005): *Az előzetesen megszerzett tudás mérése és elismerése*. Budapest: Nemzeti Felnőttképzési Intézet, www.edu.u-szeged.hu/~csapo/publ/CSB_ElozetesTudas.pdf
- Csapó Benő – Zsolnai Anikó (2015, szerk.): *Online diagnosztikus mérések az iskola kezdő szakaszában*. Budapest: Oktatáskutató és Fejlesztő Intézet, http://publicatio.bibl.u-szeged.hu/5989/1/2015_OnlineDiagnosztikus.pdf
- Csikos Csaba (2002): A pedagógiai értékelés új irányzatai. *Új Pedagógiai Szemle*, 52/7–8, 175–179, <http://epa.oszk.hu/00000/00035/00062/2002-07-vt-Csikos-Pedagogiai.html>
- Fulmer, G. W. – Lee, I. C. H. – Tan, K. H. K. (2015): Multi-level model of contextual factors and teachers' assessment practices: an integrative review of research. *Assessment in Education: Principles, Policy & Practice*, 22/4, 475–494, <https://doi.org/10.1080/0969594X.2015.1017445>
- Golnhofer Erzsébet (2003): A pedagógiai értékelés. In: Falus Iván. (2003, szerk.): *Didaktika. Elméleti alapok a tanítás tanulásához*. Budapest: Nemzeti Tankönyvkiadó, 386–416.
- Guerriero, S. (2017, eds.): *Pedagogical knowledge and the changing nature of the teaching profession*. OECD, http://www.iep.edu.gr/images/IEP/EPISTIMONIKI_YPIRESIA/Epist_Grafeia/EU_Policy/2017/2017-05-11_OECD-Pedagogical-Knowledge.pdf
- Kahl, S. R. – Hofman, P. – Bryant, S. (2013): *Assessment literacy standards and performance measures for teacher candidates and practicing teachers*. Dover, New Hampshire: Measured Progress, <http://caepnet.org/~media/Files/caep/standards/assessment-literacy-in-teacher-preparati.pdf?la=en>
- Kanjee, A. – Mthembu, J. (2015): Assessment literacy of foundation phase teachers: An explanatory study. *South African Journal of Childhood Education*, 5/1, 125–141. DOI: 10.4102/sajce.v5i1.354
- Kerber Zoltán (2011, szerk.): *Sokarcú implementáció: Kompetenciafejlesztő programcsomagok bevezetése a HEFOP 3.1.3. iskolákban*. Budapest: Oktatáskutató és Fejlesztő Intézet, http://www.ofi.hu/sites/default/files/attachments/sokarcu_implementacio.pdf

- Kotschy Beáta (2011, szerk.): *A pedagógussá válás és a szakmai fejlődés sztenderdjei*. Eger: EKF, http://www.epednet.ektf.hu/eredmenyek/a_pedagogussa_valas_es_a_szakmai_fejlo-des_sztenderdjei.pdf
- Kotschy Beáta (2016): A pedagógiai értékelés kettős arca. In: Garai Imre – Vincze Beatrix – Szabó Zoltán András (szerk.) *Hiteles pedagógia. Tanulmányok Golnhofer Erzsébet tiszteletére*. Budapest: ELTE Eötvös Kiadó. 28–35, http://real.mtak.hu/42333/1/Hiteles_pedagogia_Golnhofer_.pdf
- Kovácsné Duró Andrea (2004): Tanár szakos egyetemi hallgatók értékelési tudása és kompetenciái. *Magyar Pedagógia*, 104/2, 203–224, http://www.magyarpedagogia.hu/document/Kovacsne%20Duro_MP1042.pdf
- Kovácsné Duró Andrea (2007): Miskolci és pécsi tanárjelöltek értékelési ismeretei és attitűdjei. *Iskolakultúra*, 17/1, 135–142, http://real.mtak.hu/57245/1/EPA00011_iskolakultura_2007_01_135-142.pdf
- Knausz Imre (2008): *Mit kezdjünk az értékeléssel? Adalékok az integrációs nevelés pedagógiájához*. Budapest: Educatio Társadalmi Szolgáltató Közhasznú Társaság, https://tanitonline.hu/ftp/Mit_kezdzunk_az_ertekelessel.pdf
- Levy-Vered, A. – Alhija, F. N. (2015): Modelling beginning teachers' assessment literacy: the contribution of training, self-efficacy, and conceptions of assessment. *Educational Research and Evaluation*, 21/5–6, 378–406, DOI: 10.1080/13803611.2015.1117980
- Lénárd Sándor – Rapos Nóra (2006, szerk.): *MAGTÁR III.: Ötletek tanítóknak a fejlesztő értékeléshez és az adaptív tanulásszervezéshez*. Budapest: Országos Közoktatási Intézet.
- Lénárd Sándor – Rapos Nóra (2009): *Fejlesztő értékelés*. Budapest: Gondolat Kiadó.
- Lukin, L. E. – Bandalos, D. L. – Eckhout, T. J. – Mickelson, K. (2005): Facilitating the Development of Assessment Literacy. *Educational Measurement: Issues and Practice*, 23/2, 26–32, DOI:10.1111/j.1745-3992.2004.tb00156.x
- Molnár Edit Katalin – Vigh Tibor (2015): *A tantervelmélet és pedagógiai értékelés alapjai*. Szeged: Szegedi Tudományegyetem. http://www.jgypk.hu/mentorhalo/tananyag/Tantervelmlet_s_a_pedagogiai_rtkels_alapjai/index.html
- Scriven, M. (1967): The methodology of evaluation. In: Tyler, R. W. – Gagne, R. M. – Scriven, M. (eds.) *Perspectives of Curriculum Evaluation*. Chicago: Rand McNally, 39–83.
- Vass Vilmos (2006): A fejlesztő értékelés nemzetközi tendenciái. *Új Pedagógiai Szemle*, 56/3, 15–18, <http://epa.oszk.hu/00000/00035/00101/2006-03-ko-Vass-Fejleszto.html>
- Vidákovich Tibor (1990): *Diagnosztikus pedagógiai értékelés*. Budapest: Akadémiai Kiadó.
- Volante, L. – Fazzio, X. (2007): Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, 30/3, 749–770, <https://files.eric.ed.gov/fulltext/EJ780818.pdf>
- Xu, Y. – Brown, G. T. L. (2016): Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162.
- 8/2013. (I. 30.) EMMI rendelet a tanári felkészítés közös követelményeiről és az egyes tanárszakok képzési és kimeneti követelményeiről https://net.jogtar.hu/jr/gen/hjegy_doc.cgi?docid=a1300008.emm

EINHORN ÁGNES

Fejlesztő értékelés a felsőoktatásban - módszertani mozaik

Formative assessment in higher education – a methodological review

The paper presents feasible methods and techniques of formative assessment with a focus on higher education. The Centre for Modern Languages at the University of Technology and Economics, Budapest provides the backdrop for the methodological development scheme where the core elements were a learner-centred approach and formative assessment. During our work it was important to prepare guidelines for teaching methods and to involve pilot groups that could facilitate classroom observation by fellow teachers.

This paper also explores the features of the summative and formative assessment, followed by a brief summary of the features of the pedagogical context applied. The conclusion details the tools applied and experiences gathered throughout lessons further supplemented with student and teacher reflections.

Keywords: higher education, formative assessment, gamification, learning outcomes, error culture

Az elmúlt évtizedek egyik jelentős pedagógiai újdonsága a fejlesztő szemléletű értékelés, egyre több fejlesztési projekt és kísérlet található ebben a témában (pl. OECD 2010, Prievara 2015). Előnyeként általában azt emelik ki, hogy erősíti a tanulók motivációját és autonómiáját, továbbá egyértelműbbé teszi a tanulói felelősséget az elért eredményekben, és így támogatja a tanulók bevonódását a tanítási folyamatba (Black–Wiliam 1998, Gardner et al. 2008). A tanulóközpontú értékelés iránti igény természetes következménye az általános pedagógiai modernizációnak: a tanuló fejlődését középpontba helyező pedagógiai kultúrának része a korszerű értékelési kultúra, azaz a fejlesztést támogató értékelési eljárások alkalmazása. A fejlesztő értékeléshez kapcsolódó egyszerűbb vagy komplexebb kísérletek általában a köznevelésben jelennek meg, mivel azonban az alkalmazásától várható többlet minden oktatási területen értékes lehet, ez a szemlélet a felsőoktatásban is alkalmazható. A felsőoktatási közeg pedagógiai szempontból sajátos: a célcsoport életkorából fakadóan felnőtt, ám mégis átmeneti korosztálynak tekinthető, a felmerülő pedagógiai problémák tehát sok szempontból hasonlítanak a köznevelésben jelentkező pedagógiai kihívásokra.

A tanulmányban a fejlesztő értékelés módszertani lehetőségeit és technikáit gyűjtöm össze felsőoktatási környezetben, ennek háttere a Budapesti Műszaki és Gazdaságtudományi

Egyetem Idegen Nyelvi Központjában folyó módszertani fejlesztés. A központban olyan nyelvi kurzusokat kínálunk, amelyek kiegészítik a hallgatók magyar vagy idegen nyelvű szakmai képzését, ezek a tantárgyak tehát az idegen nyelvi kompetenciák fejlesztését szolgálják, többnyire a szakmai nyelvhasználathoz kapcsolódóan. A bemutatandó gyakorlat szempontjából fontos tényező, hogy a BME egy komoly elvárásokat támaztó egyetem, ahol a hallgatók erős szakmai terhelés mellett veszik fel ezeket a többnyire választható idegen nyelvi tárgyakat. Ha tehát egy hallgató a túlterheltsége miatt fontossági sorrend felállítására kényszerül, akkor az idegennyelv-tanulásra fordít kevesebb energiát. A hallgatók motiválása, a bevonódásuk erősítése ebben a közegben a sikeresség egyik kulcseleme, ez tehát a fejlesztő értékelés mellett szól.

A tanulmányban módszertani megoldások leírására vállalkozom, az eredményesség igazolására vagy az eszközrendszer igazolható pedagógiai hatásának elemzésére nem kerül sor. A minősítő és a fejlesztő értékelés jellemzőinek feltárása után röviden leírom a pedagógiai környezet jellemzőit, majd az alkalmazott eszközöket és a tanórai megvalósítás során szerzett tapasztalatokat.

A minősítő és a fejlesztő értékelés

A napi pedagógiai gyakorlatban a minősítő és a fejlesztő értékelés igénye egyszerre van jelen, ezért a kétféle értékelési szemlélet különbségeinek megértése és a fejlesztő szemléletű értékelés alapelveinek áttekintése fontos kiindulópont a módszertani megoldások leírásához.

A minősítő értékelés az oktatási rendszerek alapeleme, hiszen ennek eredménye biztosítja és szabályozza a tanulók továbbhaladását a rendszerben. A mérési eredményeknek azonban információs szerepe is van, ezek ugyanis nagyban segítik (vagy akadályozzák) az oktatási intézmény legitimációját, mert az eredmények alapján megjósolható a tanulók későbbi teljesítménye (Zimmermann 2018), ez pedig erős információtartalommal bír az oktatási intézmények közötti választás során. A megbízható mérési eredmények elérése annyira erős igény, hogy ennek komoly kutatási és szakirodalmi háttere van az általános pedagógiában, az idegennyelv-oktatásban ennek a területnek még erősebb a tudományos értékű recepciója, hiszen a nyelvvizsgáztatáshoz kapcsolódóan jelentős szakmai tudás és tapasztalat gyűlt össze (pl. Bachman–Palmer 1996, Bárdos 2002, Vigh 2018). Míg a minősítő értékeléssel szembeni elvárások és az ezzel kapcsolatos kötelezettségek erős korlátok közé szorítják a vizsgafejlesztőket és a vizsgáztatókat, az osztálytermi értékelés, tehát a tanulási folyamathoz közvetlenül kapcsolódó mérés és értékelés ezzel szemben sokkal szabadabb és sokrétűbb terület. Az osztálytermi értékelés bonyolult eszközrendszer, amelyben komoly pedagógiai lehetőség rejtőzik, tehát a szummatív technikák mellett lenne szerepe a diagnosztikus és a formatív funkciójú értékelésnek is, bár a tanárok nem feltétlenül élnek ezzel a lehetőséggel. Ebben szerepe lehet a megszokásoknak, a tanulóként megélt modellek automatikus

továbbvitelének, de annak is, hogy a diagnosztikus vagy a formatív funkciójú értékelés megvalósítása bonyolult és időigényes tevékenység. A napi értékelési gyakorlatra erősen hatnak a (nyelv)vizsgáztatásból ismert megszokott technikák és eszközök, ráadásul a társadalmi elvárások is inkább a minősítő értékelési technikákat erősítik.

A minősítő értékelésnek tehát megvan a helye és a szerepe a tanulási folyamatban, azonban erős korlátai is vannak, kérdéses ugyanis, hogy valóban támogatja-e a fejlődést. A tanulás alapja nagyon leegyszerűsítve a próbálkozás – hibázás – újabb próbálkozás láncolata, a következetes és szigorú minősítő értékelés azonban nem támogatja a kísérletezést, a próbálkozást és a hibázást (Knausz 2008). A jobb értékelésre (a jobb jegy elérésére) törekvő tanuló számára ugyanis az a pragmatikus megoldás, ha a tétre menő feladatok megoldásakor inkább olyan területen marad, ahol kisebb a hibázási lehetőség. Ez a kockázatkerülő magatartás az idegennyelv-tanításban azt jelenti, hogy a nyelvtanuló igyekszik egyszerűbben, ám hibátlanul fogalmazni, nem kísérletezik az újonnan megtanult szerkezetekkel, márpedig ez megakasztja vagy lassítja a nyelvi fejlődést. További probléma, hogy a minősítő értékelés általában érzéketlen a befektetett munkára (Rheinberg 2001). Általánosan elfogadott tény, hogy az értékelés három különböző referenciaponthoz viszonyítható: a tantárgyi követelményekhez (kritériumorientált értékelés), a többi tanulóhoz (normaorientált értékelés), illetve a tanuló saját korábbi vagy szokásos teljesítményéhez. A minősítő értékelés megszokott normái miatt sokszor maguk a tanulók is azt tekintik a legigazságosabbnak és a legobjektívabbnak, ha csak kritériumorientált értékelés folyik, holott sokkal több információt hordoz a fejlődésről, ha a tanár az értékelés során a referenciapontokat tudatosan váltogatja, és erről nyíltan kommunikál (Rheinberg 2001). Így érhető el az, hogy a befektetett munkát is értékelni lehessen, ám a nyelvtanulónak világos képe legyen a nyelvtudási szintjéről is.

A fejlesztő szemléletű értékelés nem választható el a pedagógiai kultúra modernizációjától, hiszen azok az általános tendenciák jelennek meg benne, amelyek a tanítási-tanulási folyamat egészének átértelmezésében megnyilvánulnak. A fejlesztő értékelés jól illeszkedik ahhoz a szemléleti változáshoz, hogy a tanárközpontú pedagógiai kultúra tanulóközpontúvá alakul (Nagy 2010), és hogy a pedagógia vagy a módszertan fókuszába egyre inkább a tanulás kerül a tanítás helyett (*shift from teaching to learning*) (Bachmann 2018). A korszerű pedagógiai kultúra lényegi elemei a tevékenységalapúság, az autonóm tanulás támogatása, a tanulók egyéni szükségletéhez való igazodás (pl. Heacox 2006, KER 2002, OECD 2010), és ezeket az alapelveket érvényesíteni kell az értékelési folyamatban is. Az egyéni tanulási utak támogatása egyértelműen szükségessé teszi azokat az értékelési technikákat, amelyek képesek az értékelés individualizációjára is: ezek az egyéni célokhoz kapcsolódó fejlődés érzékelése, az egyénre szabott visszajelzés, a tanulási folyamatba beépülő folyamatos reflexiók sora (Gardner et al. 2008).

A fejlesztő szemléletű értékelés legfontosabb alapelvei a jövőre irányulás, az értékelés fogalmának kiterjesztése és a kitűzött célok következetes bevonása az értékelési

eljárásba. A jövőre irányulás a legszembetűnőbb jellemzője ennek a szemléletnek, és ez alapozza meg a fejlesztő értékeléshez kapcsolódó eszközzrendszert is. Míg a hagyományos, minősítő értékelés fókuszja mindig a múlt, hiszen a korábbi tanulási folyamat lezárása a legfontosabb funkciója, addig a tanulási folyamat részeként értelmezett értékelés (*assessment for learning és assessment as learning*) elsősorban a tanuló további fejlődésére fókuszál. Fontos tehát, hogy akkor értékeljünk, amikor még van mód a változtatásra (nem csak az adott tanulási szakasz végén, attól gondosan elkülönítve), és hogy az értékelés hordozzon a tanuló számára elegendő jól értelmezhető információt a további teendőiről, tehát arról, hogyan tudja a további tanulási folyamatát még sikeresebbé tenni. Ebből következik, hogy az értékelés a tanulási folyamatba beépítve és nem csak a tradicionális mérőeszközök segítségével történik (Gardner et al. 2008). A fejlesztő értékelés tehát nem a minősítő értékelés kiegészítő elemeként értelmezendő (a jegy mellé a tanár fűz még egy indoklást is), hanem egy attól elkülöníthető szemléletet jelent, amelyben az értékelés alapja nem feltétlenül a hagyományos eszközökre épülő mérés.

A minősítő értékelés célja a korábbi tanulási folyamat eredményeinek ellenőrzése, így a tárgy leszűkül a jól mérhető elemekre: leginkább a tudásra és bizonyos megszorításokkal a készségekre. A fejlesztő szemléletű értékelés legfőbb funkciója, hogy a további tanulási folyamatot támogassa, ennek érdekében azonban ki kell tágítani az értékelés fogalmát olyan területekre is, amelyek ugyan a hagyományos eszközökkel nem feltétlenül mérhetők, ám a tanulási folyamat sikeressége szempontjából nagyon fontosak, ilyen például a motiváció, az attitűdök, a tanulói önállóság, az együttműködési hajlandóság vagy képesség, a fejlődésre való igény vagy a problémamegoldó stratégiák alkalmazása. Az ezekre a területekre vonatkozó értékelő visszajelzések egyrészt kiemelik ezek fontosságát a tanulók számára, másrészt erősítik a fejlődés folyamatszerűségét. Az értékelés ilyen értelemben folyamatos reflexió a tanulási folyamatra, annak az összes elemére, és természetesen nem csak a tanár privilégiuma. A későbbi siker szempontjából elengedhetetlen, hogy a tanuló is foglalkozzon a saját tanulási folyamatával, próbálja érzéklni az előrehaladását, és hogy ne csak a saját munkájára reflektálhasson, hanem a tanár által megteremtett tanulási környezetre is. Az értékelés tárgyának kiterjesztése tehát úgy is értendő, hogy az ön- és társértékelés, a kurzusértékelés és a tanári munka értékelése összefonódik a tanulói fejlődés értékelésével, természetesen úgy, hogy világosak legyenek a további teendők, és egyértelmű legyen, melyik résztvevő miért felelős. A folyamatos reflexiók ugyanis csak akkor tudják valóban támogatni a tanulási folyamatot, ha minden résztvevő számára világos, hogy mi az ő felelőssége, mely pontokon tud ő maga változtatni.

A fejlesztő szemléletű értékelés harmadik fontos alapelve a célokhoz igazodás, és ez összefügg az utóbbi évtizedek egy másik fontos pedagógiai innovációjával, a tanulási eredményekkel is. Mindenfajta értékelés alapja a kitűzött célhoz való viszonyulás: jó minőségű mérőeszközt csak a tanítási/tanulási célok alapján lehet készíteni, és a

célok elérése a minősítő értékelésnek is alapja. A fejlesztés szempontjából ez a kérdéskör még hangsúlyosabb, hiszen a tanulási, fejlődési folyamat támogatása azt is jelenti, hogy a tanulónak információkra van szüksége arról, hogy a kitűzött céljaiból melyeket érte el, továbbá el kell gondolkodnia azon, hogy mi segítette őt e célok elérésében, és mit kell tennie azért, hogy a többit is elérje. Ahogy a tanulási folyamat egészére jellemző az elmozdulás a tanítás felől a tanulás irányába, úgy ez érvényes a célokra is. Ennek fontos eszköze a tanulási eredményeken alapuló célmeghatározás, amelynek lényege az, hogy a tanár a tervezési folyamatában nem a megtanítandó tananyagra koncentrálna, hanem arra, hogy az adott tanulási szakaszban milyen fejlődést, milyen tudástöbbletet kíván elérni a tanulónál. Ez az alapvetően a tervezési kompetenciát érintő szemléleti elem hat az értékelésre is, hiszen a fejlesztő szemléletű értékelés egyben tanulásieredmény-alapú értékelés is.

A tanulási eredmény (*learning outcomes*) fogalom megjelenése egy nemzetközi projekthez köthető: az ezredforduló óta folyik egy képesítési keretrendszer fejlesztése Európában (*European Qualifications Framework*: EQF; Magyar Képesítési Keretrendszer: MKKR). A képesítési keretrendszer fő célja az Európai Unión belül kiadott különböző képesítések és bizonyítványok összehasonlíthatóvá tétele, amit a különböző képesítési követelmények tanulásieredmény-alapú leírásának összehasonlításával oldottak meg (*Explaining the European Qualifications Framework for Lifelong Learning* 2008). A tanulási eredményeken alapuló tervezés az idegennyelv-tanárok számára egyáltalán nem újdonság: erre jó példát jelentenek a *Közös európai referenciakeret* szintleírásai, amelyek azt részletezik, hogy a nyelvtanuló mire képes az adott szinten (KER 2002). A tanulási eredmények megfogalmazása a képzési követelményekben és a tantervi dokumentumokban önmagában is nagyon érdekes folyamat (*Using Learning Outcomes* 2011, Ütőné 2015), a fejlesztő értékelés eszközzrendszeréhez pedig megkevlhetetlen a tanulási eredmények használata.

Ezek a jellemzők jól mutatják, hogy a fejlesztő értékelés komplex, a tanulási-tanítási folyamat egészét érintő szemlélet, amely erősen összefügg a tanári és a tanulói autonómia problémájával és a tanári szerepvállalással. A minősítő értékelés a hagyományos tanári eszköztárban leginkább kényszerítő eszköz. Megkockáztatható ez alapján, hogy általában azokban a tanároknak van hajlandóság arra, hogy az értékelést ne a nyomásgyakorlás eszközeként tekintsék, akiknek van más eszköze a tanulók motiválására, és akik számára természetes a reflektív hozzáállás a saját munkájukhoz és szerepükhöz is. A fejlesztő szemléletű értékelés következetes alkalmazása során ugyanis a tanulók egyenrangú félként reflektálnak a tanulási-tanítási folyamatra magára is, és ezek a reflexiók hordozhatnak olyan kritikai elemeket is, amellyel a tanárnak meg kell birkóznia.

A tanítási környezet jellemzői

A BME Idegen Nyelvi Központjában az idegen nyelvekhez kapcsolódóan többféle tevékenység folyik: a központ fő profilja a hallgatói nyelvoktatás, de emellett működik egy nyelvvizsgaközpont és egy nyelviskola, továbbá folyik tolmács- és fordítóképzés is. A hallgatói nyelvoktatás kínálatában szaknyelvi vagy a munkához kapcsolódó nyelvhasználatot fejlesztő kurzusok vannak, az angol és a német mellett újlatin nyelveket is lehet tanulni, illetve az egyetem nemzetköziesítési törekvései miatt egyre több a magyar mint idegen nyelv kurzus is. A hallgatók nyelvtanulását az egyetem támogatja: a választható (ingyenes) kredités tárgyak mellett a hallgatók egy bizonyos keret erejéig a 0 kredités kurzusokat is térítésmentesen, a keretük túllépése után pedig kedvezményes tarifával vehetik fel. Ebben a nagy intézményben félévente mintegy 2500 hallgató vesz fel valamilyen idegen nyelvi tárgyat, több mint 40 főállású nyelvtanár vagy oktató dolgozik itt.

A központban 2016 óta folyik egy módszertani és tartalmi fejlesztés, amelyben az elmúlt évtizedekben elterjedt csoportos, hálózati tanuláson alapuló tanártovábbképzési formákat alkalmazzuk (Gordon Győri 2007, Suñer 2018, Szabó et al. 2011). Megpróbálunk közös munkával megoldásokat találni a legfontosabb pedagógiai problémáinkra, amelyek közül kiemelkedik a hallgatói motiváció és a heterogén csoportok kezelése. A motiváció problémájáról a bevezetőben már esett szó: a hallgatók élni próbálnak a mai felsőoktatási gyakorlatban bőkezűnek tekinthető nyelvoktatási kínálattal, ám a szakmai tárgyak elsőbbséget élveznek, mert a szokásos kimeneti követelményeken túlmenően (nyelvvizsga-kötelezettség a képzés végén) többnyire nincsenek kötelező nyelvi szigorlatok vagy vizsgák a tanulmányi kötelezettségekbe illesztve. A hallgatók szabadon választják a kurzusokat, nincsenek kötelezően egymásra épülő nyelvi kurzusok, és a bemeneti szint csak ajánlatként jelenik meg. Ez a helyzet azonban azt a megoldandó problémát veti fel, hogy minden csoport egy-egy szemeszterre összeállt közösség, különböző szakmai és nyelvtudási háttérrel. Ezért többnyire heterogén csoportokkal kell dolgozni, tehát a differenciálás, az egyéni szükségletekhez igazodó nyelvtanítás technikáinak ismerete kiemelkedően fontos az itt dolgozó tanárok számára.

Ezekre a kihívásokra a válasz a pedagógiai megújulás: többféle továbbképzési formát alkalmazva (gyakorlatközösség, tanórákutatás, segítő célú hospitálás) próbáljuk a tanárainkat támogatni a kísérletezésben és a fejlődésben (Einhorn 2017). Ennek eredményeként komoly változások történtek: több tanárunk repertoárjába is bevonult a tanulási eredményeken alapuló óratervezés és az ezekről való kommunikáció a hallgatókkal, a játékosítás különböző formái, illetve a kurzusértékelés a hallgatók bevonásával. Mivel a fejlesztés központi eleme a tanulóközpontúság és a fejlesztő szemlélet alkalmazása, fontos volt, hogy legyenek olyan nyelvi csoportok, ahol ezeket az alapelveket következetesen alkalmazzuk. Kifejezetten cél volt, hogy sokféle módszert és eszközt próbáljunk ki ezekben a csoportokban, ezzel előkészítve a tanárok számára

készült módszertani ajánlásokat, és ilyen módon hospitálási lehetőséget is tudunk biztosítani a tanároknak. Az érintett német nyelvi kurzusok a következők voltak: Műszaki szaknyelv, Interkulturális kommunikáció, Kommunikációs készségfejlesztés, Felkészülés külföldi tanulmányokra. Ezek ajánlott bemeneti szintje B2, 18–25 fős csoportokkal dolgozunk, a kurzusok 2–3 kreditesek, heti 2 órásk, és egy szemeszter 14 hétig tart.

A fejlesztő értékelés alkalmazott technikái

A fejlesztő szemléletű értékelésről szóló tanulmányok és kutatások a formatív értékelési technikák vagy alternatív értékelési technikák közé általában az értékeléshez közvetlenül kötődő eszközöket sorolják: például portfóliók, ön- és társértékelés, részletes visszajelzés a tanulóknak, tanulási naplók. Az eszközrendszert ebben a kontextusban a szokásosnál kicsit tágabban értelmezem, hiszen a következetesen alkalmazott tanulóközpontú, egyéni szükségletekhez igazodó fejlesztéshez kapcsolódó értékelési eszközök sokszor nem választhatók el élesen a fejlesztés módszertanától. A következőkben azt gyűjtöm össze, hogy a fejlesztő szemléletű értékelést milyen módon lehet érvényesíteni felsőoktatási környezetben, ebből a szempontból fontos, hogy a célok és/vagy tanulási eredmények miként jelenhetnek meg az értékelési folyamatban, és hogy milyen módon lehet erősíteni a hallgatók bevonódását. Bár a hibakezelés és a hibakultúra problémáját nem szokták kifejezetten a fejlesztő értékeléshez sorolni, de erre is kitérek. Ezt a legnagyobb kihívás tárgyalása követi: hogyan lehet a minősítő és a fejlesztő értékelést együtt alkalmazni. A különböző technikák leírásába lehetőség szerint bevonom a hallgatói és a tanári véleményeket is, a leírások a saját tanórai tapasztalataimon alapulnak.

A tanulási eredményeken alapuló értékelés

A (nyelv)tanulók felelősségvállalásának egyik fontos alapfeltétele, hogy legyen világos számukra, hogy egy adott órának vagy egy-egy feladatnak egyértelműen van célja, hogy a tanár tudatos tervező tevékenységet folytat, és az is legyen egyértelmű számukra, hogy a tananyagok és feladatok meghatározásában nekik is lehet szerepük.

A felsőoktatásban a szemeszter első óráján többnyire minden kurzusban tájékoztatják a hallgatókat a követelményekről, a célokról és a tantárgyi tematikáról, és ehhez kapcsolódóan sokszor folytatnak megbeszélést a hallgatók elvárásairól is. Erről a tanárok gyakran úgy nyilatkoznak, hogy főleg idősvesztés, mert nincsenek a hallgatóknak pontos elvárásaik, és ezért a nyelvtanításban sokan a szükségletelemzést egy szintfelmérő teszttel próbálják helyettesíteni. Azonban a hallgatók elvárásairól és az egyéni céljaikról folytatott megbeszélés az első egyértelmű jelzés lehet arról, hogy ők maguk (is) felelősek a tanulási folyamatukért, azaz a féléves kurzus eredményességért. Kétségtelen tény, hogy a hallgatók nem feltétlenül tudnak árnyalt nyelvtanulási célokat meghatározni, ez az absztrakciós szint azonban nem is várható el egyértelműen egy

nyelvtanulótól, hiszen ez az a terület, amelyen sokszor a gyakorló tanároknak is nehézségeik vannak. Ezért egyszerűbb elindítani ezt a beszélgetés onnan, hogy mit nem szeretnének a kurzus során, és ebből kialakítható egy olyan beszélgetés, amelyben világossá lehet tenni a kurzus céljait és a korlátokat is, tehát kompromisszumot lehet kötni a kurzus keretei és a hallgatói igények között. Erre a célra sokféle módszer ismert, ha kisebb csoportokban dolgozhatnak a hallgatók, akkor egyrészt jobban inspirálhatják egymást, másrészt nem kell feltétlenül az egész közösség vagy a tanár előtt vállalni a véleményüket.

Egy-egy tanóra fejlesztési céljai is fontos információt jelenthetnek a hallgatóknak. Ezért minden óra elején megmutatom (kivetítem vagy előkészített papíron kiteszem) az adott órához meghatározott két-három legfontosabb tanulási eredményt, az órai aktivitások, a feladatok során pedig szükség esetén nagyon röviden magyarázom, hogy erre miért van szükség, mit tanulhatnak belőle, hogyan függ össze a fő céllal. Minden óra végén valamilyen formában visszatérünk ezekhez a célként kitűzött tanulási eredményekhez: vagy önértékelés történik, például úgy, hogy értékelje egy 10-es skálán, hogy mennyire érte el az adott tanulási eredményt, vagy tanári értékeléssel zárom a tanórát. Így arra is van mód, hogy a motivációjukra, az együttműködés minőségére vagy más fontos elemre is kitérjünk, továbbá ez lehetőséget ad az óratervezésre vonatkozó tanári önreflexióra is (utalás arra, ha valami túl nehéznek bizonyult, vagy a kiválasztott szöveg nem volt elég érdekes a hallgatóknak). Elsőre ez az óra végi rövid (ön)reflexió többnyire furcsa vagy vicces a hallgatóknak, de ha röviden megbeszéljük, hogy ennek a visszacsatolásnak az eredményét mire használhatják ők és mire használhatja a tanár (akár a következő óra megtervezése során), nagyon hamar részt vesznek ebben is aktívan.

Az órát lezáró, „levezető” aktivitások közé sokszor olyanokat illeszttek, amelyekben az órai folyamatra, így a saját tanulási folyamatukra kell reagálniuk. Például be kell fejezniük megkezdett mondatokat (célnyelven):

Ma azt tanultam, hogy...

Ma az volt a legérdekesebb, hogy...

Ma az volt nehéz, hogy...

Ennek egészen egyszerű formája az is, ha egy záró körben mindenkinek meg kell említenie egy-két szót vagy kifejezést, amit az adott órán tanult.

A legfontosabb tanulási eredményeket egy önértékelő lista formájában is összegyűjtöm, és ezekkel a kurzus közepén és a végén még egyszer dolgoznak a hallgatók. Ezek a listák arra szolgálnak, hogy az európai nyelvi portfólió logikája szerint (*Európai nyelvtanulási napló* 2001) azon gondolkozzanak el a hallgatók, hogy a kitűzött tanulási eredmények közül melyeket érték el, és mely területeken kell még dolgozniuk, fejlődniük. Ezek a listák jó lehetőséget adnak az értékelés tárgyának kiterjesztésére is, hiszen

felvehető rá olyan stratégiai és attitűdbeli területek, amelyek fontosak a nyelvtanulás szempontjából (pl. nem állok meg egy szövegben, ha nem értek egy szót; vannak ötleteim arra, hogyan fejleszem tovább a nyelvtudásomat). Mivel a tanulási eredmények olyan pedagógiai, módszertani szakszókincset tartalmazhatnak, amelyet a hallgatók nem feltétlenül ismernek, ezért ezeket mindig anyanyelven adom meg, bár a róluk szóló beszélgetés idegen nyelven folyik. Azokban az angol nyelvű csoportokban, ahol vannak külföldi hallgatók is, ezeket a listákat is a célnyelven lehet csak használni, ilyenkor annyira le kell egyszerűsíteni a nyelvezetét, hogy könnyen érthető legyen. Ezek a tanulási eredményen alapuló értékelések kiváló alkalmat adnak az önértékelés és a kurzusértékelés összekötésére is, azaz arra, hogy a hallgatókat bevonjuk az értékelésbe.

A hallgatók bevonása az értékelésbe

A tanulás sikeressége szempontjából fontos, hogy a (nyelv)tanuló mennyire tudja a nyelvtanulást olyan aktív folyamatnak érzékelni, amelyben neki is van szerepe. Sok tanárunk alkalmaz egy hárompillérű értékelési rendszert, amelyhez módszertani ajánlásokat készítettünk közös munkában. Ennek az az érdekessége, hogy egyrészt a tanulói önértékelést összekötjük a kurzusértékeléssel, másrészt, hogy a szokásos évkezdő megbeszélésen és a szintén szokásos év végi kurzusértékelésen túl a folyamat közepén is tartunk egy értékelést (ez a harmadik pillér). Mindhárom fázisban változatos módszereket használunk, erre számtalan példa ismert.

Meglepő módon a kurzus közepén beiktatott értékelések általában több információval szolgálnak, pontosabbak és részletgazdagabbak, mint a szemeszter végiek. Talán a hallgatókra is hat az, hogy ilyenkor még van tétje a véleményüknek, hiszen akkor értékeljük a folyamatot, amikor még van mód a változtatásra.

Ezeknek a negyedéves értékeléseknek általában az a menete, hogy először az első tanulási szakaszra meghatározott legfontosabb tanulási eredmények összefoglaló listájával dolgoznak a hallgatók egyéni munkában: a nyelvi portfólió alapötlete alapján arról döntenek, hogy az adott dolgot jól elsajátították, tudják, vagy még dolgozniuk kell rajta. Ezeket az egyéni listákat nem beszéljük meg, viszont valamilyen formában (kisebb csoport, plenáris beszélgetés stb.) foglalkozunk a következő három kérdéssel:

- Miben fejlődött a némettudásom a félév eleje óta? Miben vagyok jobb?
- Mit tehetnék még a félév végéig a némettudásomért?
- Min kellene változtatnunk az órákon?

Magát a beszélgetést többnyire célnyelven folytatjuk le, bár ilyenkor lehet magyarul is hozzászólni. Ezeknek a beszélgetéseknek van tanácsadó jellege is, hiszen arról is lehet beszélgetni a hallgatókkal, hogy milyen módon lehet használni, gyakorolni egy olyan idegen nyelvet, amely nem annyira van jelen a hallgatók mindennapjaiban, mint az angol. A negyedéves értékelés (amely 15–30 percnél nem vesz igénybe több

időt), általában azzal záródik, hogy egy névtelen levélben leírhatják a véleményüket a kurzusról.

Ezt a negyedéves értékelést igyekszünk a központ sok kurzusában lebonyolítani, és arra is volt példa, hogy ennek segítségével sikerült oldani a tanár és a hallgatók közti feszültséget is: a hallgatóknak módjuk nyílt arra, hogy egyértelműen jelezzék a problémájukat, változtatási igényüket, a tanárnak pedig volt lehetősége arra, hogy reagáljon erre, és tényleg változtasson.

A hallgatók bevonása azt is jelenti, hogy az értékelés a nyelvtanuló számára értelmezhető és jövőbeli cselekedetekre átfordítható legyen. Ezért minden kisebb órai aktivitáshoz és minden leadott munkához kapcsolódik egy néhány mondatos (szóbeli vagy írásbeli) értékelő visszajelzés. Ebben a klasszikus „értékelési szendvics” logikája alapján először arról esik szó, hogy mi sikerült jól, aztán a kérdések, problémák következnek, majd utalás arra, hogy mit kellene még tenni azért, hogy jobb legyen a produktum (Ziebell–Schmidjell 2012). Valóban rövid, tehát csak néhány mondatos visszajelzésről van szó, ez tehát egy olyan szöveg esetében, amelyet a tanár egyébként is alaposan átnézett és értékelt, nem különösen megterhelő nagyobb csoport esetében sem.

A szemeszter végén minden hallgató írásban is kap egy rövid záró értékelést (mivel a beadandó feladatok miatt élénk írásbeli levelezés folyik köztünk, levélben). Ebben kiemelem, hogy miben láttam fejlődést, és hogy milyen további javaslatom van a nyelvtanulásával kapcsolatban. Sok tanár számára ez is riasztóan hat, de a fejlesztő értékelés mindenképpen arra kényszeríti a tanárt, hogy sok információt gyűjtsön a hallgatók nehézségeiről és erősségeiről, és ezeket az információkat valamilyen formában őrizze is a maga számára, ráadásul sok kisebb produktum alapján értékel (erről később még esik szó), így minden hallgatónál van elég muníció egy nagyon rövid visszajelzéshez.

Az önértékelés és a tanári értékelés egyensúlyát ilyen módon jól meg lehet teremteni, a társértékelés kérdésében azonban figyelembe kell venni, hogy rövid ideig vannak együtt a csoportok, így nem mindig alakul ki az a bizalmi légkör, amely feltétlenül szükséges a társértékeléshez. A tanácsadó típusú társértékelésre, hibajavításra azonban van mód, ez azonban már átvezet a hibakezelés problémaköréhez.

A hibák kezelése, hibakultúra

A hibakultúra kérdését általában nem szokták a formatív értékelési technikák körébe sorolni, de mivel a hibajavítás is visszajelzés, ilyen értelemben meghatározó terület a tanulók fejlesztésében. A hibák kezelése alapvető abból a szempontból is, hogy a korábban említett próbálkozás–hibázás–próbálkozás láncolatot mennyire tudja a tanítási környezet támogatni. A hibázás, hibakezelés komplex terület, az általánosabb problémák megemlítése után csak néhány konkrét alkalmazott technikát sorolok fel.

A hibakezelés terén vannak általános kulturális elemek, tehát hogy egy adott kultúrában mennyire van hagyománya a hiba elfogadásának, mennyire elfogadott a hiba

lehetőségként való értelmezése. Ebből a szempontból lényeges, hogy a magyar kulturális közeg nem igazán elfogadó a hibákkal kapcsolatban. Ezt súlyosbítja, hogy a hagyományos, tanárközpontú pedagógiai szemléletből következik egy olyan tanári szerepértelmezés is, hogy nem maradhat hibás elem kijavíthatatlanul, mert ennek rossz hatása lesz a tanulókra. Ez a tradicionális felfogás egyébként nemcsak a tanulói fejlődést gátolja, hanem a tanárokét is: a tanári fejlesztések és továbbképzések során komoly problémát jelent az, hogy a tanárok sokszor a saját hibáikkal sem tudnak pozitívan bánni, nem megérteni, elemezni akarják azokat, és ilyen módon tanulni belőlük, hanem sokkal inkább elrejtteni őket. Ezen a ponton tetten érhető a pedagógiai kultúráváltás egyik legalapvetőbb akadályja: a tanároknak valami olyasmit kellene fejleszteniük, kialakítaniuk a tanulóknak, amit maguk sem tudnak feltétlenül, mert az ő iskolai szocializációjuk során egészen más normák szerint jártak el a tanárok.

A hiba szerepe tanítási nyelvenként is némiképp eltérő. A némettanításban erős pedagógiai hagyománya van a folyamatos hibajavításnak, ami természetesen összefügg a nyelv sajátosságaival is, hiszen sok olyan területe van a nyelvnek (névelők, melléknévragozás), amelyeken általában sokat tévesztenek a nyelvtanulók, bár ezek a pontatlanságok a kommunikációt nem feltétlenül zavarják. Ezért a némettanításban a hibák és tévesztések kezelésének motivációs szerepe is lehet. A német egyre inkább a második idegen nyelv pozíciójába kerül Magyarországon: a gazdasági képzésben részt vevő hallgatók számára egyértelműen, a műszaki képzésben részt vevők számára többnyire ez egy olyan nyelv, amelyet ritkábban használnak, mint az angolt. A kurzuskezdő órákon folytatott beszélgetésekből kiderül, hogy sok hallgató évekkor korábban a középiskolában vagy a nyelvvizsgára való készülés során beszélt utoljára németül, és a legtöbb csoportban van olyan hallgató, akinek a szakmai karrierje miatt meg kell ugyan tanulnia németül, de ezt a nyelvet csúnyának és nehéznek tartja. Ebben a helyzetben fontos a motiváció erősítése vagy megteremtése: meg kell tapasztalniuk a hallgatóknak, hogy a német is használható közvetítő nyelvként, és hogy ezen a nyelven keresztül is meg lehet ismerni izgalmas dolgokat a világból.

Az órai hibakezelés egyik legfontosabb alapelve az, hogy sok felszabadító kommunikációs helyzet adódjon, tehát olyan beszélgetések, amelyekben nem kell nyilvánosan megnyilvánulni: el kell egy kicsit távolítani a „tanári fület”, így ugyanis csak annyi a feladata a nyelvtanulónak, hogy a lehető legvilágosabban kifejezze a gondolatait, és megélje azt, hogy a közvetítő nyelven akkor jöhet létre kommunikáció, ha a közlése eléri az érthetőség küszöbét. Ebben az segít, hogy előnnyé tudjuk fordítani a kurzusok legnagyobb hátrányát, a heterogenitást. Mivel különböző karokról és szakokról veszik fel a hallgatók a nyelvvizsgát, többnyire nem ismerik egymást, így egyszerű kérdések, témák is eredményezhetnek izgalmas beszélgetéseket, csak arra kell ügyelni, hogy mindig az idegen nyelvet használják. Az órákon tehát sok kisebb beszélgetés folyik párban vagy csoportban, sokszor kell egy kérdésről több emberrel beszélgetni

röviden, egymás után. A heterogenitás komplexebb feladatokra is lehetőséget ad, hiszen így van mód a különböző szakterületek látásmódját feltárni és szintetizálni bonyolultabb kérdésekről.

A szóbeli közlés során többnyire nincs direkt hibajavítás, ez általában egy külön hibajavító blokkban történik. A hibajavító blokk mottója mindig az, hogy „szeretem a hibáimat”, amikor ez a felirat megjelenik, akkor egyértelmű a hallgatónak, hogy néhány tipikus hibát fogunk összegyűjteni, és ezeken fogunk dolgozni. A hallgatók javítási javaslatokat fogalmaznak meg, megpróbáljuk összeszedni a kapcsolódó szabályszerűségeket, ebben is sokat segít az egymástól eltérő nyelvtanulási múltjuk. Ezen a módon a hibák leválaszthatók a hibázóról, tehát nem egyéni fogyatékként jelennek meg, hanem lehetőségként a tanulásra. A tipikus problémák kiemelése pedig abban is megerősíti a hallgatókat, hogy mások is hasonló problémákkal küzdenek (például angol interferenciából adódó hibák, a német pontosabb szóhasználatából adódó problémák). Ezek alapján meg lehet határozni azt is, hogy mely nyelvi területeken kell célzottan gyakorolni vagy magyarázni a nyelvhasználatot. Nyilvánvaló, hogy ezen a módon egy módszertani tabut figyelmen kívül hagyunk, miszerint hibás alak nem kerülhet a táblára, ezek a közös, páros vagy csoportos feladatok azonban mindig úgy zárulnak le, hogy a végén a helyes alakokat látják a hallgatók. Időnként az írásban beadott feladatok hibáival is hasonlóan dolgozunk.

Az írásban beadott szövegek javítása mindig két fázisban történik: először csak megjelölöm a szövegben a nyelvileg hibás helyeket, és szükség esetén megjegyzésben utalok a problémára vagy adok megoldási lehetőségeket, de nem adom meg a helyes alakot. Megjegyzésben írok még szükség esetén szövegalkotásra vonatkozó javaslatokat is (kellene egy néhány mondatos összefoglalás, tisztázni kellene a két gondolat kapcsolatát stb.). Ezt követően a hallgató elkészítheti a javított változatot, amelynek értékelésekor már minden hibát egyértelműen javítok, megjegyzésben csak további jobbító javaslatokat adok meg. A végső értékelés a javított változat alapján történik. Ha van órai írásbeli szövegalkotás, akkor a társértékelést olyan segítő, tanácsadó módon alkalmazzuk, hogy egymásnak adnak ötleteket a szöveg javítására, a tanári értékelés csak azt követően történik.

A hibák kezelésében megfigyelhető valamiféle törekvés az egyensúlyteremtésre: az órai munkában, a szóbeli kommunikációban nagyvonalúan járunk el, míg az írásbeliségben a pontosságra törekszünk. A hallgatói visszajelzésekben nagyon sok megjegyzés vonatkozik erre a területre: sokan emelik ki, hogy jó volt félelem nélkül beszélni, hogy bátrabban beszéltek a kurzus végére. További visszatérő elem az értékelésekben, hogy jó alkalom volt a kurzus arra, hogy beszélgessenek, hogy megismerjenek más szakokra járó egyetemistákat.

A minősítő értékelés

A legnagyobb dilemma a fejlesztő értékelés alkalmazása során az, hogyan építhető be ez a minősítő értékelésbe, azaz hogyan tud a fejlesztő értékelést alkalmazó tanár megfelelni a minősítési kötelezettségének is. Erre a problémára általában két irányból keresik a megoldást a szakemberek: az egyik elképzelés szerint teljesen külön kell választani a két dolgot, a másik pedig arra keres lehetőséget, hogy a két utat valahogy ötvözze.

A fejlesztő és a minősítő értékelés szétválasztása azt jelenti, a tanulók számára világossá kell tenni, hogy az adott értékelés során a fejlődésüket akarjuk értékelni és segíteni, vagy egy világos, kritériumorientált tudásszintmérést akarunk végezni (Knausz 2008, Rheinberg 2001). Ez az adott pedagógiai helyzetben azt jelentené, hogy a folyamat során a fejlesztő értékelés technikáit alkalmazzuk, de a minősítésben a felsőoktatásban szokásos zárthelyik és beadandók kombinációjára építünk. A nyelvi tárgyak funkciója és célja alapján ez a megoldás nem volt elég ígéretes, sokkal inkább a másik út, tehát a fejlesztő és a minősítő értékelés kombinációja tűnt járhatónak, és erre jó megoldás a játékosítás (*gamification*) egyes elemeinek alkalmazása.

A játékosítás alkalmazására a közoktatásból sok kísérlet ismert (pl. Bíróné 2017, Prievara 2015), a felsőoktatásban, így a BME-n is sok szakmai tárgyban használják ennek az egyszerűsített, pontszerzésen alapuló változatát. A minősítő értékelés, tehát a kurzust záró jegy alapja az, hogy a hallgatók sok különféle feladatból válogatva hozhatnak létre különböző produktumokat, és ezzel folyamatos pontszerző tevékenységet folytatnak. A kurzus lezáró szakaszában (általában két héttel a szemeszter vége előtt, tehát a 12. héten) befejeződik a pontszerzés, és ekkor az előre meghatározott átváltás szerint érdemjegyre váltjuk a pontokat. Sok különböző feladatból lehet pontot gyűjteni, a hallgató maga választja meg, hogy milyen módon és milyen témában akar megnyilvánulni, és folyamatos javítási lehetősége is van, hiszen ha valóban jobb lett a szövege, akkor további pontokat kaphat.

Ez a gyakorlatban azt jelenti, hogy a kurzus elején megbeszéljük az alapelveket, ezt utána írásban is megkapják a hallgatók, majd minden órán kapnak különböző feladatlehetőségeket, előre rögzített pontszámmal, és ezek közül maguk választhatnak. A pontozás maga többféle módon alakítható ki, ezekben a kurzusokban egy 30 pontos rendszerrel dolgoztunk: összesen ennyire volt szükség a jeleshez, és 15 pont kellett az elégségeshez. Pontszerzésre folyamatosan van lehetőség, néhány tipikus feladat: esszé vagy kifejtő szöveg 1–1,5 oldalban egy konkrét témához (10 pont), műszaki szöveg fordítása magyarra (10 pont), kisebb opcionális (szókincs)teszt (5 pont). Szóbeli feladatok: szóbeli egyéni beszámoló valamilyen témáról (5–10 pont), komolyabb prezentáció 15 pont. Az órán is lehet pontot szerezni: órai szóbeli csoportmunka (2–3 pont), közös órai szövegalkotás (2–3 pont), és sokszor lehet egy nagyon aktív, jól sikerült óra után 1–1 pontot szerezni. A pontszámok természetesen elérhető maximumként értendők, és az adott produktum minőségétől függ a valós eredmény.

Ez az értékelési mód lehetőséget ad az egyéni célokhoz való igazításra: a minimális pontszámot össze lehet szedni egy-két nagyobb beadandóval, míg a jóhoz vagy a jellehez komolyabb munkára és aktív órai részvételre van szükség. Minden csoportban vannak olyan hallgatók, aki a szükségesnél több feladatot is beadnak, tehát megpróbálják kihasználni a tanulási lehetőséget.

A pontrendszer kialakítása hosszas kísérletezés eredménye, hiszen úgy kellett a rendszert beállítani, hogy ne lehessen munka nélkül jó jegyeket szerezni, de ne is legyen riasztó a munkamennyiség. A kurzusvégi jegyek csoportokbani eloszlása általában teljesen normális, nem látszik eltolódás egyik irányban sem.

A játékosítást általában komolyabb IKT-háttérrel szokták lebonyolítani (például Prievara 2015), mert ez egyszerűsíti a szervezést, a feladatok számon tartását, az eredmények közlését. Ebben a konkrét helyzetben azonban nagyon fontos volt a tanárok bátorítása, ezért semmilyen külön platformot nem használtunk, ez az értékelési rendszer egyszerű levelezéssel is működtethető.

A hallgatói visszajelzések az értékelési rendszerrel kapcsolatban nagyon elfogadók, általában pozitívként jelenik meg, hogy a rendszer igazságos, és jó benne, hogy több szabadságot hagy a hallgatóknak. Ez azonban többször negatívként is megjelenik: egy-egy hallgatói vélemény szerint kellene néhány keményebb dolgozat is. Sok visszajelzésben megjelenik valamilyen formában az a gondolat, hogy jó volt végre egy olyan tárgy, amelyben könnyebben lehetett jó jegyet szerezni. Ezt többnyire pozitívként írják le a hallgatók, de ez a rendszer egyik problémájára is rávilágít, hiszen a fejlesztő szemléletű értékelés ebben a szigorú, versengésen alapuló pedagógiai környezetben kockázatokat is rejt: a fejlesztő értékelésből adódó kisebb „nyomás” teljesítményromláshoz vezethet. Ebből a szempontból nagyon fontos, hogy a hallgatók mennyire értik meg és fogadják el ezt a szemléletet. Valószínű, hogy ez a szigorú zárt-helyihez szokott célcsoport igazodott volna ahhoz is, ha néhány szigorú dolgozatra alapozzuk a jegyét, és közvetlenül a zárthelyi előtt intenzíven készültek volna a hallgatók. Ezzel azonban az általánosabb kommunikációs fejlesztési célokat nem értük volna el, és a motiválást, a nyelvtanulási stratégiák erősítését is nehezebben lehetett volna megvalósítani.

Összegzés

Mivel nem kutatási céllal alkalmaztuk a módszertani elemeket, hanem fő célunk a tanárok módszertani fejlesztése volt, az eredményességgel kapcsolatban csak a hallgatói véleményekre lehet támaszkodni, amelyek különböző típusú reflexiókból származnak: értékelő megbeszélések, írásbeli hallgató visszajelzések, kisebb kérdőívek, amelyek a három éve folyó fejlesztés különböző szakaszaiban keletkeztek. Ezeket a tanári reflexiókkal és a szakmai viták eredményeivel lehet még kiegészíteni.

A hallgatói visszajelzések között egyértelműen dominálnak az elfogadó, pozitív vélemények. Hangsúlyosan jelent meg a véleményekben a siker, a sikerélmény és a fejlődés gondolata valamilyen formában, illetve a szabadság és az önállóságra való lehetőség. Sokkal kisebb számban (csoportonként legfeljebb 1–2 említés) ezzel pontosan ellentétes vélemények is megjelentek a válaszokban: ezekben mindig az a gondolat merült fel, hogy a szabadság, az interaktivitás, a felelősség sok munkával jár, és a hallgató elvárásainak jobban megfelelt volna, ha lehetőséget kap a passzív jelenlétre. Több olyan szöveg is készült, amelyben a hallgató részletesen leírta azt a folyamatot, ahogy ki tudott lépni a passzivitásból, tehát meg tudta haladni a saját kezdeti elvárásait.

Sok hallgatói reflexió vonatkozott a személyességre, az egyéni visszajelzésekre, a rövid személyre szabott értékelésekre. Úgy tűnik ez alapján, hogy a tömegoktatás keretei között továbbhaladó hallgatók számára az értékelés individualizálódása már önmagában is érték. Ezt kiegészítette sok arra vonatkozó megjegyzés, miszerint a nyelvi fejlődést leginkább az segítette, hogy kellemes környezetben, kellemes emberekkel lehetett kommunikálni.

A tanári reflexiók alapvetően a megnövekedett feladatokra és terhekre vonatkoztak: a tanácsadás, a kommentálás, az értékelés, a többszöri javítás sok tanár számára riasztó többletterhernek tűnt. Ez azonban sokkal inkább félelem, mint valós tapasztalás volt, mert a különböző elemeket kísérletképpen alkalmazó kollégák mindig arról számoltak be, hogy az eredeti félelmeiknél kevésbé volt megterhelő a munka. Sok kolléga kis lépésekben bevezette a fejlesztő értékelés egyes elemeit, egy-egy csoportban átalakította az értékelési rendszerét, sok csoportban a hallgatók több lehetőséget kaptak az oktatási folyamat értékelésére. A tanári visszajelzések alapján a legnagyobb problémát a tanári szerepek változása jelentette: a visszahúzóadás, a központi szerep elhagyása egyes munkafázisokban. A kapcsolódó szakmai vitákban sokszor felmerült még a hibák kezelése, a pontosság szerepe: több kolléga számára nem tűnt elfogadhatónak vagy célravezetőnek a hibakultúra megváltoztatása.

A tanulói autonómia kérdésével kapcsolatban is vannak érdekes tanulságok. A fejlesztő értékelés egyik legfontosabb hozadéka, hogy erősíti az autonómiát és a felelősségvállalást, ugyanakkor ez az egyik legnagyobb csapdája is, hiszen ha a tanár adottságnak tekinti az autonómiát, és nem fejleszti azt, akkor az ön- és társértékelés, az önálló tanulási útra és a saját szükségletekre való koncentráció kudarchoz vezethet. Ez a terület a szaknyelvtanításban még hangsúlyosabb, hiszen a tanár és a tanuló közötti felelősség- és feladatmegosztás a szakmai és a nyelvi kompetenciák megosztása miatt sokkal szimmetrikusabb.

Az autonómia fejlesztésének klasszikus eszköze, hogy választási lehetőségeket kell hagyni a nyelvtanulónak, és támogatni kell abban, hogy vállalja választásának következményeit. A választás a differenciált fejlesztésnek is alapvető eszköze: így az órai munkában is sokszor jelenik meg olyan szakasz, amikor a hallgatók alternatív feladatok

között választhatnak. Az alternatív feladatok többféle módon szervezhetők: téma alapján (például egy adott probléma megközelítése, műszaki, környezetvédelmi, gazdasági stb. oldalról), a tevékenység alapján (például olvasás, film, szókincsmunka), a feladatnehézség alapján (például ugyanahhoz a szöveghez több vagy kevesebb nyelvi segítséget tartalmazó szövegértési feladat). Ezekkel a lehetőségekkel a hallgatók különböző tudatossággal tudnak élni, és erre a hallgatói visszajelzésekben is sokszor kitérnek: általában sokan kiemelik pozitívként, hogy volt mód választani, ám szinte minden csoportban megjelenik egy-egy vélemény, hogy túl sok volt ez a szabadság, vagy hogy túl sok munkával járt.

A felsőoktatási kurzusok példája alapján megállapítható, hogy a tanulóközpontú értékelés a fejlesztő szemlélet következetes alkalmazását jelenti, és ez a tanároknak és a tanulóknak egyaránt feladatokat jelent. A fejlesztő szemléletű értékelés legfontosabb hozadéka az, hogy a nyelvtanuló kényszerűen kilép a passzív, a tanulást és az értékelést csak elszenvedő szerepből, aktivitásra kényszerül: döntéseket hoz, létrehoz dolgokat, véleményez, reflektál, értékel. Ez az, ami a legnagyobb lehetőség a fejlesztő szemléletű értékelést megvalósító pedagógiai innovációkban.

IRODALOM

- Bachmann, H. (2018): Hochschullehre neu definiert – shift from teaching to learning. In: Bachmann, Heinz (Hrsg.): *Kompetenzorientierte Hochschullehre. Die Notwendigkeit von Kohärenz zwischen Lernzielen, Prüfungsformen und Lehr-Lern-Methoden*. Eine Publikation der Abteilung Hochschuldidaktik und Erwachsenenbildung der Pädagogischen Hochschule Zürich. [Forum Hochschuldidaktik und Erwachsenenbildung. Band 1.] Bern: hep verlag ag., 14–33.
- Bachman, L. F. – Palmer, A. S. (1997): *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press.
- Bárdos Jenő (2002): *Az idegen nyelvi mérés és értékelés elmélete és gyakorlata*. Budapest: Nemzeti Tankönyvkiadó.
- Bíróné Várhegyi Judit (2017): Osztálytermi kísérlet egy alternatív értékelési módszer kipróbálására. *Modern Nyelvoktatás*, 23/1, 37–44.
- Black, P. – Wiliam, D. (1998): Inside the Black Box: raising standards through classroom assessment. *Phi Delta Kappan*, 80/2, 139–148.
- Einhorn Ágnes (2017): Gyakorlatközösség a felsőoktatásban? Egy kísérleti jellegű fejlesztés bevezető szakasza. In: Besznyák Rita (szerk.) *Porta Lingua – 2017. Szaknyelvhasználat: a tudomány és a szakma nyelvének alkalmazása. Cikk, tanulmányok a hazai szaknyelvoktatásról és -kutatásról*. Budapest: Szaknyelvoktatók és Kutatók Országos Egyesülete, 293–304.
- Európai nyelvtanulási napló* (2001): Veszprém: Kiss Árpád Országos Közoktatási és Szolgáltató Intézmény, Nodus Kiadó.

- Explaining the European Qualifications Framework for Lifelong Learning.* (2008): Luxembourg: Office for Official Publications of the European Communities.
- Gardner, J. – Harlen, W. – Hayward, L. – Stobart, G. (2008): *Changing Assessment Practice. Process, Principles and Standards.* Assessment Reform Group.
- Gordon Györi János (2007): Tanórákutató. Egy elterjedőben lévő oktatásfejlesztési módszer magyarországi adaptációjának kérdései. *Új Pedagógiai Szemle*, 57/2, 15–23.
- Heacox, Diane (2006): *Differenciálás a tanításban, tanulásban. Kézikönyv a 3–12. évfolyam számára.* Budapest: Szabad Iskolákért Alapítvány.
- KER = *Közös európai referenciakeret: nyelvtanulás, nyelvtanítás, értékelés* (2002): Pedagógus-továbbképzési Módszertani és Információs Központ Kht.
- Knausz Imre (2008): *Mit kezdünk az értékeléssel? Adalékok az integrációs nevelés pedagógiájához.* Budapest: Educatio. https://tanitonline.hu/ftp/Mit_kezdjunk_az_ertekelessel.pdf (2019. július 12.)
- Nagy József (2010): *Az új pedagógiai kultúra.* Szeged: Mozaik Kiadó.
- OECD (2010): *The Nature of Learning: using research to inspire practice.* Paris: OECD Publishing.
- Prievara Tibor (2015): *A 21. századi tanár. Egy pedagógiai szemléletváltás személyes története.* Budapest: Neteducatio Kft.
- Rheinberg, F. (2001): Leistungsbeurteilung im Schulalltag: Wozu vergleicht man was womit? In: Weinert, F. E. (Hrsg.). *Leistungsmessung in Schulen.* Weinheim: Beltz, 59–71.
- Suñer, F. (2018): Planung und Durchführung von Fort- und Weiterbildungsmaßnahmen. In: Roche, J. – Einhorn, Á. – Suñer, F. (Hrsg.) *Unterrichtsmanagement.* Tübingen: Narr Francke Attempto Verlag GmbH. [Kompendium DaF/DaZ 6], 306–320.
- Szabó Mária – Singer Péter – Varga Attila (2011, szerk.): *Tanulás hálózatban. Elméleti összefoglaló és gyakorlati tanácsok az eredményes hálózati tanulás megvalósításához.* Budapest: Oktatókutatató és Fejlesztő Intézet.
- Using learning Outcomes. European Qualifications Framework Series: Note 4.* (2011): Luxembourg: CEDEFOP, Publications Office of the European Union.
- Ütöné Visi Judit (2015, szerk.): *A Magyar Képesítési Keretrendszer és a validáció lehetőségei a köznevelés szemszögéből. A TAMOP-3.1.8 keretében megvalósuló fejlesztés.* Budapest: Oktatási Hivatal.
- Vígh, Tibor (2018): Sprachtests im kommunikativen Fremdsprachenunterricht. In: Roche, J. – Einhorn, Á. – Suñer, F. (Hrsg.) *Unterrichtsmanagement. Grundlagen der Qualitätsentwicklung und des Qualitätsmanagements.* Tübingen: Narr Francke Attempto Verlag GmbH. [Kompendium DaF/DaZ 6], 85–122.
- Ziebell, B. – Schmidjell, A. (2012): *Unterrichtsbeobachtung und kollegiale Beratung.* Neu-Berlin etc.: Langenscheidt. [Fernstudieneinheit 32.]
- Zimmermann, T. (2018): Durchführen von lernzielorientierten Leistungsnachweisen. In: Bachmann, H. (Hrsg.) *Kompetenzorientierte Hochschullehre. Die Notwendigkeit von Kohärenz zwischen Lernzielen, Prüfungsformen und Lehr-Lern-Methoden.* Eine Publikation der Abteilung Hochschuldidaktik und Erwachsenenbildung der Pädagogischen Hochschule Zürich. [Forum Hochschuldidaktik und Erwachsenenbildung, Band 1.] Bern: hep verlag ag., 50–84.

ERDEINÉ BACHER KATALIN

A fejlesztő értékelés a nyelvórán - példák a szakgimnáziumból

Formative assessment in the language class

There are hardly any better means of pedagogy to benefit students, or to motivate them to achieve further progress than assessment itself. The aim of formative assessment is to help teachers identify the learning needs of students and to adjust their teaching techniques accordingly. The article provides a selection of methodology ideas and secondary school language class best practices which can be of use in the theory or practice of primary school language teaching. The article not only gives examples of how the teacher using formative assessment can adjust teaching techniques to learning needs, but it also includes diverse techniques of continuous assessment, with special regard to the assessment of new student groups, self-evaluation and ways of utilising today's digital assets in the classroom.

Keywords: formative assessment, assessment strategies, improvement of language skills, constant feedback, motivation

Szókratész, akinek két legjellemzőbb módszere a sétálva kérdezés, kérdezéssel tanítás volt, előszeretettel nevezte magát a gondolatok bábájának, hisz ezekkel a módszerekkel hívta elő tanítványaiiban a gondolatokat, a tudást, így láttatta meg az összefüggéseket, így serkentette őket gondolkodásra. Vajon hogyan értékelhette a tanítványait, akikből tudjuk, kiválóságok lettek? Bizonyára nem vetette oda nekik séta közben: „hármás”, de ugyanígy, nem intézte el annyival: „ötös”! Feltételezem, a rávezető, irányító, gondolkodtató kérdéseken túl részletesen reagált arra, amit hallott, s értékelő szavaival továbbfejlesztette, magasabb szintre emelte tanítványai készségeit, képességeit, utat mutatva nekik a folyamatos fejlődéshez.

Be kell látnunk, hogy egyéni fejlődési útvonalakban kell gondolkodnunk a tanulási folyamat szervezésekor: a korszerűen értelmezett tanulás egy olyan, együttműködésen, kölcsönös reflektáláson alapuló folyamat, amelyben a személyre szabott célok, a folyamatos visszajelzések, a segítő stratégiák, módszerek, a hosszabb távú motiváltság (*drive*) ösztönző ereje, a bevonódás dominál. A fejlesztő értékelést alkalmazó tanár a tanulási szükségletekhez alakítja a tanítást, folyamatosan, sokszínűen gyakorol, ezt osztályzat nélkül, tét nélküli értékeléssel leméri, visszajelez, figyel, hogy tudna a gyerek többet teljesíteni, közösen feltárják az okokat, módszereket ajánl, önbizalmat, helyes önértékelést épít, csak ezt követően értékel osztályzattal. A diák megtanul tanulási célokat

megfogalmazni, saját munkájára reflektálni, önmagát és társait értékelni. Az alábbiakban a témához tartozó módszertani ötleteket, a középiskolai angolórákon bevált jó gyakorlatokat gyűjtöttem össze.

A folyamatos értékelés technikái

Új csoport esetén az első időszak, a „beszokás” fontos mozzanata a gyerekek számára az adott tanár értékelési szempontjának a megismerése. Az év elején próbafelélések során lehet közösen végigbeszélni a szóbeli felelet erényeit, gyengéit; kiemelni, hogy legközelebb mi mindennel lehet azt még erősíteni. Mikor a gyerekek már érzik a tanár elvárásait, akkor is, év közben is előbb véleményt kérek a többi diáktól és azt követően magától a felelőtől, majd következik a szaktanári értékelés. A folyamat akkor teljeseedik be, amikor a három fél véleménye már rendre alapvetően egybecseng. Az így közösen kialakított értékelés-önértékelés alapján a későbbi évek alatt is belátással-megértve elfogadják az osztályzataikat. A középiskolába frissen bekerült tanulók érzékenysége nagy, minden értelemben keresik a helyüket a csoportban, figyelik a tanár és a többiek reakcióit, a többiekéhez mérik a nyelvtudásukat. „*Mutasd be magad érdekes, izgalmas, színes módon!*”, ez szokott nálam az év eleji első házi feladatok egyike lenni nyelvórákon szóban vagy írásban. A tanulók írásbeli bemutatkozását a padokra tesszük a teremben, a többiek körbejárnak, odaülnek, elolvassák, s ha tetszett, egy tetszőleges jelet (virág, smiley stb.) rajzolnak a lapra. Utána összeszedjük, megszámloljuk, kinek hány elismerő 'like'-ja van. Érdekes a gyerekek számára, mert többet megtudhatnak egymásról, emellett a többiek munkáit elolvassva, el tudják helyezni a magukét a mezőnyben anélkül, hogy kellemetlenül éreznék magukat, és végül: ha gyengébben szerepeltek a társak előtt, nem a tanárnak kell rájuk pírítania, legközelebb biztosan több időt, figyelmet szánnak rá.

Ne idegenkedjünk kérdőívben (*1. ábra*) kikérni a diákjaink véleményét arról, hogy az előző év(ek)ben milyen órai és egyéb módszerek váltak be az esetükben a nyelvtanulásban, hogyan tudunk nekik személyre szabottan segíteni céljaik elérésében, mi az, amit szeretnek az óráinkban, s mi az, amit szeretnének megváltoztatni.

Ha igen, hogyan? (pl. filmek angol nyelven, külföldi levelezőtárs, interneten angol cikkek; applikációk, magántanár/nyelvtanfolyam, nyelvvizsga-feladatsorokat oldok meg, stb.):

- | | | | |
|--|---------------|-------|-----|
| 1. Az angolt szívesen tanulom : | | igen | nem |
| Szívesebben tanulom, mint az általános iskolában: | | igen | nem |
| 2. Érzem, hogy a beszédem év eleje óta fejlődött : | nagyon sokat | sokat | nem |
| Magabiztosabban, bátrabban szólok meg: | sokkal inkább | igen | nem |
| A szókinccsem is fejlődött: | nagyon sokat | sokat | nem |
| Az órákon oldottan, jól, derűsen érzem magam: | nagyon | igen | nem |
| 3. Órán kívül is (szívesen) foglalkozom az angollal: | | igen | nem |
| 4. a) Angolból az erősségeid (felsorolásszerűen) (pl. szótanulás, beszédkészség, szorgalom stb.): | | | |
| b) Gyengébb pontjaid, amit fejleszteni szeretnél (felsorolásszerűen): | | | |

5. Milyen tanórai **munkamódszer** a leghatékonyabb a számodra?
Számold be (1-es a leghasznosabb)!
- ...tanári magyarázat
 - ...egyéni feladatmegoldás
 - ...pármunka
 - ...csoportmunka
 - ...tanulói kiselőadás (pps-sel)
 - ...órai kreatív, projekt jellegű feladatok
 - ...önálló kutatómunka
 - ...tanulás IKT-eszközök használatával (mobil, tablet, kivetítő stb.)
 - egyéb:
6. Amit **szerettem**, értékelttem, hasznosnak/érdekesnek találtam az eddigi órákon (felsorolásszerűen):
7. **Kérés, javaslat** az évre (pl. több szóbeli számonkérés, változatos szorgalmi feladatok stb.):
8. **Céljaid**, terveid az angol nyelvvel:
- a) rövid távon (pl. turistákkal való szót értés, dalszövegek/könyvek megértése, nyári diákmunka stb.):
 - b) hosszabb távon (pl. valamilyen nyelvvizsga, közép/emelt érettségi, jó szintű társalgási készség, felvételi külföldi egyetemre stb.):
9. **Miben, hogyan segíthetek Neked céljaid elérésében?**
- Örömteli, sikeres munkát kívánok ! 😊*

1. ábra. Nyelvtanulást segítő kérdőív

Óra közben többször (különbözőképpen, változatosan) bizonyosodjunk meg arról, hogy ki érti/mennyire érti a tananyagot. Korszerű IKT-eszközökkel azonnali visszajelzést is tudunk kapni/adni arról, hogy hol tartanak diákjaink (lásd gyors, tudásmérő alkalmazások pl. Kahoot, Redmenta). Óra végén értékkártyát, önértékelő, illetve kulcskártyát alkalmazhatunk. Gyakoroljunk sokféle képpen, türelmesen, ne adjunk egyből „éles” jegyet. Folyamatosan, írásban is értékeljük a tanulók teljesítményét, hiszen az osztályzatnál jóval inkább fontosnak tartják a megjegyzéseinket, szöveges értékelésünket, hogyan, merre haladjanak tovább.

Fontos (kisebбекnél még inkább), hogy láttassuk a fejlődés mértékét, hogy azt a diák nyomon követhesse, akár vizuálisan, grafikus megjelenítéssel is. Manapság nagyon népszerű a különféle pontrendszerek, jelvények bevezetése, a „játékosítás” is. A fiatalabb tanulók szeretik, ha Oklevelet (Certification) kapnak. A www.certificatemagic.com oldalon könnyen-gyorsan készíthetünk névre szóló, változatos, színes okleveleket. Nagyon ösztönző a diákok számára a különféle jelvények adása valamilyen tudásanyag elsajátításáért, illetve készségbeli fejlődésért. Ezeket adhatjuk mi is, de társak is, sőt saját magának is megíthatja egy diák, ha úgy érzi, teljesítette a kitűzött szintet, ezzel reflektál saját teljesítményére is.

Hasznos a portfólió is, a tanulói munkák meghatározott, közösen kialakított elvek szerinti gyűjteménye (otthoni külön dossziében kérem gyűjteni a kiegészítő jellegű, kreatív típusú munkákat). Természetesen ma már digitális formában is könnyen kivitelezhető, s a hagyományos formától a digitális mappán át eljuthatunk akár a diák ilyen funkciójú weblapjáig is. Bármikor áttekinthető a tanulási folyamat, állomásairól, az előrehaladás egészéről ad képet, növeli a tanulói felelősséget.

A folyamatos értékeléshez arra is szükség van, hogy a tanár információkat gyűjtsön a tanulókról. A diákok fejlődését táblázatok formájában is nyomon követhetjük, a sorok/oszlopok tárgyát (kompetenciák fajtája) szabadon megválaszthatjuk. Az efféle táblázatok könnyen átláthatóvá-értelmezhetővé teszik a teljesítményt. A teljesítmények alakulását természetesen sokféle grafikonon is megjeleníthetjük, sőt össze is vethetjük a korábbi személyes eredményekkel, csoportátlaggal stb.

Tanuló: K. Judit 10. G

Esemény: Projektnapok Téma: London Projekt időpontja: 2019. 03. 04–09.

TUDÁS	KÉPESSÉG	ATTITŰD	AUTONÓMIA
✓ Ismeretei szerteágazóak a témában, tudását folyamatosan és szívesen bővíti	✓ Ügyesen továbbfejlesztette csoportja eredeti alapötletét	✓ Bátran képviseli csoportja érdekeit	✓ Kezdetben útmutatással dolgozott (munkája precíz); később egyre inkább önálló javaslatokat fogalmazott meg
✓ Könnyen felismeri és megnevezi azokat a részterületeket, ahol továbblépésre van szükség	✓ Kommunikációs és szervezési képességei magas színvonalúak	✓ Nyitott a problémák újfajta megoldására, kreatív	✓ Felelősséget érez a rábízott feladat precíz teljesítése iránt
✓ Érti és átlátja a témakör problémáit, megoldandó feladatait	✓ Szívesen bekapcsolódik kisebb alcsoportokhoz egy-egy részfeladat elvégzésére	Szem előtt tartja ugyan a csoport érdekeit, de vitás kérdésekben nem mindig törekszik megegyezésre	✓ Nem esik nehezére korrigálni saját hibáit
	✓ Aktivitásával hatékonyan motiválja a csoportot	Túlzottan kritikusan szemléli társai tevékenységét	✓ Önálló javaslatokat fogalmaz meg, sok ötlete van
	Nem mindig használja fel eredményesen a társai által kidolgozott részeredményeket	Nem mindig tartja be a csoport által elfogadott szabályokat, néha a saját feje után megy	Az érvekkel megalapozott döntéshozatal terén még fejlődhet

2. ábra. Feljegyzés tanulási folyamatról

Összetettebb értékelést is adhatunk már új eszközök segítségével fogalmazások, esszék fejlesztő értékelése során pl. videoértékeléssel (<http://www.utipu.com>).

Az ön- és társértékelés néhány lehetősége

Hasznos-örömteli dolog, ha egy nagyobb tanulási egység, egy tankönyv befejezése után a diákok haladási naplót készítenek: My Progress Diary. Ebben a tanulók tematikusan összegyűjtik, hogy mivel léptek előre, mit tudnak, mire képesek már az adott nyelvtanulási szakasz után (3. ábra).

1. Fejezd be a mondatokat a nyelvtanulással kapcsolatban:

Már jóval magabiztosabb vagyok.....
 Úgy érzem, sokkal könnyebben.....
 Sokat fejlődtem.....
 Korábban gondot okozott, de most már.....
 Jelenleg a legnagyobb erősségem.....

Még nem vagyok elégedett.....
 Jó lenne még fejlődni.....
 Gyakran nehézséget okoz.....
 Sokat segítene, ha.....

Már tudok beszélgetni az alábbi témákról:

Írás

Már tudok:

	Igen	Nem
Folyékonyan olvasok angolul.	✓	
Könnyen megértem a hallott szöveget.		✓
Szívesen szólok meg a csoport előtt.	✓	
Könnyen tanulok szavakat.		✓

	Gyakran	Néha	Soha
Angolul is szoktam filmeket, videókat nézni	✓		
Sok dalszöveget tudok angolul		✓	
Szoktam angolul kommunikálni az interneten (e-mail, Facebook, Skype, kommentek stb.)	✓		
Keresem a lehetőséget, hogy angolul beszélgethessek			✓
Szoktam olvasni angolul (honlapok, újságcikkek, hírek, könyvek)		✓	

3. ábra. Példa egy fejlődési naplóra

A tanulási napló (ez praktikusán lehet egy blog is, amelyet a tanár is egyszerűen, folyamatosan figyelemmel kísérhet, megjegyzésekkel, tanácsokkal láthat el) egy tanulást segítő önreflektív módszer. A tanuló leírja, dokumentálja, az adott órákon mi történt, mit tanult, miben fejlődött, mi nem sikerült, esetleg mi érdeklí még az anyaggal kapcsolatban, hol nézett még utána, hogyan, milyen módszerekkel gyakorolt még, stb.

Értékelőív is segítheti a visszacsatolást, ezt a tanulók maguk, párban, illetve csoportban tölthetik ki, egyszavas, egész mondatos vagy skálán bejelölt módon az óra utolsó egy-két percében. Én a füzet hátuljába ragasztattam a lapot, a heti óraszámra elkészítve. Ez a módszer nagyon jól bevált rövid távú (1–4 hét) időtartammal, egy-egy olyan csoport esetében, ahol valamilyen probléma mutatkozott. Ilyenkor a fejlesztési területekre fókuszálva készítettem el a táblázatot, és ezzel szinte ráirányítom a diákok figyelmét egy-egy fejlesztendő területre (4. ábra).

<i>Értékelj teljesítményedet 1–5 ponttal /tegyél pipát vagy ikszet / jó kedvű-szomorú smiley-t stb.</i>	2019. 05. 06. Hétfő	2019. 05. 07. Kedd	2019. 05. 08. Szerda	2019. 05. 09. Csütörtök	2019. 05. 10. Péntek
Az órán gyakran megszólaltam angolul, aktív voltam					
Párommal jól eljátszottam az adott párbeszédet					
Magam is ki tudtam találni változatos párbeszédet az adott témában					
A ma tanult új szavakat tudtam használni a beszédemben is					
Az olvasmányt megérttem, és a tartalmát el tudtam mondani a páromnak					
A csoport előtt már kevésbé izgulok, ha meg kell szólalni					

4. ábra. Minta a füzet hátuljába ragasztott nyelvvoravégi értékelőívre, például a beszédkésztség fejlesztésére

Különösen nagy a jelentősége ebben az életkorban a kortárs csoport véleményének, ezért is jó módszer a csoportmegbeszélés. Az adott feladatot együtt végző diákok megbeszélése/írásbeli értékelése különösen szükséges a hosszabb, projekt jellegű tevékenységek értékelésekor. Egy kialakult, stabilabb értékelési gondolkodás alapján egy csoportmunka után, a team tagjai esetenként maguk határozzák meg, ki mennyivel járult hozzá a munka eredményességéhez, a végső értékelés ez alapján is történhet (a csoportmunka után a csoport összes tagja ugyanazt az értékelést is kaphatja – szöveget, illetve érdemjegyet). Ezt követően a csoport írásos kérdésekkel vezetett módon elemzi, hogy ki mivel járult ehhez hozzá (5. ábra). Egymásnak üzeneteket írnak/megbeszélnek, hogy kitől mit várnak a következő együttes munka során.

Néha egy tanuló értékeli a csoport munkáját: szintén előre megadott szempontok alapján mindig más végzi az értékelést. Fontos, hogy módot adjunk a többiek hozzászólására is. A tanulók párban (általában közösen) megállapított szempontok alapján is értékelhetik a tanulási folyamatot. Ez erősen megnöveli a felelősséget, tudatosítja a tanulási folyamat egyes elemeinek szerepét.

ÉRTÉKELŐÍV

Kérlek, értékeld egy-egy külön íven a magad és csoporttársaid munkáját az alábbi szempontok szerint 0–3 ponttal. Alaposan és őszintén gondold át a véleményed, próbáld valóban igazságos lenni. Nem kell, hogy mindenkinek maximum pontszámot adj mindenre, hiszen társaidnak más és más erősségei, ill. fejlesztendő területei lehetnek. Ha megjegyzéseket is írsz, segítesz nekem, hogy értelmezni tudjam a pontszámaidat. Dolgozz egyedül az értékelőív kitöltésekor. Köszönöm.

Pontozd saját magad és társaid teljesítményét az alábbi pontok szerint:

- 3 – Az adott területen jobban teljesített, mint a csoport többi tagja
- 2 – Körülbelül a csoport többi tagjával egy szinten teljesített
- 1 – A csoport többi tagjánál gyengébben teljesített
- 0 – Nem volt a csoport segítségére az adott területen

Az értékelt személy neve:

Az értékelő személy neve:

<i>Szempontok</i>	<i>Pont</i>	<i>Megjegyzések, példák, kiegészítések</i>
Részvétel a csoport munkájában Személye a közös munkában ösztönző; aktív, tevékeny volt; a célok érdekében mindig megegyezésre törekedett, ha tudott, szívesen dicsért másokat a munka során.		
Időbeosztás és felelősség Feladatait időben, pontosan elkészítette; a rábízott területeken alapos munkát végzett, gondos, lelkiismeretes munkát adott be.		
Kreativitás és eredetiség A felmerülő problémákat megoldotta; a közös munkát új ötletekkel gazdagította, érdekes megoldásokkal vitte előre.		
Kommunikációs képesség Hatékony volt a megbeszéléseken, figyelmesen meghallgatta a többiek véleményét is, lényegre törően tartott beszámolót a feladattal kapcsolatban; írásban jól dokumentálta a munkaeredményeket.		

5. ábra. Értékelőív csoportmunkához
Forrás: Goldfinch (1994) és Lejk-Wyvill (2001)

Egyéni és csoportmunkák értékelésekor szoktam kiosztani fejenként 3-4 kis színes lapot, amelyeket mindenki tetszőlegesen használhat fel (ha akarja, mindet akár egy embernek/csoportnak is adhatja, de akár szét is oszthatja), hogy értékeljen egy teljesítményt. Ha a feladat az volt például, hogy minden csoport ismertesse egy hotel általa tervezett különleges, nyári kínálatát, akkor a végén körbejárnak, és a cédulákkal szavaznak a legjobbakra, így a kapott lapok száma jelöli ki a végén a legjobbakat.

Sokszor a projekt- vagy egyéni munka elkészülte után az osztályteremben kiteszük a padokra a munkákat, körbejárva minden tanuló/kiscsoport végigolvassa a többiekét, és egy lapra részletesen leírja az adott munkáról a véleményét (kiegészítésül ponttal is értékelhetnek). Ezután bő értékelés vár a helyén minden csoportot. Óriási izgalommal olvassák, majd vitatják meg mindezt ilyenkor egymás között.

Változatok a dolgozatírásra

Inspiráló, izgalmas a diákok számára, ha egy kicsit átvehetik a tanár szerepét. A következő módszerrel (kivitelezhető mind hagyományos, mind digitális formában) háromszor gyakoroltatjuk észrevétlenül az anyagot a gyerekekkel. Feladatuk, hogy az adott anyagrészből állítsanak össze egy változatos feladattípusokból álló dolgozatot (ehhez minden felszerelésüket használhatják), instrukciókkal, pontozással, osztályzási kategóriákkal, százalékokkal stb. A helyes megoldásokat a füzetükbe kell lejegyezniük. Mire a feladatlappal elkészülnek, alaposan átismétlik az anyagot a tankönyvből, munkafüzetből, füzetből. A megadott idő lejártá után az elkészített dolgozatokat kicseréljük, és mindenki valaki másét oldja meg – minden felszerelést használhatnak hozzá. Legutolsó lépésként következik a dolgozatok javítása: a feladatlapok visszakerülnek az eredeti gazdájukhoz, aki kijavítja azokat, immár harmadszor tekintve át a kérdéseket, a válaszokat és a helyes megoldásokat. Természetesen pontozza és „osztályozza” is a munkát, majd visszaadja a „gazdájának”, aki átnézi, mit hogyan kellett volna írnia. Ha esetleg kérdés merül fel, azt megvitatják (megvitatjuk). Még ha nem is kerülnek be az osztályzatok a naplóba, az egymástól kapott jegyek fontosak a számukra.

Amikor egy adott csoportnál erősíteni akarom ennek a módszernek a hasznosságát, akkor beszédem a próbadolgozatokat, és átnézem, mi az, amit a tanulók leginkább fontosnak tartottak az anyagból, amivel legtöbbször sikeresen birkóztak meg; ezekből aztán a magam által összeállított dolgozatba is jócskán átemelek. Mindez megerősíti a gyerekekben, hogy mindannak, amit órán csinálunk, haszna van, hogy érdemes együttműködni, gyakorolni.

Az is jó módszer, hogy diákok véletlenszerűen vagy általunk kiválasztott párral is írhatnak dolgozatot, közben válaszaikat megbeszélhetik, egymás erősségeit, tudását, készségeit szabadon kiaknázhatják. A kapott pontszámot aztán a maguk belátása szerint válthatják osztályzatra, mérlegelve, ki mennyivel járult hozzá a közös munkához, ki mennyit érdemel. Például 10-ből 7 pontot értek el: ki kap ebből hármat (vagyis hármast), s ki négyet (vagyis négyest).

A tanulói motiváltságot nagyban fejleszti, ha a diák úgy érzi, beleszólhat a tanulási folyamatába, sőt a számonkérésbe is. Módszertani vezetőként módomban volt végigkísérni iskolámban az alábbi innovatív gyakorlatot. Az elmúlt évek egy, módszertani pályázaton díjat nyert, már itthoni és külföldi szakmai karriert is befutott, több konferencián, pedagógiai napon sikeresen szerepelt módszertani újítás (*Student Voice*) az a matematikai gyakorlat, amely kollégám, Eőri János nevéhez fűződik, aki így összegezte módszerét:

„Facebook csoportban lehet szavazásra bocsátani feladatokat, hogy mik kerüljenek be a témazáróba. Amelyik feladat megkapja az osztálylétszámnak megfelelő szavazatok több mint felét, az a típus automatikusan bekerül a dolgozatba. A dolgozatba kerülő feladatok kiválasztásakor egyfajta stratégiai játék kezdődött” (Ijjas 2018). „A legjobb tanulók eleinte nem voltak hajlandóak támogatni a legkevesebb pontot érő feladatokat, hiszen ezzel korlátozódhat az elérhető osztályzat. Hasonló helyzet alakult ki a másik oldalon a legnehezebb feladatokkal kapcsolatban. Miután mindenki elsajátította a legkönnyebb alappéldákat, a tanulók a közepes pontértékűek felé nyitottak. Mindenki igyekezett az általa elsajátítottak mellé szavazótársakat keresni. Ezen feladatokat vagy én magyaráztam el, vagy egymást képezték a diákok. Mindkét esetben a tanítási idő meghosszabbodásáról beszélhetünk, sőt utóbbi esetben erősödött az együtt tanulás élménye, melynek során a tanulók egymást segítették. Mivel a stresszmentes matekdolgozat Facebook csoportjába befotóztam jegyzeteket, illetve megoldott feladatokat is beküldhetek, ezáltal a telefonjukon a diákok bárhol ránézhetnek a társaik által megosztott tananyagokra, és akár közlekedési eszközön is foglalkozhatnak a matematikával. Kisebb tanulócsoportok alakultak, szünetekben az iskola termeit, folyosóit járva láthatjuk a közösen matekózó diákokat. A gyakorlás során és a visszajelzésekből a tanuló pontosan tudta, hogy melyik részre kell a későbbiekben több energiát fordítani. A diákok gyakrabban mertek kérdezni, hiszen jobban átlátták a tananyagot és annak nehézségeit is. A sikerélmények következtében az oktatáshoz is pozitívabban viszonyultak” – idézi cikkében Ijjas Tamás egy matematikatanár szavait a tapasztalatokról.

Komplex értékelésre alkalmas tevékenységek

Bármilyen téma alkalmas lehet az írásbeli kifejezőkészség komplex munkává ötvözésére: mindenki készít egy terjedelmes, színes képes folyóiratot saját írásokkal, novellákkal, versekkel, receptekkel, rejtvényekkel stb. Mindig külön órát szánunk rá, hogy egymás gondos kivitelű, tartalmas, gazdag, szeretettel készített újságját csereberéljék, elolvassathassák, spontán dicsérjék egymást, szöveges megjegyzésekkel értékeljék – ezek az értékelések lehetnek kötetlenek is, de megadhatunk előzetes szempontokat is.

A dramatizálás, hosszabb-rövidebb jelenetek előadása nagy népszerűségnek örvend. A mi iskolánkban tanév végén maguk írnak és adnak elő egy idegen nyelvi darabot a tanulók. Ez segít, hogy a nyelvhasználat „elszakadjon” a tanulástól, és immár valós, alkalmazott, színes, komplex eszközhasználatú lépjen elő.

Előre egyeztetett időpontban a tanulók sokszor előadást, vetítéssel egybekötött prezentációt tartanak a tanév folyamán országismereti témákról, idegen nyelven olvasott könyvről vagy egy őket érdeklő területről. Az értékelés szempontjait előre kézbe kapják, illetve közösen kialakítjuk. A csoport tagjai ezek alapján pontszámokkal, kulcsszavakkal írásban vagy szóban kifejtve értékelik egymást (6. ábra).

<i>KRITÉRIUMOK</i>	<i>MEGFIGYELÉSI SZEMPONTOK</i>	<i>Pontok 1–20</i>
FELKÉSZÜLTSG	Az előadó nagyon alaposan felkészült, láthatólag gazdag anyagot gyűjtött, a témában alaposan elmélyült.	
IDEGEN NYELVI MEGFORMÁLTSÁG	Mondanivalóját megfelelő nyelvi szinten formálta meg, a hallgatóság érti a kiselőadást, a szöveg nyelvi nem hibás. Az esetlegesen fontos új szavakat előzetesen felírta a táblára / menet közben megtanította.	
ELŐADÁSMÓD	Az előadó folyamatosan, megakadás nélkül beszél, látszik, hogy begyakorolta a szöveget. Beszéde jó tempójú, értelmező, magyarázó, folyamatosan kapcsolatot tart hallgatóságával (pl. szemkontaktus).	
TECHNIKAI MEGVALÓSÍTÁS	A kiselőadást kísérő vetítés jól kiegészíti a beszédet, azt teljesebbé teszi. A képek, ábrák, grafikonok esztétikusak, jól láthatóak, jól követhetőek.	

6. ábra. Prezentáció értékelési szempontjai

A fentieket összegezve elmondhatjuk, mindezen értékelési formának csak akkor van értelme, ha hozzáférhető a diákok és szülők számára is, és áttekintésük, afféle előrevivő értékelésük után kijelölik a fejlődés további útját. Karácsony Sándor szerint a „nevelés” azt jelenti: „növelés” (Karácsony 1940). Növeljük hát tanítványainkat tudásban, képességeikben-készségeikben, önbizalomban, nyelvtanulási kedvben.

IRODALOM

- Goldfinch, J. (2006): Further Developments in Peer Assessment of Group Projects, *Assessment & Evaluation in Higher Education*, 19/1, 29–35. (2019. 06. 10.)
<https://doi.org/10.1080/0260293940190103>
- Ijjas Tamás (2018): *Hogyan alakítják a tanulási folyamataikat egy budapesti szakgimnázium diákjai?* *TanTrend Oktatási Portál*, 2018. 08. 07. <http://tantrend.hu/hir/hogyan-alakitjak-tanulasi-folyamataikat-egy-budapesti-szakgimnazium-diakjai> (2019. 05. 04.)
- Karácsony Sándor (1940): *Demokrácia és pedagógia*. Budapest: Exodus.
- Lejk, M. – Wyvill, M. (2010): Peer Assessment of Contributions to a Group Project: A comparison of holistic and category-based approaches, *Assessment & Evaluation in Higher Education*, 26/1, 61–72. (2019. 06. 10.) <https://doi.org/10.1080/02602930020022291>

VARGA ZOLTÁN BALÁZS - SALAMON ANIKÓ

A fejlesztő értékelés gyakorlata a nyelvoktatásban egy szakképzési centrumban

Language teachers' practice of assessment for learning in a VET center

Our study aims at analyzing assessment for learning in teachers' everyday pedagogical practice. Before the investigation, 72 teachers participated at the assessment training organized by our center. In our questionnaire, we asked about their assessment principles and elements of supportive learning environment. 45 teachers answered the questions. In line with this, a similar questionnaire was administered with the participation of 81 language teachers. Among the results of the study, we show similarities and differences between the two target groups and give an insight into teachers' experience stemming from students, colleagues and parents.

Keywords: assessment, vocational education, practice, language teaching, teaching and learning

A Budapesti Gazdasági Szakképzési Centrum kiemelt figyelmet fordít a pedagógiai munka modernizálására, az innovatív oktatási és pedagógiai módszerek elterjedésének támogatására. A centrumban ősszel megrendezésre kerülő Pedagógiai Napok és a pedagógusok számára tanévenként meghirdetett módszertani pályázatok megeremtik a korszerű pedagógiai ötletek és eszközök iskolák közötti megismertetésének lehetőségét. Az elmúlt tanévek folyamán szervezett továbbképzések, valamint az egységesen bevezetett bemeneti mérés a pedagógiai értékelés területét a pedagógusok figyelmének fókuszába helyezte.

Hagyományosan az értékelési funkciók közül leginkább a szummatív, illetve a diagnosztikus értékelést alkalmazzák a pedagógusok, utóbbira jó példa az iskolába való belépéskor a nyelvi szintfelmérő íratása, amelynek eredménye alapján csoportokba sorolják a tanulókat. Kevésbé alkalmazznak tudatosan formatív értékelést a pedagógusok, miközben az osztályzat nélküli visszajelzés nélkülözhetetlen eleme az oktatásnak és a tanulásnak. A formatív értékelés nemzetközi szintű kutatásaiban a tanulás eredményének értékelése (*assessment of learning*) mellett egyre inkább teret nyer a tanulást támogató értékelés (*assessment for learning*). Utóbbi során bevonjuk a tanulókat a tanulási célok meghatározásába, ösztönözzük őket az önértékelésre, megerősítjük őket abban, hogy mindenki képes fejlődni. Ugyanakkor a mindennapokban kihívást jelent, hogy a pedagógusok hajlamosak a tanulás minősége helyett inkább az elvégzett munka mennyiségét értékelni, nagyobb hangsúlyt helyeznek az osztályozásra, mint a fejlődést szolgáló tanácsokkal való támogatásra, illetve gyakran egymáshoz hasonlítják a tanulókat,

ami a kevésbé sikeres diákokra demoralizáló hatással van (Broadfoot et al. 1999). A fejlesztő értékelés a tanulók fejlődésének és tanulásának gyakori, interaktív módon történő értékelését jelenti, melynek célja a tanulási igények meghatározása, és a tanítás azokhoz való igazítása (OECD 2005). A fejlesztő értékelést nem tekinthetjük egy az egyben formatív értékelésnek, hiszen nem a külső célok megvalósítására, hanem a fejlődés folyamatára összpontosít (Lénárd–Rapos 2009).

Gyakorlati igény egy új típusú értékelésre

A centrum iskoláiban a fejlesztő értékelést a 2017 tavaszán elindított Itemműhely foglalkozás folytatásaként terveztük közelebb hozni a pedagógusokhoz. Az Itemműhely egy kétórás műhelymunka volt, amit az elmúlt két tanévben hét alkalommal hirdettünk meg a pedagógusok számára, összesen 44 kolléga vett részt ezeken az alkalmakon. Tematikájában a mérés-értékelés osztálytermi lehetőségeire fókuszált. Rövid elméleti áttekintés után gyakorlati feladatszerkesztési rész következett konkrét példákkal. Áttekintettük a feladattípusokat, megvizsgáltuk a feladatok egyes részeit. A zárt típusú feladatoknál felmerült, hogy mennyire sokféle feladat létezik, amiből válogathatunk és megbeszéltük, hogyan lehet ezeket – céljainknak megfelelően – a lehető legjobban összeállítani. Ugyanakkor az itt tárgyalt tanulóértékelés a feladatlapokra korlátozódott, így gyakran merültek fel további értékelési dilemmák (objektivitás, igazságosság, méltányosság, motiváció) a pedagógusok részéről. Időközben a centrum pedagógiai elképzelései is a tanulási eredmény alapú megközelítés felé fordultak: felmerült annak a létjogosultsága, hogy a tanulók és szüleik megismerhessék egy-egy tantárgyra vonatkozóan egy adott tanóra, egy témakör, egy félév vagy az egész tanév tanulási eredményeit, tisztában lehessenek azzal, mire készíti fel a tanulókat az iskola. Ez az igény találkozott a belső továbbképzés megtartására felkért szakember elképzeléseivel, aki egy Itemműhely-foglalkozásunkon meghívott vendégként megismerte több pedagógusunk értékelési dilemmáit, és kidolgozott számunkra egy, a tanulási eredményekre is építő tréninget „A fejlesztő szemléletű értékelés a gyakorlatban” címmel. A tréning során a résztvevők megismerték a jövő fókuszú, tanulási eredmény alapú értékelés jelentőségét. Beszéltek a hibázás, a kísérletezés tanulást segítő szerepéről (Knausz 2008). Megismerték a tanulásieredmény-alapú értékelés lényegét. Gyakorolták a fejlesztő hatású értékelő kommunikációt, és megismertek néhány, fejlesztő értékelésre használható eszközt.

A tréning négy alkalommal került megszervezésre, az első két alkalmon iskolánként két-három vállalkozó pedagógusnak nyílt lehetősége részt venni, ezeket követte az igazgatók számára szervezett továbbképzés (vezetői szemmel újragondolva a témát), majd az egyik iskola teljes nevelőtestülete részt vett a tréningen. Így összesen 72 pedagógus nyert betekintést a 2018. május és szeptember közötti időszakban a fejlesztő értékelés gyakorlatban is alkalmazható eszközeibe és módszereibe.

A tréning utánkövetése és felmérés a nyelvtanárok körében

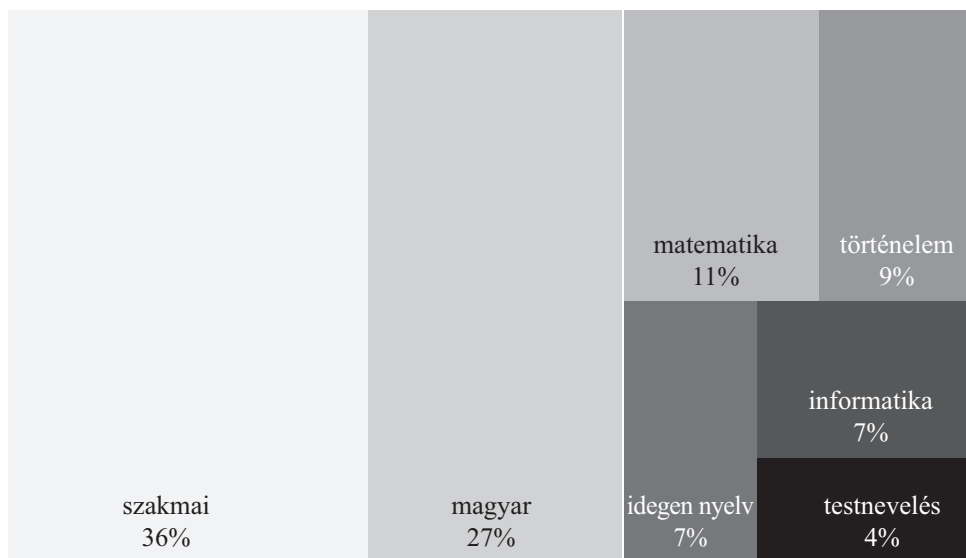
A képzésen részt vett kollégák körében kérdőíves utánkövetést végeztünk, amelyben rákérdeztünk az egyes értékelési elvek fontosságára, illetve érvényre juttatásukra az oktatás során. Felmértük, hogy a támogató tanulási környezet elemeit milyen mértékben építik be óráikba a pedagógusok, végül nyitott kérdéseket fogalmaztunk meg a fejlesztő értékeléssel kapcsolatos negatív/pozitív tapasztalatokra vonatkozóan.

Feltételeztük, hogy a nyelvtanárok értékelési gyakorlata közelebb áll a fejlesztő szemlélethez, ezért az utánkövetéssel párhuzamosan a centrumban tanító nyelvtanárok számára is elkészítettünk egy hasonló kérdőívet, amelyet kiegészítettünk az érdemjeggyel, illetve érdemjegy nélkül alkalmazott szóbeli és írásbeli értékelési módszerek használatára vonatkozó kérdéssel.

Jelen tanulmányban röviden bemutatjuk a tréningen részt vett kollégák utánkövetésére létrehozott kérdőív eredményeit. Ezt követően a nyelvtanárookra fókuszálva megvizsgáljuk, hogy a centrum 19 iskolájában oktató nyelvszakos kollégák számára mennyire ismert a fejlesztő értékelés, milyen gyakorisággal és mélységben alkalmazzák azt a tanórákon. Kitérünk arra is, milyen sikerélményeik és kihívásaik vannak ezen a területen napi munkájuk során.

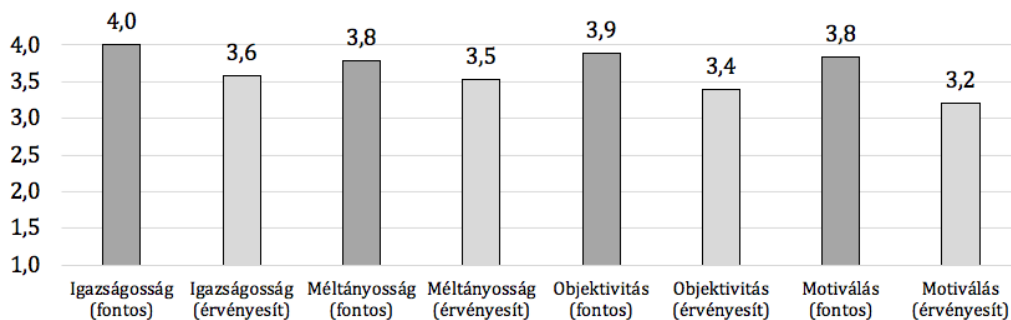
Eredmények

Kérdőívünket 45 kolléga töltötte ki, tanított tantárgyuk szerint az *1. ábrán* látható megoszlásban.



1. ábra. A kérdőívet kitöltők tanított tantárgy szerinti összetétele

Az igazságosság, a méltányosság, az objektivitás és a motiválás mint értékelési elvek tekintetében négyfokozatú Likert-skálán kértük a tanárok válaszát arra vonatkozóan, hogy a) mennyire tartják fontosnak ezeket az elveket, b) mennyire tudják érvényre juttatni ezeket munkájuk során. A válaszok (2. ábra) alapján megállapítottuk, hogy mind a fontossághoz, mind pedig az érvényre juttatáshoz magas átlagok kapcsolódnak, ugyanakkor az elvek érvényre juttatása rendre alacsonyabb, mint a fontosság.



2. ábra. Az értékelési elveknek tulajdonított fontosság és érvényre juttatásuk a mindennapokban

Arra a kérdésünkre, hogy a tréningen elhangzottakat be tudták-e építeni munkájukba a kollégák, 36 fő válaszolt igennel és 9 fő nemmel. A felmérést március közepén végeztük, eddig az időpontig az első két csoport tréningje óta tíz, a nevelőtestületi tréninget követően hat hónap telt el. Az elvek tekintetében a fenti két csoport válaszai között egy esetben adódott szignifikáns különbség: a tanultakat nem beépítő kollégák esetében a méltányosság mint elv fontosabbnak bizonyult, mint a tanultakat beépítő tanárok között ($x_1 = 4$, $x_2 = 3,72$, $t = -3,25$, $p < 0,05$).

Azok, akik beépítették a tréningen tanultakat napi munkájukba, az alábbi elemeit nevezték meg a megszerzett tudásnak: célok megfogalmazása, ön- és társértékelés, motiválás, tevékenység alapú tanulás. A tréning fent részletezett tematikája szempontjából megállapítható, hogy a pedagógusok által szabadon megnevezett elemek nemcsak a gyakorlatban alkalmazható eszközökre, hanem a képzésen elsajátított mélyebb összefüggésekre (célok, motiváció) is utalnak.

Kíváncsiak voltunk arra is, hogy a pedagógusok milyen mértékben illesztik be munkájukba a támogató tanulási környezet egyes elemeit. A kérdőívet kitöltő tanárok saját bevallásuk alapján

- gyakran használják az egyértelmű elvárásokat (négyfokozatú skálán az átlag: 3,85; szórás: 0,4), a tanulói együttműködést (átlag: 3,47; szórás: 0,5) és a tevékenység alapú tanulást (átlag: 3,22; szórás: 0,5),

- ritkábban adnak teret a tanulói autonómiának (átlag: 2,91; szórás: 0,6) és az egyéni tanulási utaknak (átlag: 2,67, szórás: 0,7).

Megvizsgáltuk, hogy a támogató tanulási környezet egyes elemeinek használata összefügg-e az értékelési elvek érvényre juttatásával. Külön-külön vizsgáltunk valamennyi lehetséges kapcsolatot, és számos, statisztikailag jelentős összefüggést kaptunk (1. táblázat):

- a motiválás érvényre juttatása összefügg az egyéni tanulási utak gyakoribb használatával, a tanulók közötti együttműködés gyakoriságával és a tanulói autonómia növelésével,
- a méltányosság nagyobb mértékben érvényesül ott, ahol nagyobb a tanulói autonómia, illetve gyakrabban van lehetőség egyéni tanulási utakra,
- az igazságosság fokozottabban jut érvényre, amikor egyéni tanulási utak is megjelennek a pedagógiai gyakorlatban.

	Igazságosság	Méltányosság	Objektivitás	Motiválás
Tevékenységalapú tanulás	n. s.	n. s.	n. s.	n. s.
Tanulói autonómia	n. s.	0,36	n. s.	0,33
Együttműködés	n. s.	n. s.	n. s.	0,34
Egyéni tanulási utak	0,37	0,43	n. s.	0,44
Egyértelmű elvárások	n. s.	n. s.	n. s.	n. s.

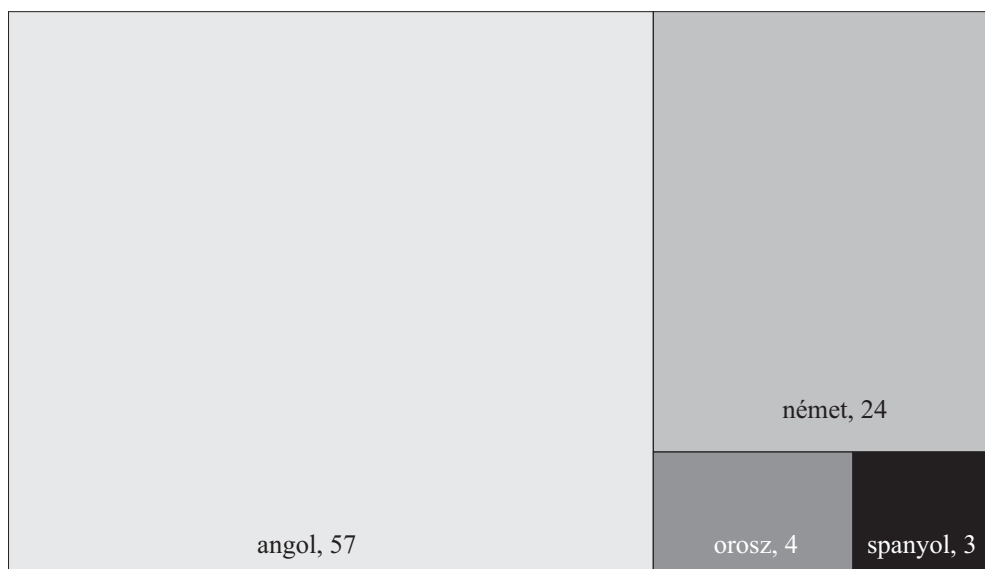
1. táblázat.¹ Összefüggés a támogató tanulási környezet egyes elemeinek használata és az értékelési elvek érvényre juttatása között

Az objektivitás érvényre juttatásával nem találtunk szignifikáns kapcsolatot egyik támogató tanulási környezeti tényező esetén sem. Megállapíthatjuk, hogy az egyéni tanulási utak alkalmazásával a pedagógusok válasza (tapasztalatai) alapján az igazságosság, a méltányosság és a motiválás elve is jobban érvényesül az értékelési gyakorlatban. Fontos azonban megjegyezni, hogy vizsgálataink eredménye együttjárásra és nem ok-okozati kapcsolatra utal a fenti tényezők között.

¹A táblázatban szereplő Spearman-féle korrelációs együtthatók $p = 0,05$ szinten szignifikánsak

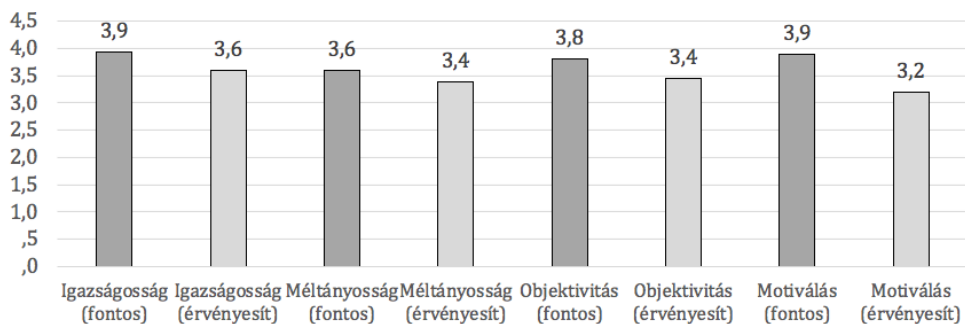
Fejlesztő szemléletű értékelés a nyelvoktatásban

A tréning utánkövetésével egyidejűleg mértük fel a centrumban tanító nyelvtanárok körében az értékelés fejlesztő szemléletű összetevőit. Kérdőívünkre 81 válasz érkezett, ez a 19 tagintézményben tanító nyelvtanárok kicsit több mint felét jelenti. (Amennyiben egy nyelvtanár a tréningen is részt vett, azt kértük, hogy ő csak a nyelvtanárok részére készített kérdőívet töltsse ki, ugyanakkor ezt nem minden esetben vették figyelembe a kollégák.) A kérdőív kitöltői közül 45 fő nyilatkozott úgy, hogy vett már részt fejlesztő értékelés témájú továbbképzésen, míg 24-en nemleges választ adtak. 12 fő a talán válaszlehetőséget jelölte meg, ami utal a téma újszerűségére, esetleg a kifejezés homályos voltára is. A kérdőívet kitöltők angolt vagy németet tanítanak, kisebb számban jelentek meg a spanyol és az orosz nyelvet oktatók. A 3. ábra a válaszadók tanított nyelv szerinti összetételét mutatja.



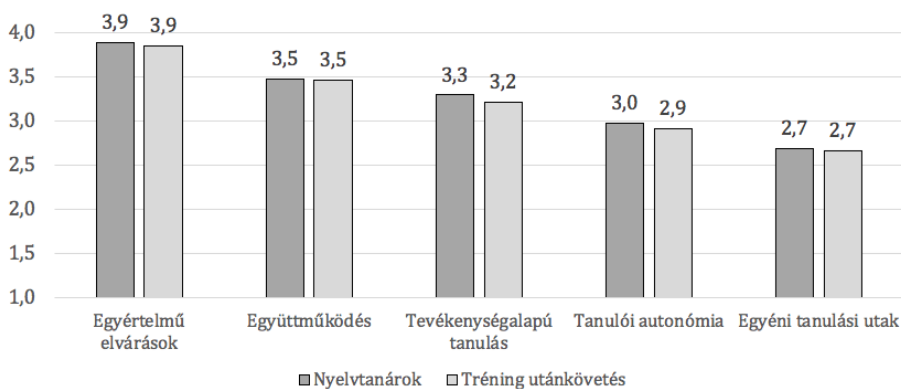
3. ábra. A kérdőívet kitöltők tanított nyelv szerinti összetétele

Az értékelési elvek hasonló képet mutattak, mint az utánkövetéses vizsgálat esetében (4. ábra). Ugyanakkor a csak nyelvtanárokból álló mintában a fontosság és a gyakorlati érvényre juttatás közötti különbségek kisebbnek bizonyultak. A nyelvtanárok tehát (a motiválás kivételével) következetesebbnek tűnnek válaszaik alapján.



4. ábra. Az értékelési elveknek tulajdonított fontosság és érvényre juttatásuk a gyakorlati munkában a nyelvtanárok esetében

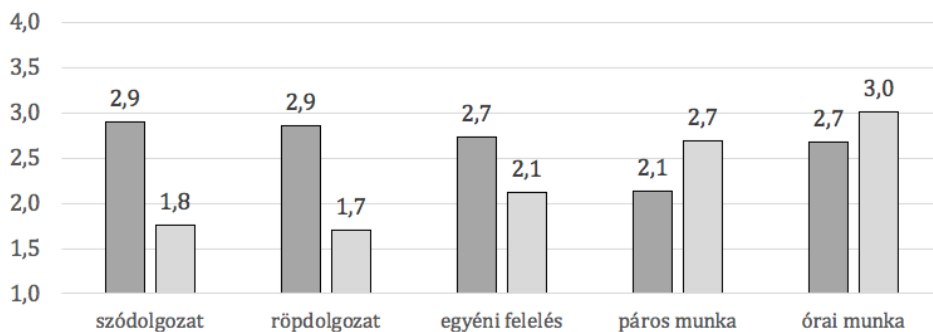
A támogató tanulási környezet elemei a nyelvtanárok között hasonló gyakorisággal jelennek meg a pedagógiai gyakorlatban, mint ahogyan azt a tréningen részt vett pedagógusok utánkövetésénél tapasztalhattuk (5. ábra).



5. ábra. A támogató tanulási környezet egyes elemeinek használati gyakorisága

Az 5. ábra megerősíti azt a feltételezésünket, miszerint a nyelvtanárok módszertanában a tréningtől függetlenül is érvényesül a fejlesztő szemléletű megközelítés. A tizedpontosságig hasonló átlagok megítélésünk szerint inkább egy közös gyökerekből táplálkozó pedagógiai kultúrát tükröznek.

A nyelvtanárok kérdőívét bővítettük egy kérdéssel. A hagyományosan alkalmazott értékelési eszközökkel (szódolgozat, röpdolgozat, egyéni felelés, páros munka, órai munka) kapcsolatban megvizsgáltuk, hogy használják-e azokat nyelvtanáraink. A kérdést két formában tettük fel: kíváncsiak voltunk arra, hogy érdemjeggyel való értékelésre milyen gyakran használják a fenti eszközöket, illetve érdemjegy nélküli értékelés milyen gyakran történik (ha történik) segítségükkel (6. ábra).



6. ábra. A hagyományos értékelési eszközök alkalmazási gyakorisága

A 6. ábrán közölt eredmények alapján a nyelvtanárok az órai munkát és a páros munkát gyakrabban értékelik szövegesen, mint érdemjeggyel. A szódolgozat, a röpdolgozat és az egyéni felelés értékelése gyakrabban történik érdemjeggyel, mint a nélkül. Az írásbeli számonkérések esetén természetesen gyakoribb az osztályozás, ugyanakkor az egyéni felelés nagyobb lehetőséget biztosít a fejlődés támogatására, ha érdemjegy nélkül, szöveges visszajelzést kap a tanuló, ami társainak is segítségül szolgálhat. Emellett érdemes lehet megfontolni a páros és az órai munka osztályozását is, hiszen a tanulók közötti együttműködést, a közös munkát így lehetséges felmérni, jutalmazni.

A tanulási eredményeket saját bevallásuk szerint három pedagógus kivételével valamennyien használják, azonban a további kérdésekre adott válaszaik alapján nem feltétlenül 'learning outcomes' értelemben. Utóbbi szerint „a tanulási eredmények tudás, képesség, felelősség és autonómia kontextusában meghatározott kijelentések arra vonatkozóan, hogy a hallgató mit tud, mit ért, és önállóan mire képes, miután lezárt egy tanulási folyamatot, függetlenül attól, hogy hol, hogyan, mikor szerezte meg ezeket a kompetenciákat” (CEDEFOP 2008). A tanulási eredmények inkább mint szókapcsolat ismert számukra, a továbbképzésen és egyéb fórumokon feltett kérdések alapján 'learning outcomes' értelemben még csak kevesen találkoztak ezzel a kifejezéssel. Jelen kutatásban a tanórákon, a középiskolai oktatásban alkalmazható tanulási eredményekkel foglalkoztunk, így most eltekintünk a fogalom képesítési keretrendszerekben alkalmazott jelentésétől. A tréningen pontosan meghatározásra került, hogy a tanulási eredmények arra a tudásban, képességben, hozzáállásban, illetve autonómiában bekövetkezett változásra vonatkoznak, amivel a tanuló a tanóráról való kilépéskor rendelkezik ahhoz a pillanathoz képest, amikor belépett oda. (A nyelvtanárok, akik közül sokan nem vettek még részt fejlesztő értékelés témájú tréningen, nem biztos, hogy ebben a kontextusban gondolkodtak a fogalomról.)

A tanulási eredmények gyakorlati alkalmazásával kapcsolatban 43 fő válaszolta azt, hogy akár tanórai szintű eredményeket is megfogalmaz, 26 pedagógus havi vagy

témához kapcsolódó eredményeket, kilenc tanár pedig csak éves vagy féléves eredm-nyeket állít össze. A 81-ből 50 pedagógus a tanulókkal együtt, 28 tanár egyedül foglal-mazza meg a tanulási eredményeket. Utóbbiak – egy kivétellel – megosztják a tanulókkal az általuk megfogalmazott tanulási eredményeket. A fenti arányok pozitív képet mutatnak, és bizakodásra adnak okot, hiszen a közösen meghatározott vagy a tanulókkal megosztott tanulási eredmények növelik a tanulók felelősségvállalását saját tanulási folyamataik-ban. A számok utalnak arra a fajta szabadságra is, amivel egy nyelvtanár rendelkezik a tanórákon (például egy más közismereti vagy szakmai tantárgyat oktató kollégájához képest).

Kérdőívünkben 66 nyelvtanár nyilatkozott arról, hogy alkalmaz fejlesztő értékelést pedagógiai gyakorlatában, 15-en adtak nemleges választ a kérdésre. A fejlesztő értékelést saját bevallásuk szerint alkalmazó és nem alkalmazó pedagógusokat két részmintának tekintve összehasonlítottuk válaszaikat, és kizárólag egyetlen esetben találtunk a két csoport között statisztikailag jelentős eltérést. A fejlesztő értékelést alkalmazó tanárok gyakrabban értékelik érdemjegy nélkül, szövegesen az órai munkát, mint az ezt az értékelési formát nem alkalmazó kollégáik ($x_1 = 3,1$, $x_2 = 2,6$, $t = -2,56$, $p < 0,05$). További eltérés nem adódott sem a támogató tanulási környezet elemeinek, sem a többi (érdemjeggyel vagy a nélkül alkalmazott) értékelési eszköz használatának gyakorisága, sem pedig az egyes értékelési elvek fontossága vagy érvényre juttatása terén.

A nyelvtanárok válaszai alapján is vizsgáltuk az együttjárást az értékelési elvek érvényre juttatása és a támogató tanulási környezet elemei között.

	Igazságosság	Méltányosság	Objektivitás	Motiválás
Tevékenységalapú tanulás	n. s.	n. s.	n. s.	0,38
Tanulói autonómia	0,27	0,29	n. s.	0,36
Együttműködés	n. s.	n. s.	n. s.	0,23
Egyéni tanulási utak	n. s.	n. s.	n. s.	0,26
Egyértelmű elvárások	n. s.	n. s.	n. s.	0,23

2. táblázat². Összefüggés a támogató tanulási környezet egyes elemeinek használata és az értékelési elvek érvényre juttatása között a nyelvtanároknál

² A táblázatban szereplő Spearman-féle korrelációs együtthatók $p = 0,05$ szinten szignifikánsak

A 2. táblázat alapján megállapítható, hogy a nyelvtanárok mintáján a motiválás érvényre juttatása összefügg valamennyi támogató tanulási környezeti elem használatának gyakoriságával. A gyakoribb tevékenységalapú tanulás, a megnövelt tanulói autonómia, a tanulók közötti fokozottabb együttműködés, a gyakrabban alkalmazott egyéni tanulási utak és egyértelmű elvárások mind-mind együtt járnak a motiválás pedagógiai gyakorlatban történő jobb érvényesítésével. Ez azért is lehet fontos, mert a 4. ábra alapján a motiválásnál tapasztaltuk a legnagyobb különbséget a fontosság és az értékelési elv érvényre juttatása között. A tanári válaszok alapján a támogató tanulási környezet valamennyi eleme támogatja a motiválás gyakorlati érvényesülését. Ugyanakkor itt is fontos elmondani, hogy a korreláció nem ok-okozati kapcsolatra, hanem együttjárásra utal. Beavatkozásra viszont a táblázat két dimenziója közül inkább a tanulási környezet változtatásával nyílik lehetőségünk.

A fejlesztő értékelés hasznosnak bizonyult számomra, amikor	A fejlesztő értékelés alkalmazása problémát okozott számomra, amikor
...egy érdemjeggyel nem lehet kifejezni a tanuló teljesítményét, vagy ha motiválni kellett a jobb eredmény elérésére, mert képessége alatt teljesített.	...egyidejűleg sok tanulót kellett értékelni.
...olyan diákkal vagy diákokkal találkoztam, akik valamilyen oknál fogva tanulási nehézségekkel küzdöttek.	...egy diák nem figyelt, nem koncentrált.
...új célokat fogalmaztam meg.	...nem voltak egyértelműek ezek az új célok.
...diákom bizonytalan volt saját teljesítményét illetően, többnyire negatívan értékelte magát, illetve nem volt tisztában az erősségeivel és fejlesztendő területeivel.	...diákom elzárkózott az együttműködéstől, az értékelést támadásnak vette.
...a tanulóval fogalmazzuk meg közösen, hogyan és miben kell változtatnia a sikerebb eredmény elérése érdekében. Ebből én is pontosabban fogom tudni, hogy az adott tanulónál melyik az a számonkérési módszer, amiben ő, a tőle telhető legjobb eredményt/tudását tudja megmutatni.	...a tanulónak nincs elég önismerete ahhoz, hogy elfogadja a fejlesztő útmutatást
...ugyanazon csoportban nagyon eltérő képességű tanulók munkáját kell összehangolnom.	...a diák hajlamos elkényelmesedni, ha nem kap érdemjegyet.
...a tanulóknak elmondhatom, hogy egyéni képességeiket figyelembe véve mennyit fejlődtek egy adott időszakban	...a tanulók nem motiváltak, csak a kettes elérése a cél.

3. táblázat. A fejlesztő értékelés sikere és nehézségei a tanárok válaszai alapján

Kérdőívünk végén a saját bevallásuk szerint fejlesztő értékelést alkalmazó pedagógusoktól konkrét esetek, példák felsorolását is kértük az alábbi témákban:

- Mikor bizonyult hasznosnak számukra a fejlesztő értékelés?
- Milyen esetekben okozott problémát?
- Mikor szerették a tanulók a fejlesztő értékelést?
- Milyen esetekben érkezett visszajelzés a szülőktől?
- Mikor keltette fel kollégáik figyelmét a téma?

Az alábbi két táblázatban kiemeltünk néhány választ. Megállapíthatjuk, hogy a legtöbb sikert a motiváció, a (tanulóval közös) célkitűzés, illetve a tanulási nehézségek legyőzése kapcsán tapasztalták meg a pedagógusok.

A 3. táblázat alapján nehézséget okoz, ha több tanulót kell egyszerre értékelni, illetve önmagában az új értékelési módszer (nem érdemjeggyel történő értékelés) bevezetése is problémákba ütközhet. A 4. táblázat a tanulói, szülői és a kollégák által adott visszajelzéseket tartalmazza a fejlesztő értékelésről, ahogyan azt a pedagógusok érzékelik a tanórákon, a fogadóórákon vagy a tanáriban, illetve az értekezleteken.

A fejlesztő értékelést szerették a tanulók, amikor	A fejlesztő értékeléssel kapcsolatban visszajelzés érkezett a szülőktől, amikor	A fejlesztő értékelés felkeltette a kollégáim figyelmét is, amikor
...négy szemközti, vagy legfeljebb a munkapartnere vagy kiscsoportja előtt hangzik el az értékelés, mint az osztály előtt, különösen, ha több negatívumra számítanak.	Nem volt ilyen.	...a tantestület előadást hallgatott a témában.
...közös megbeszéljük.	...a diák/diákok nem teljesítettek jól az iskolában.	...a diák nem jól vagy kitűnően teljesített.
...úgy érezték, ezek méltányosabbá tehetik a közös munkát, különösképpen az értékelést.	Ilyesmit nem tapasztaltam.	Ilyesmit nem tapasztaltam.
...megértették, hogy olyan területeken is sikert érhetnek el, amelyekről azt gondolták, hogy nem képesek teljesíteni.	Nem érkezett visszajelzés.	...a munkaközösség tagjai ötleteket adtak egymásnak.
...az erősségeikkel kezdtem az értékelést.	...a tanuló otthon is angolul válaszolt a szülőknél vagy társalgott a testvérekkel	...munkaközösségi értekezleten beszélgettünk róla
Alapvetően népszerű a tanulók körében.	Nem érkezett visszajelzés.	...továbbképzésen vettek részt.
...megértették, hogy a jegyeken kívül létezik más forma is arra, hogy lássák mégiscsak haladnak.	Nem érkezett.	Nem beszéltünk róla.

4. táblázat. A tanulók, szülők, pedagógustársak visszajelzései a fejlesztő értékelésről

A tanulók számára fontos, hogy személyesen vagy kiscsoportban kapjanak visszajelzést munkájukról. A szülőktől ritkán kaptak tanáraink visszajelzést. Amikor a mintában szereplő pedagógusok kollégáinak a visszajelzéseire kérdeztünk rá, azt tapasztaltuk, hogy ők legtöbb esetben egy továbbképzésen vagy előadáson elhangzottak kapcsán mondtak véleményt a fejlesztő értékelésről. Érdekes volt látni, hogy többször is megjelent a „nem beszéltünk róla” válasz. A kérdőívnek ebben a részében megosztott nyelvtanári tapasztalatok alapján viszonylag ritkán fordul elő közvetlen módszertani tapasztalatcsere a kollégák között, ami további lehetőségekre hívta fel a figyelmünket.

Összegzés

A fejlesztő szemléletű értékelésre mint modern pedagógiai eszköztárra, illetve megközelítésre hangsúlyt helyező, 19 tagintézménnyel működő szakképzési centrumban lehetőségünk adódott a nyelvtanárok értékelési módszereinek vizsgálatára egy, a témában tréninget végzett tanárcsoport által kitöltött kérdőív eredményeinek fényében. Kutatásunkban a tréning utánkövetésére használt kérdőív kérdéseit nyelvtanárainknak is feltettük, így két minta válaszainak összehasonlítására nyílt módunk. A vizsgálat rámutatott, hogy a támogató tanulási környezet elemeinek alkalmazása nagyon hasonló a két csoportot összevetve. Feltételezzük, hogy az eredmények háttérében a vizsgálatban részt vevő tanárok közös pedagógiai kultúrája állhat. Ugyanígy hasonlóságot tapasztaltunk az értékelési elvek fontossága és a mindennapi pedagógiai gyakorlatban történő érvényre juttatásuk között is a két csoport vonatkozásában. Az eredmények igazolták hipotézisünket, miszerint a nyelvtanárok módszertanában fontos megközelítés a fejlesztő szemlélet. A nyelvtanári mintán csak egy mutató esetében (órai munka szöveges értékelése) adódott eltérés a fejlesztő értékelést bevallásuk szerint alkalmazó és nem alkalmazó tanárok csoportja között. Feltételezhetően a támogató tanulási környezet, a szöveges értékelés, visszajelzés a nyelvtanári módszer-tannak jobban részét képezik. Ezzel az a lehetőség is megerősítést nyert, miszerint a nyelvtanárok értékelési szemlélete, tudása adaptálható szakmai és egyéb közismereti tárgyat tanító kollégáik körében, ezzel is gyarapítva a múlt helyett a jelen-, illetve jövőfókuszú értékelés (*assessment for learning*) egyre népszerűbb követőit iskoláinkban.

IRODALOM

- Broadfoot, P. – Daugherty, R. – Gardner, J. – Gipps, C. – Harlen, W. – James, M. – Stobart, G. (1999): *Assessment for Learning: Beyond the Black Box*. Nuffield Foundation and University of Cambridge.
http://www.nuffieldfoundation.org/sites/default/files/files/beyond_blackbox.pdf (2019. 07. 30.)
- CEDEFOP 2008: *Terminology of European education and training policy. A selection of 100 key terms*. Luxembourg: Publications office of the European Union.
https://www.cedefop.europa.eu/files/4064_en.pdf (2019. 07. 29.)
- Fejlesztő értékelés. A tanulást fejlesztő osztálytermi módszerek a középfokú oktatásban 2005*. OECD. CERI. Országos Közoktatási Intézet. Budapest, <http://www.ofi.hu/tudastar/tanulas-tanitasi/fejleszto-ertekeles>, (2019. 07. 23.)
- Knausz Imre (2008): *Mit kezdünk az értékeléssel? Adalékok az integrációs nevelés pedagógiájához*. Budapest: Educatio, https://tanitonline.hu/ftp/Mit_kezdzunk_az_ertekelessel.pdf, (2019. 07. 29.)
- Lénárd Sándor – Rapos Nóra (2009): *Fejlesztő értékelés*. Budapest: Gondolat Kiadó.

PROJEKTSZEMLE

Az eTransFair projekt portáljai: értékelési technikák és módszertani ötletek fordítást oktató szakemberek számára

A BME Idegen Nyelvi Központja (INYK) 2016. szeptember 1. és 2019. augusztus 31. között vett részt abban a nemzetközi projektben, amely az Erasmus+ Stratégiai Partnerség keretében nyert támogatást. Az *eTransFair* (<https://etransfair.eu>) névre keresztelt együttműködésben a két partnerrel, a bécsi Univerzitát Wien Fordítástudományi Központjával (Zentrum für Translationswissenschaft) és a madridi Hermes Traducciones fordítóirodával azt a célt tűztük ki, hogy a szakfordítóképzésben érdekelt három főszereplő – az egyetemek, a hallgatók és a fordítópia – közötti sinergiára építve egy innovatív tudásmegosztó adatbázist hozunk létre a szakfordítóképzés korszerűsítésének elősegítésére¹.

A projekt fontos eredménye volt a minden szakember számára ingyenesen elérhető két online platform, a *Módszertani portál*, valamint az *Értékelési portál*, amelyeket a fordítóképzésben érdekelt feleknek, elsősorban oktatóknak szánunk. A két portált azzal a céllal hoztuk létre, hogy a szakfordítást oktatók ötleteket meríthessenek más kollégák gyakorlati tapasztalataiból és osztálytermekben, valamint online felületeken alkalmazott technikából, és ezzel kibővíthessék és változatosabbá tehessék saját módszertani kellékárukat. Mindkét portált kulcsszavas módszerrel kereshetővé tettük, így minden felhasználó találhat számára érdekes feladatokat: a platformok szűrhetők az alkalmazandó nyelvpárra, a fejlesztendő kompetenciaterületekre, az elérendő tanulási eredményekre; adott szakterületre vagy a feladatra szánt időkeretre is.

Az alábbiakban elsőként az Értékelési portált mutatjuk be, majd röviden összefoglaljuk a Módszertani portál ismérveit, végül beszámolunk arról is, hogyan alakítottuk a projekt más eredményeit (ön)értékelési eszközzé.

¹ A projekt hivatalos angol címe: „How to Achieve an Innovative, Inclusive and Fit-for-Market Specialised Translator Training”.

Az Értékelési portál²: értékelési módszerek és technikák a fordítást oktatásában

Az eTransFair projekt egyik legértékesebb eredménye az *Értékelési portál (Pool of Assessment Techniques, PAT)*, melynek célja, hogy a tudásmegosztás lehetőségeit kihasználva egy olyan online elérhető, többnyelvű gyűjteményt hozunk létre, amelyben a hagyományos értékelési technikákon túl a nem hagyományos, fejlesztő jellegű értékelési módok, elsősorban az ön- és a társas értékelési technikák is hangsúlyos szerepet kapnak. Az értékelési technikákat elsősorban a fordítás oktatásában lehet alkalmazni, de több olyan technika is megtalálható a felületen, amely a tolmácsképzésben vagy az idegennyelv-oktatásban is felhasználható. Az értékelési technikák összegyűjtését célzó folyamat első szakaszában ki kellett alakítanunk egy olyan fejlécet (sablont), amelynek segítségével minden lényeges információt beszerezhetünk, s amely alapján a portálon hatékonyan kereshetnek a felhasználók. Hasonlóan a Módszertani portál fejlécéhez, a sablon hosszas konzultációs folyamat eredményeképpen jött létre, melynek során az első szakaszban segítségünkre voltak a BME TFK szakfordítást oktató belső munkatársai, később pedig a társ-képzőintézményekben oktató kollégák (lásd BME TFK Tanár–Tanár Találkozók). Itt szeretnénk köszönetet mondani azoknak a szaktanácsadóknak is, akik a fejléc véglegesítésének utolsó fázisában értékes javaslatokkal szolgáltak (Dróth Júlia – KRE, Einhorn Ágnes – BME INYK és Alexandra Krause – UniVie).

Az Értékelési portálra ötleteket feltöltő oktatóktól azt kérjük, hogy a *kötelezően kitöltendő mezőben* az alábbi adatokat adják meg:

1. az *alkalmazási területet*, amelyben meghatározzák, hogy jellemzően hol használható az adott ötlet, például a) tanteremben (vagy online), b) fordítói megbízás során a piacon, c) félévi vagy év végi számonkérés részeként; illetve d) diplomafordítások értékelésére záróvizsgákon (diploma- vagy oklevélszerzés esetén);
2. az *értékelő személyét*: önértékelés, társértékelés, oktató vagy külső személy, pl. fordítóirodai lektor;
3. az *értékelés célját*: diagnosztikus, szummatív vagy fejlesztő céllal értékelünk;
4. az *alkalmazott értékelési technikát*: pontozás, megjegyzések, ezek jegyekké alakítása, hibák súlyozása, esetleg egyéb;
5. a *képzési programot és a kurzust*, ahol az ötlet alkalmazható (amennyiben meghatározható);
6. illetve kérjük a *feladat leírását* – itt azt várjuk a feltöltőktől, hogy világosan elkülöníthető lépésekben írják le az ötletet/feladatot, lássák el címmel, javaslatokkal és/vagy megjegyzésekkel, és határozzák meg, hogy a szóban forgó ötlet miként fejleszt a tudást, a készségeket, illetve az attitűdöt/szemléletmódot.

²A portál közvetlenül elérhető ezen a linken: <https://hu.etransfair.eu/e-portalok/ertekelesi-portal>

The screenshot shows the eTransFair.eu website header with navigation tabs: RÓLUNK, EXTRA, HÍREK, e-Portálok, and KAPCSOLAT. The main content area is titled "2. Értékelés típusa:" and contains two sections:

2.a Az értékelő személye alapján:

- hallgatói önértékelés
- hallgatói társas értékelés
- tanári értékelés
- külsős értékelés (pl. fordítóirodai lektor)

2.b Az értékelés célja alapján:

- diagnosztikus értékelés (felvételi, csoportbesorolás)
- fejlesztő értékelés (előmenetel nyomon követése és mérése a tanulási folyamat során történő visszajelzéssel)
- szummatív értékelés (tanulási eredmény számszerűsített mérése)

1. ábra. Az Értékelési portál megjelenítése az eTransFair honlapján

A fejléc *opcionális* része tartalmazza a feladatot feltöltő(k) nevét, továbbá a feladathoz felhasznált forrást (hivatkozás), a szükséges eszközöket, és a feladat elvégzéséhez szükséges becsült időkeretet. A keresőbe írt szavak alapján a rendszer kigyűjti a kulcsszóhoz kapcsolódó feladatokat. (Pl. a „társas” szó beírásakor minden olyan feladat megjelenik, ahol ez a szó szerepel a leírásban.) Látható (lásd 2. ábra), hogy egy adott feladat/ötlet akár több nyelven is elérhető.

Módszertani portál: tantermi és online keretben alkalmazható ötletek fordítást oktatóknak

Az eTransFair *Módszertani portálján*³ elsősorban olyan módszertani, didaktikai ötleteket gyűjtöttünk össze, amelyeket a fordítást oktató kollégák alkalmaznak a tanórákon, vagy blended képzésben/távoktatásban online platformokon. Hasonlóan az értékelési portálhoz, az ötletek megosztására itt is kialakítottunk egy sablont, amelyben a feltöltést végző kollégáktól kérünk néhány alapadatot (*kötelező mező*), illetve megadhatnak egyéb, a kereshetőség szempontjából kevésbé fontos információkat (*opcionális mező*). A kötelezően megadandó elemek között található a feladat címét, a fejlesztendő kompetenciaterületeket, tanulási eredményeket, szakterületet (pl. építészet, jog) és a nyelvpárt (vagy jelezni lehet, ha a feladat nyelvfüggetlen); míg az opcionális információk között

³ Közvetlenül elérhető az alábbi linken: <https://hu.etransfair.eu/e-portalok/modszertan>

The screenshot shows the eTransFair.eu website interface. At the top, there is a navigation bar with the logo and several menu items: RÖLUNK, EXTRA, HÍREK, e-Portálok, and KAPCSOLAT. The main content area is titled "Keresési találatok "társas" kifejezésre". It lists three search results, each with a folder icon and a brief description:

- TÖBB SZEM TÖBBET LÁT**: ... megoldásait (puha készségi). Fischer Márta és Szabó Csilla, BME TFK 5 perc (önellenőrzés) + 10 min (társas) + 10 perc pármunka megbeszélés + 10 perc csoportos megbeszélés = összesen 35 perc nyomtató, vagy PC, Google ...
- THE MORE YOU LOOK THE MORE YOU SEE**: ... = 35 min in total paper or PC, Google Drive or simple word document (using track changes) Fischer, M. 2011. „A társas és önértékelés szerepe a fordítás oktatásában”. In: Dróth J. (szerk.) Szaknyelv és Szakfordítás. ...
- A legjobb megoldás**: A legjobb megoldás hu hu - x javítás és jobbitás nappali és/vagy esti, kontaktórák/blended (szak)fordítás óra A legjobb megoldás 1. A hallgatók az otthon elkészített fordításaikat feltöltik a Moodle-ba. A tanár ...

On the right side of the third result, there is a circular icon with an upward-pointing arrow.

2. ábra. A keresési eredmények megjelenítése az Értékelési portálon

megadhatjuk az alkalmazott osztálytermi formát (egyéni, pár- vagy csoportmunka), az alkalmazott értékelés típusát, a feladatra szánt becsült időkeretet, illetve a felhasznált anyagokat; szakirodalmat. A feltölteknek világosan elkülönített lépésekben kell vázolni a feladatot, megadni a megfelelő instrukciókat, illetve feltölthetnek bármilyen kiegészítő (kapcsolódó) dokumentumot.

Mindkét portál regisztráció és jelszó nélkül elérhető, tehát bárki letölthet bármilyen anyagot, és fel is tölthet ötleteket. Várjuk a további ötleteket/javaslatokat, nemcsak a képzőintézményekben oktató kollégáktól, hanem a fordítópia képviselőitől is, akik a szakmai gyakorlatok és mentori tevékenységük során számos „irodai” fordítási feladatot adnak a szakfordító-gyakornoknak, s ezek az életszerű feladatok már az egyetemi oktatás során is megjelenhetnek a klasszikus tanfordítások mellett. Érdekes lehet a képzőintézmények számára az is, hogy az irodákban a mentorok, projektmenedzserek, lektorok miként értékelik belső munkatársaikat vagy alvállalkozásban dolgozó fordítóikat napi munkájuk során, azaz milyen konkrét visszajelzéseket adnak. Szívesen látnánk tehát az Értékelési portálban ilyen leírásokat az irodák részéről annak érdekében, hogy az oktatók értékelési eszköztárát ki tudjuk tágítani, és közelíteni tudjuk a piaci gyakorlatban alkalmazott lektorálási szempontokhoz.

Önértékelési eszközök szakfordítóknak

Az eTransFair projektben két önértékelési eszközt is létrehoztunk, amelyet a hallgatók a képzőintézménybe történő bemenetkor, illetve a diplomaszerezést követően tölthetnek ki. Az első kérdőívet az eTransFair szakfordítói kompetenciamodelljének felhasználásával készítettük.⁴ A kérdőív arra keres választ, hogy bemenetkor a hallgatók saját értékelésük alapján milyen kompetenciákkal rendelkeznek a nyolc kompetenciaterület mindegyikén (<https://hu.etransfair.eu/rolunk/szellemi/szakforditoi-kompetenciaprofil>); majd a végzős hallgatók kimenetkor felmérhetik, hogy mennyit fejlődtek a képzés során az adott kompetenciaterületeken, ezeken belül pedig a részelemekre vonatkozó tudás, készségek, attitűd és felelősségvállalás tekintetében.

Kompetencia ¹	Tanulási eredmények/deszriptorok	
1. FORDÍTÁSI KOMPETENCIA		
Tudás ²	A folyamatokra fókuszálva (procedurális készségek elsajátítása) számos fordítási modellt és módszert átlát.	1 2 3 4 5
	Ismeri a fordítási stratégiákat és technikákat.	1 2 3 4 5
Készségek ³	Megérti a forrásnyelvi szövegeket, és az ügyfél igényeinek, valamint a szkoposznak megfelelő fordítást tud létrehozni.	1 2 3 4 5
	Képes felismerni és megoldani fordítási problémákat.	1 2 3 4 5
	Felismeri és helyesen alkalmazza a fordítási stratégiákat és technikákat, megfelelő metanyelv használatával.	1 2 3 4 5
2. NYELVI KOMPETENCIA		
Tudás	Tisztában van a forrásnyelv és a célnyelv nyelvi és (inter)textuális konvencióival	1 2 3 4 5
	Tudja, hogyan kell helyesen alkalmazni ezeket a konvenciókat a forrásnyelven és a célnyelven egyaránt.	1 2 3 4 5

3. ábra. Szakfordítói kompetenciák önértékelése az eTransFair kompetenciaprofilja alapján (mintaoldal)

Ugyanezt a mérőeszközt elkészítettük az eTransFair projekt során összeállított puha készségek (*soft skills*) listájára vonatkozóan is⁵ – hallgatóink ezt a kérdőívet is kitöltik bemenetkor és kimenetkor egyaránt, és ezzel is képet kapnak saját fejlődési ívükről, erősségeikről és esetleges hiányosságaikról.

⁴A teljes szakfordítói kompetenciaprofil elérhető itt: <https://hu.etransfair.eu/rolunk/szellemi/szakforditoi-kompetenciaprofil>

⁵ Mindkét önértékelési mérőeszköz megtalálható az Értékelési portálon

“PUHA KÉSZSÉGEK”	
1. elemzési képesség	1 2 3 4 5
2. problémamegoldás	1 2 3 4 5
3. döntéshozatal	1 2 3 4 5
4. empátia	1 2 3 4 5
5. tolerancia	1 2 3 4 5
6. etikai megfontolások	1 2 3 4 5
7. rugalmasság	1 2 3 4 5
8. kreativitás	1 2 3 4 5
9. önbizalom	1 2 3 4 5
10. gyors reagálás /reaktív/proaktív készségek	1 2 3 4 5
11. megfelelő önismeret	1 2 3 4 5
12. megbízhatóság	1 2 3 4 5
13. instrukciók megértése, követése	1 2 3 4 5
14. stressztűrő képesség	1 2 3 4 5
15. kitartás, állóképesség	1 2 3 4 5

4. ábra. Puha készségek önértékelése az eTransFair lista alapján (mintaoldal)

A fordító- és tolmácsoló intézmények munkájuk során számos belső oktató kollégával, külsős óraadóval és a szakmai gyakorlat révén sok piaci szereplővel kerülnek kapcsolatba, akik valamilyen módon részt vesznek a képzésben, ugyanakkor eltérő háttérismeretekkel, pedagógiai ismeretekkel és gyakorlattal, piaci tapasztalattal és habitussal rendelkeznek. A projekt segítségével két olyan klasszikus értelemben vett tudásmegosztási felületet sikerült létrehozni, ahol a fordítóképzésben érdekelt felek, elsősorban az oktatók és a piaci képviselők megoszthatják egymással gyakorlati tapasztalataikat. Várjuk tehát az ötleteket, feladatokat, javaslatokat, illetve személyes tapasztalatokat a módszertani és az értékelési portálra egyaránt. Arra buzdítjuk felhasználóinkat – mind az egyetemi, mind pedig a piaci kontextusban dolgozó szakembereket –, hogy keressék fel az eTransFair portáljait, és próbáljanak ki minél több érdekes, számukra új technikát, és bővítsék értékelési, módszertani kellékárukat.

Szabó Csilla

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 illusztrálendő, 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

Beszámoló a XVII. Pedagógiai Értékelési Konferenciáról (PÉK 2019) Szeged, 2019. április 11-13.

A Szegedi Tudományegyetem Neveléstudományi Intézete és Neveléstudományi Doktori Iskolája, valamint az MTA–SZTE Képességfejlődés Kutatócsoport szervezésével 2019. április 11–13. között került megrendezésre a XVII. Pedagógiai Értékelési Konferencia (PÉK). A konferencia 2003 óta évente meghívja a pedagógiai értékelés területének kutatóit s a vonatkozó tudományos eredmények iránt érdeklődőket. Már a kezdetektől jellemző az a szándék, hogy a konferencia országos és nemzetközi szinten is teret adjon a tudományos diskurzusnak. A rendezvény rendszerint szegedi szervezésű, a plenáris előadásokon nemzetközileg elismert, gyakran külföldről érkező kutatók szólalnak fel. 2012 óta a konferencia kétnyelvű, a tudományos eredmények bemutatása a magyar mellett angol nyelven történik.

A megnyitó angolul hangzott el, ahol először a konferencia elnöke, Molnár Edit Katalin köszöntötte a résztvevőket, majd átadta a szót Kónya Zoltánnak (az SZTE rektorhelyettesének), Gécseg Zsuzsannának (az SZTE oktatási dékánhelyettesének), majd Csapó Benőnek (az SZTE Neveléstudományi Doktori Iskola vezetőjének, a konferencia alapítójának). A köszöntőn középpontba került a konferencia kétnyelvűségének ténye, ugyanis ez alkalommal az elfogadott idegen nyelvű előadások száma meghaladta a magyar nyelvűekét.

Számos külföldi vendég érkezett például Finnországból, akik közül Patrik Scheinin, a Helsinki Egyetem Magatartástudományi Karának vezetője tartotta meg az első plenáris előadást a finn érettségi rendszer változásainak tanulságairól (*Reforming the Matriculation Examination in Finland: Problems, Prospects and Conclusions for Europe*). Az előadás kérdéseket vetett fel, s a hazai kollégákat közös gondolkodásra hívta. Mi az a műveltségi szint, tudásszint, amelyet érdemes és reális elvárni a tanulóktól? Mely tantárgyaknak volna igazán fontos megjelenni az érettségi vizsgán? Összesen mennyi tantárgyból lenne valóban szükséges vizsgát tenni? Milyen típusú teszt lehet alkalmas a megbízható és valid vizsgáztatáshoz? Az egyértelmű válaszok helyett az elmúlt évek finn megoldási irányainak bemutatása történt, amelyek jelentősen eltérnek a hazánkban megszokottól. A finn érettségizők mozgástere nagyobb: a vizsgákon szerzett érdemjegyek és a vizsgára megválasztott érettségi tárgyak hányadosából egy összesített pontszám

képezhető, amely a felsőfokú továbbtanulás alapja. Finnország digitalizált érettségit vezetett be, ezzel számos lehetőséget (így például az adaptív tesztelést) megteremtve. Kihívást jelent ugyanakkor továbbra is például az automatizált pontozás szakszerű megoldása, különösképp a nyílt végű kérdések esetén.

A konferencia középső egységében a Springer gondozásában 2019-ben megjelent *International Handbook of Mathematical Learning Difficulties: From the Laboratory to the Classroom* című könyv bemutatójára került sor, mely kötetnek világszerte a matematikatanulással, illetve a tanulási nehézségekkel foglalkozó, összesen 99 kutató a szerzője. E program keretein belül felszólaltak a szerkesztők: Annemarie Fritz, Vitor G. Haase és Pekka Räsänen; továbbá a szerzők közül Rausch Attila és Molnár Gyöngyvér. A 843 oldalas könyv 46 fejezetből és öt részből áll. A könyv szerkezetével összhangban e hat tágabb témakör főbb mondanivalóit ismertette a könyvbemutató. Az alábbiakról volt szó: 1. a számfogalom megértésének fejlődése különböző nézőpontok alapján, 2. a matematika tanulása és annak nehézségei a világban, 3. a matematikatanulás nehézségeinek kognitív, motivációs és emocionális alapjai, 4. az aritmetikai készségek fejlődésének nehézségei, valamint 5. a matematika tanulási nehézség felismerése és intervenciós lehetőségei. A könyv különbséget tesz a matematika tanulási nehézség és diszkalkulia között, rámutatva arra a világszerte jellemző paradoxonra, mely szerint a jó matematika teljesítménnyel rendelkező országokban alacsony a motiváció, míg az alacsony teljesítményű országokban magas.

A második plenáris előadás során Halász Gábor, az ELTE Pedagógiai és Pszichológiai Kar egyetemi tanára, a Felsőoktatás- és Innovációkutatás Csoport vezetője, az oktatási innovációk mérési modelljét mutatta be szintén angolul (*Measuring innovation in education: Theoretical framework and survey outcomes*). Az Innova kutatás esettanulmányokon és kérdőíves vizsgálaton alapuló, kevert (kvalitatív és kvantitatív) módszertanú vizsgálatsorozat, mely a mikroszintű (intézményi) és a makroszintű (rendszerszintű) innovációk, változások közötti összefüggéseket vizsgálja. Az előadás metaforákkal érzékeltette az egyes lehetséges helyzeteket, például üres vasdoboz (alacsony innovációs aktivitás mindkét szinten), gyöngy a vasdobozban (alacsony innovációs aktivitás makroszinten, de magas mikroszinten). A bemutatott kategóriákkal a *top-down* mellett a *bottom-up* változások is kifejezhetők, ennek vizsgálata pedig ugyancsak célja az elméleti kereteivel ismertetett projektnek.

A háromnapos konferencián kilenc angol nyelvű és kilenc magyar nyelvű tematikus szekció, valamint két angol nyelvű szimpózium is gazdagította a programot. A rendszerint két-két párhuzamos szekcióban, összesen 45 angol és 36 magyar nyelvű előadás hangzott el az alábbi témákban (a teljesség igénye nélkül): a tanulás szociális és kognitív aspektusai, tanulási motiváció, anyanyelvi és idegen nyelvi kutatások, nyelvtanítás, a 21. századi készségek, pedagógiai diagnosztika, társadalomtudományi műveltség, zenei nevelés, egészség és jóllét az iskolában, oktatási rendszer, tantárgypedagógia, a tanárok

szakmai háttere, szelf és közösség működése. A konferenciát lezáró beszédében Molnár Edit Katalin Dancs Katinkával (a konferencia titkárával) köszönetet mondtak a konferencia megvalósulását lehetővé tevő valamennyi kollégának, a konferencia résztvevőinek, majd jelképesen átadták a szervezői feladatokat a konferencia következő évi elnökének. A 2019-ben publikált absztraktkötet az alábbi linken elérhető: http://edu.u-szeged.hu/pek2019/download/CEA_2019_PROGRAMME_ABSTRACTS.pdf

Kiss-Kovács Renáta

Kiegészítés

A Modern Nyelvoktatás 2019/2. számában P. Márkus Katalin *Szótárhasználati munkafüzetek az oktatásban – használatban a Tanulószótárak* című tanulmányának 47. oldalán a szerkesztés folyamán az alábbi rész mellől

„A makrostruktúra a szótár címszavainak valamilyen szempont alapján sorba rendezett együttese. A rendezés különböző szempontok alapján történhet: betűrend szerint, tematikusan, időrendben, gyakoriság szerint stb.”

kimaratdak az alábbi hivatkozások (Hartmann 1984: 24, Hartmann–James 1998, Fóris 2002: 18).

IRODALOM

- Hartmann, Reinhard Rudolf Karl (1984, szerk.): *Workbook on lexicography*. Exeter Linguistic Studies. 8. Exeter: University of Exeter.
- Hartmann, Reinhard Rudolf Karl – James, Gregory (1998): *Dictionary of lexicography*. London: Routledge
- Fóris Ágota (2002): *Szótár és oktatás*. Budapest: Iskolakultúra.