

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 vizsgaszempontjainak 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 nyelvvizsgák 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.

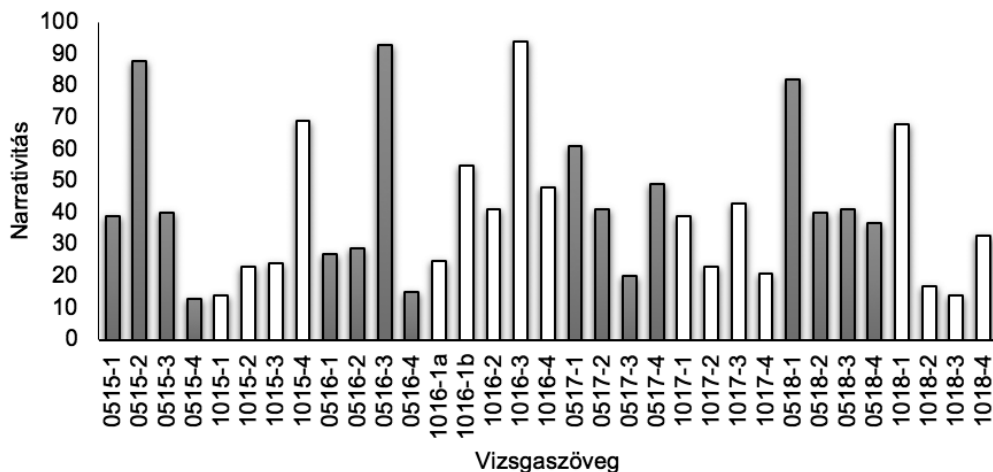
Az elemzés menete a következő volt. Először a feladatokban használt szövegeket elemezhető formátumúvá kellett alakítani. Erre azért volt szükség, mert az érettségi különböző feladattípusai között olyanok is szerepelnek, amelyek esetében a szöveg egyes részei hiányoznak, és a vizsgázók feladata a megfelelő 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 normáeloszlás 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

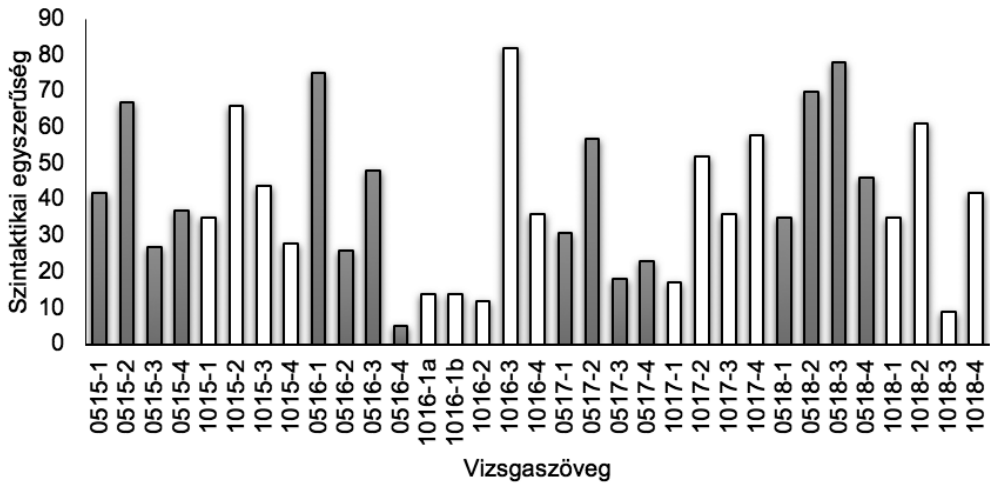
A következőkben áttekintjük a vizsgálat eredményeit. Elsőként vegyük szemügyre a *narrativitás* kapcsán született Coh-Matrix TERA eredményeket, melyeket az 1. ábra mutat be.



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.

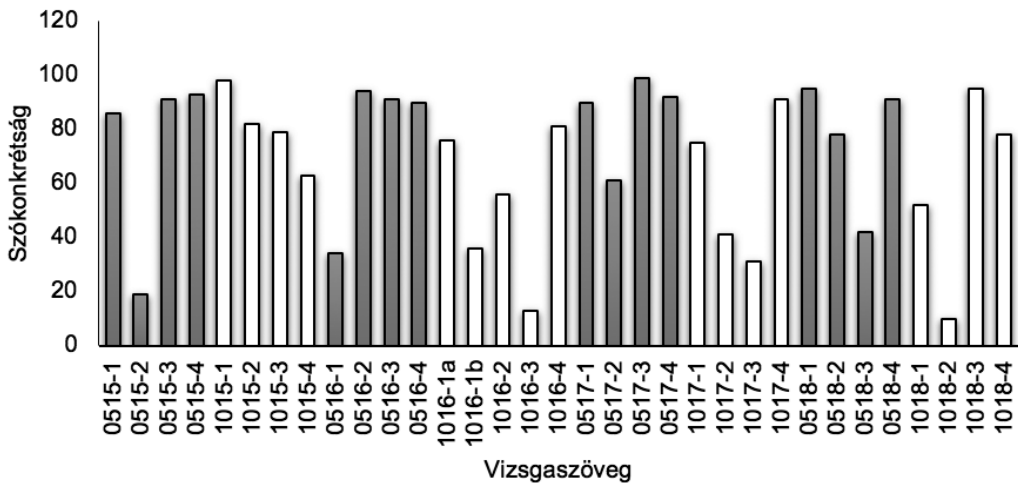


2. ábra. Coh-Matrix TERA eredmények: szintaktikai egyszerűség

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. Érdekes 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



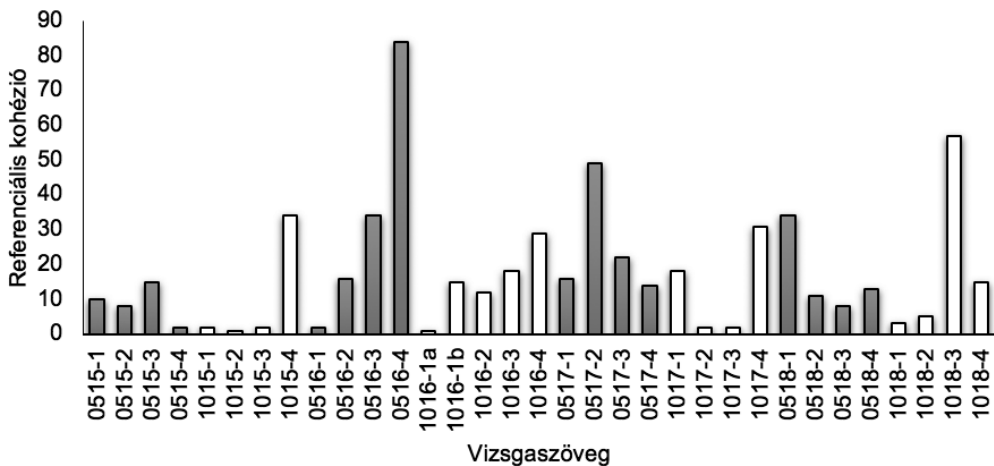
szerepeltek. A szavak konkrétsága általában könnyebbé teszi a szövegek érthetőségét,

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

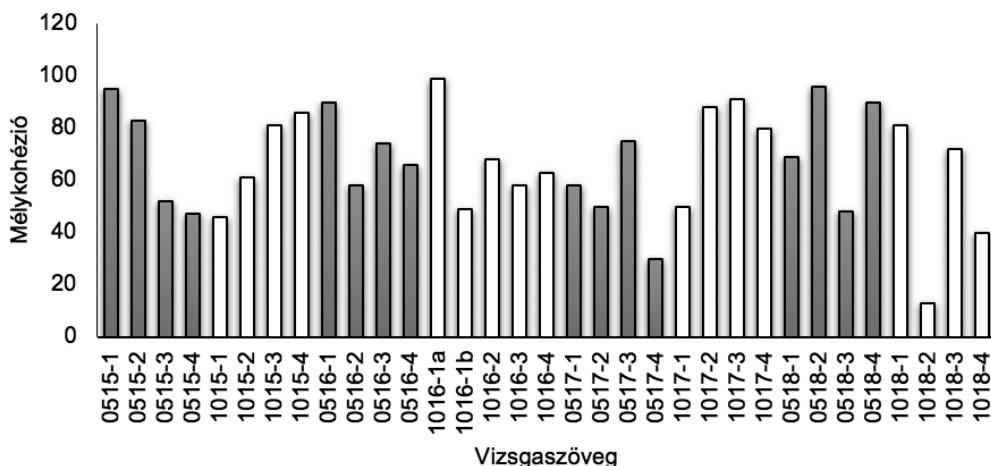


4. ábra. Coh-Metrix TERA eredmények: referenciális kohézió

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

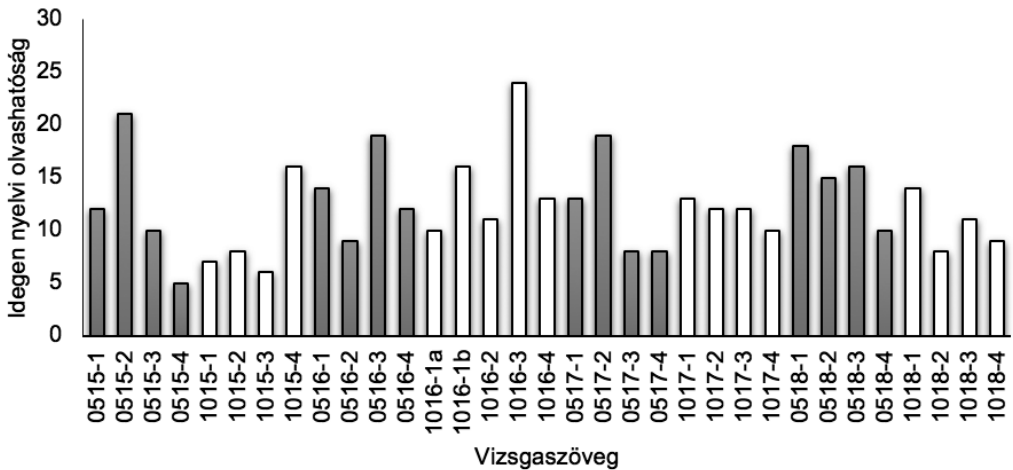


5. ábra. Coh-Matrix TERA eredmények: mélykohézió

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 szókincssel 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 kiegyenlítődték.

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 érettségien 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.)