

Angol nyelvű olvashatósági formulák magyar nyelvi adaptálásának lehetséges irányai

A magyarországi olvasáskutatás bár impozáns múltra tekint vissza, módszertani eszköztárának, a kutatás területeinek bővítése az olvasástanítás sikerességének érdekében elengedhetetlen. Azért, hogy az olvasás sikerességét objektíven meghatározhassuk, az olvasóközpontú megközelítések mellett a szöveg alkalmasságának objektív mércéjére is szükség van. Ehhez hiánypótló megoldást nyújthatnak az olvashatósági formulák, melyek magyar nyelvű adaptációját ezen tanulmány hivatott előkészíteni.

Bevezetés

Az olvashatóságformula-vizsgálatok segítségével egy adott szöveg olvasásának sikerességét a szöveg nyelvi jellemzőiből adódóan matematikai úton számolhatjuk ki és jelezhetjük előre, így a tankönyvi szövegek önálló tanulásra történő felhasználhatóságát is pontosabban határozhatjuk meg. A szövegek nyelvi adottságaiból eredő értéket a szöveg olvashatóságának nevezzük. Az olvashatóságot tehát az oktatás szempontjából kiemelten fontos kutatási területnek tekinthetjük, mivel az közvetlenül befolyásolja, hogy a diákok mennyire könnyen tudják megérteni az írott szövegeket, és mennyire hatékonyan tudnak önállóan tanulni azokból. Ha az oktatási folyamatban alkalmazott anyagok nehezen olvashatóak, a tanulók egyrészt nehezen értik meg azok tartalmát, ami további problémákhoz, frusztrációhoz, motivációvesztéshez és lemaradás-hoz vezethet. A könnyen olvasható, az információk tartalma és strukturálása, nyelvi szerkesztettség, elrendezés szerint a megértést, befogadást megkönnyítő szövegeket ezzel szemben könnyebben befogadják, jobban megértik ezek tartalmát, ami hozzájárulhat a tanulmányi eredményeik, teljesítményük javulásához (Kaposi, 2012).

A tanítás során elsajátítandó ismeretanyag jelentős része a tankönyvek szövegein keresztül jut el a tanulóhoz, így ezen szövegek minősége a tanulás minőségének egyértelmű befolyásoló tényezője. A tankönyvek szerepe az oktatásban a tanárok és a diákok szempontjából egyaránt megingathatatlanak tűnik. A tankönyvek egyfajta uniformizált tanítást tesznek lehetővé, jelentősen csökkentve a tanárok oktatásszervezési feladatát. Ebből adódóan a leckéről leckére történő haladás módszerének gyakori alkalmazását figyelhetjük meg, bizonyos esetekben még akkor is, ha a tankönyvi ütemezés az adott helyi tantervvel nem teljesen összeegyeztethető (M. Nádasi, 2011). A tankönyv azonban nem csupán az oktatásszervezés szempontjából kulcsfontosságú, hanem a tanulás és tanítás folyamatában is központi szerepet játszik. A pedagógusok többsége a tankönyveket az önálló otthoni tanulás fontos eszközének tartja. A tanári magyarázat és az ebből származó tanulói jegyzet célja nem a tankönyvi leckék helyettesítése, hanem azok kiegészítése, ebből adódóan a számonkérések előtti készülés szerves részét képezi a tankönyvi lecke

önálló értelmezése is. Mindezek ellenére a tanulók rendszeres otthoni tankönyvhasználatának mértéke a várthoz képest alacsony (Szemerszki 2018). Ezen jelenség magyarázatát részben a számonkérési szokások jellegében találhatjuk, de kétségkívül figyelembe kell vennünk azt a lehetőséget is, mely szerint a tankönyvi szövegek az önálló tanulásra nem, vagy csak részben alkalmasak, ugyanis szövegeik nem illeszkednek a tanulók olvasási igényeihez és esetenként szövegértési képességeik szintjéhez. A szövegek potenciális inadekvátságának kérdése különösen fontossá válik a tanulási zavarokkal vagy nyelvi akadályokkal küzdő diákok esetében, mivel nekik további támogatásra lehet szükségük az írott anyagok feldolgozása során. Az olvasható anyagok használatával a pedagógusok hozzájárulhatnak ahhoz, hogy minden tanuló egyenlően hozzáférjen az oktatási forrásokhoz és lehetőségekhez.

Az olvasás sikerességének szöveg felől történő megközelítése, tankönyvi szövegek olvashatóságának kutatása az Amerikai Egyesült Államokban már az 1920-as évektől kezdve folyamatos. A közel százéves múlttal rendelkező olvashatósági formulák sikerességéről és népszerűségéről tanúskodik, hogy folyamatosan fejlődő módszertanuknak és növekvő számuknak köszönhetően az 1980-as évekre a legkülönbözőbb területeken: oktatási, orvosi, jogi, marketing- és szórakoztató célú szövegek vizsgálatára már több mint kétszáz formula volt használatban az Amerikai Egyesült Államokban (Dubay, 2004). Bár használatuk százéves történelmük jelentős részében az angol nyelvterületekre lokalizálódott, a 20. század végéhez közeledve a formulák egyre népszerűbbé váltak Európában és Ázsiában is. Magyarországon azonban a formulákkal kapcsolatos diskurzus, mint már utaltunk rá, még nem bontakozott ki, ezért jelen írás célja az olvashatóság alapvető elméleti keretének és lehetséges adaptációs irányainak felvázolása.

Az olvashatóság fogalma

Mielőtt az olvasható szöveg attribútumait meghatározhatnánk, szükséges megkülönböztetnünk egymástól két fogalmat: az olvashatóságot (*readability*) és a kiolvashatóságot (*legibility*). Bár mindkét kifejezés egy adott szöveg olvasásával kapcsolatos nehézségek objektív meghatározását szolgálja, módszertanuk merőben eltérő attribútumok és számítási módszerek köré épül fel.

Utóbbi fogalmunk, a kiolvashatóság az olvasási nehézség kérdését alapvetően a szöveg mechanikus (papírminőség, háttérszín, lapméret stb.) és tipográfiai (betűméret, sor- és bekezdésköz mérete, betűtípus stb.) attribútumai felől közelíti meg, módszertana pedig főként az olvasó és a szöveg közötti távolság vagy az olvasás sebességének mérésére összpontosít (Burt, 1949). Az olvashatóság ezzel szemben bármely szöveg olvasásának nehézségét lexikai és szintaktikai adottságait alapul véve matematikai számítások segítségével fejezi ki (Dubay, 2004). Ezen látszólag egyszerű megközelítés azon a feltételezésen alapul, hogy a hosszabb, több szótagból álló szavak kiolvashatósága több időt és nagyobb erőfeszítést igényel, a több szóból álló, hosszabb mondatok pedig több információt tartalmaznak, így értelmezésük nehezebb, mint a rövidebb tömondatoké (Klare, 1974). Bár mindkét megközelítés potenciális segítséget nyújthat a tankönyvi szövegek adekvátságának pontosabb megismerésében is, a Covid19-világjárvány idején előtérbe kerülő és azóta is folyamatosan fejlesztett digitális taneszközök és digitalizált tankönyvek térnyeréséből adódóan az olvashatósági vizsgálatok fejlesztése a kiolvashatóságával szemben előnyösebb törekvésnek bizonyulhat korunkban is.

Természetesen az olvashatósági vizsgálatok önmagukban nem elegendőek ahhoz, hogy a tankönyvi szöveg olvasásának sikerességét teljes pontossággal megjósolják, és így a magyarországi olvasáskutatások által kedvelt módszereket sem hivatott felváltani, csupán kiegészíteni. Mivel az olvasás rendkívül összetett folyamat, sikerességéről teljes és pontos képet csupán minden változó összjátékának teljeskörű ismeretében kaphatunk.

Ezen változókat Dale és Chall (1949) három csoportba osztva, azok kölcsönhatásai alapján állapítja meg. Ezek közül az első maga a szöveg, annak tartalma és szerkezete; a második az olvasó, annak olvasásértési képessége, olvasási motiváltsága, tudása; a harmadik pedig maga a mérni kívánt paraméter (motiváció, olvasási sebesség, szövegértés stb.) és a mérés mikéntje.

A fent említett csoportok közül az elsöre, azon belül is a szöveg szintaktikai és lexikai tulajdonságainak olvasási nehézség szempontjából történő vizsgálatára magyarországi és magyar nyelvű példát nehéz volna találni, ezen hiátus betöltésére az olvashatósági indexek magyar nyelvű adaptálása nyújthat megoldást.

Olvashatósági kutatások az Egyesült Államokban

Az olvashatósági formulák egyszerű működési elveik ellenére – vagy épp azoknak köszönhetően – százéves múltjuk során újra és újra bizonyították megbízhatóságukat és rugalmasságukat, melynek köszönhetően mára számuk meghaladja a kétszázat. Bár a formulák alapvetően azonos elven alapulnak, a számítási módok között jelentős eltéréseket figyelhetünk meg. Az elérhető olvashatósági formulák megdöbbentően magas száma az Amerikai Egyesült Államok olvashatósági kutatásait övező versenynek és piaci vetélkedésnek köszönhető.

A 20. század első felének formulái mögött álló kutatásokról csupán keveset tudunk, mivel ezek kvázi üzleti titokként rejtve maradtak, elsőként az '50-es években felbukkanó formulák kapcsán nyerünk részletes betekintést az azok mögött rejlő kutatásokba. Az '50-es évektől kezdődően robbanásszerűen megnövekedik az olvashatósággal kapcsolatos kutatások – és velük együtt a formulák – száma. A cél nem csupán az lesz, hogy az olvasási nehézség szintjét lehető legpontosabban mérő formulákat hozzanak létre, hanem a formulák egyszerű használhatósága is fő szemponttá válik.

A leggyakrabban használt olvashatósági formulák

Az elérhető angol nyelvű formulák magas száma ellenére csupán néhány formulára lehet azt mondani, hogy a század végén számos különböző szektor széles körben egyaránt alkalmazza őket (Tekfi, 1987).

1. Flesch Reading Ease (Flesch, 1948)

A Flesch Reading Ease formulát az 1940-es években Rudolf Flesch fejlesztette ki azzal a céllal, hogy könnyebben értelmezhető újságcikkeket tegyen elérhetővé az amerikai olvasók számára. A kétváltozós formula egyszerűségének és megbízhatóságának köszönhetően viszont mára számos különböző szektor (oktatás, egészségügy, marketing stb.) eszköztárának részét képezi.

$$206,835 - 1,015 \left(\frac{\sum \text{szó}}{\sum \text{mondat}} \right) - 84,6 \left(\frac{\sum \text{szótag}}{\sum \text{szó}} \right) = \text{olvashatósági érték}$$

Ez a formula egy nulla és száz közötti értéket ad, ahol a nulla a legbonyolultabb szöveget jelenti, és az Egyesült Államok felnőtt lakosságának átlagos iskolázottsági szintjét alapul véve az angol nyelv általános érthetőségi szintjének a körülbelüli 70-es pontértéket határozta meg. A skálán mérhető lehető legbonyolultabb nullás értékű szöveg átlagos szótagszáma meghaladja a kettőt, és az átlagos mondat 37 szóból áll.

2. Gunning Fog Index (Gunning, 1969)

A formula első változatát Robert Gunning az 1940-es években kezdte el kifejleszteni, széles körben viszont csak a formula második változata terjedt el az 1960-as években. Mára pedig a formulát leggyakrabban üzleti írással kapcsolatos tankönyvek vizsgálatára használják.

$$0,4 \left[\left(\frac{\Sigma \text{ szó}}{\Sigma \text{ mondat}} \right) + 100 \left(\frac{\Sigma \text{ komplex szó}}{\Sigma \text{ szó}} \right) \right] = \text{"fog" olvashatósági érték}$$

A Gunning által „fog”-nak nevezett érték, amiről a formula a nevét is kapta, az átlagos mondathosszság és a három vagy több szótagból álló „komplex” szavak százalékos gyakoriságából számolható ki, és eredménye egy olyan 0 és 20 közé eső szám, ami azt mutatja, hogy hány év befejezett iskolai tanulmány szükséges ahhoz, hogy az olvasó a szöveget egy végigolvasás után is képes legyen értelmezni.

3. SMOG Grading (McLaughlin, 1969)

G. Harry McLaughlin formulájának megalkotása mögött az egyszerű használatra való mindenekfeletti törekvés búvik meg, és ezen törekvés eleinte számos kritika célpontjává is tette, viszont megbízhatóságát bizonyítja, hogy ma, több mint 50 évvel megalkotását követően még mindig a leggyakrabban használt formulák közé tartozik.

$$3 + \sqrt{\frac{\Sigma \text{ komplex szó}}{30 \text{ mondat}}} = \text{"smog" olvashatósági érték}$$

A teszt a FOG indexhez hasonlóan a három vagy több szótagból álló „komplex” szavak gyakoriságát nézi, viszont ezen szavaknak nem a százalékos előfordulását veszi alapul, hanem a harminc mondatból álló szövegrészletekben lévő „komplex” szavak számát

Bár az olvashatósági formulák adaptáció nélkül történő nem angol nyelvi használata egy esetben sem mondható egyértelműen tanácsosnak, a fent tárgyalt három formula a magyar nyelv esetében különösen problematikusnak bizonyul. Mind a három formula a szöveg szótagokban számolt szóhosszúságát veszi alapul, ami magyar szövegek vizsgálata esetén minden esetben a szöveg drasztikusan magas olvashatósági értékét eredményezi. Egy gyermekirodalmi szövegeket vizsgáló kutatás esetében – ahol a vizsgált szövegeket azok szerzője egyenként óvodakezdő, óvodás, iskolakezdő és felső tagozatos közönségnek szánt – a formulák mindegyike a szövegeket középiskolát végző, egyetemista vagy egyetemet végző olvasók számára ítélte ideálisnak. Ennek egyik fő oka a szövegek átlagos szóhosszúságából ered, amely szótagokban mérve 2,14, 2,19, 2,24, 2,28 volt (Bogdán, 2022). A szótagokban mért szóhosszúság azonban nem csak a nem-angol nyelvi használat során bizonyul problematikusnak, hanem az angol nyelven belül is, mivel adott szavak, regionális kiejtésbéli különbségek miatt, több lehetséges szótagszámmal is rendelkezhetnek, a vizsgáló személyétől függően.

4. Automated Readability Index (Smith és Senter, 1967)

Ezen problémát kiküszöbölendő a 20. század második felében több olyan formulát is létrehoztak, amelyek a szóhosszúságot a szótagszám helyett a karakterek számában állapítják meg. Ilyen formula volt például az Automated Readability Index (ARI) is, amit az Egyesült Államok hadserege számára hoztak létre a gépirás valós idejű olvashatósági vizsgálataira.

$$4,71 \left(\frac{\Sigma \text{ karakter}}{\Sigma \text{ szó}} \right) + 0,5 \left(\frac{\Sigma \text{ szó}}{\Sigma \text{ mondat}} \right) - 21,43 = \text{olvashatósági érték}$$

A formula sikerét gyors és egyszerű alkalmazhatósága eredményezte, melyet eleinte egy külső számlálóegység tett lehetővé, amit az írógépekhez erősítettek. A karakterszámon alapuló formulák a digitális olvashatósági vizsgálatok esetében is előnyösebbnek bizonyultak a szótagszámon alapuló formulákkal szemben, mivel az automatikus, gépi vizsgálatok nehezen tudják számításba venni a különféle szótagolási variációkat.

5. Dale–Chall (1948)

Az átlagos szóhosszúság azonban egyáltalán nem kötelező eleme az olvashatóság kiszámításának. Egyes formulák esetében a hangsúly ugyanis a szavak familiaritására helyeződik.

$$0,1579 \left(100 \frac{\Sigma \text{ nehéz szó}}{\Sigma \text{ szó}} \right) + 0,0496 \left(\frac{\Sigma \text{ szó}}{\Sigma \text{ mondat}} \right) = \text{olvashatósági érték}$$

Az első ilyen olvashatósági vizsgálatok Thorndike (1921) nevéhez kötődnek, aki 41 különböző forrásból összegyűjtött 4 565 000 szó alapján határozta meg az angol nyelv 10 000 leggyakrabban használt szavát. Ezen metódus legnépszerűbb alkalmazója azonban az 1948-ban megjelent és folyamatosan korszerűsített Dale–Chall-formula volt, ami a Dale–Chall „nehéz” szavak listáján alapul. A lista célja azon szavak összegyűjtése, amiket az átlagos amerikai 4. osztályos tanulók legalább 80%-a ismer. A formula népszerűségét magas szintű megbízhatóságának köszönheti, mivel eredményei Bormuth-féle hiányösszöveg-vizsgálatokkal $r = 0,92$ -es korrelációt is mutatnak, ami a legerősebben validált formulák közé sorolja (Dubay, 2004).

A formula azonban a többi eddig tárgyalt módszerhez hasonlóan csupán az angol nyelvű szövegek vizsgálatára alkalmazható, a formula adaptálása pedig előfeltételként igényli a szógyakorisági lista adaptálását is. Annak ellenére, hogy magyar nyelvre vonatkozó szógyakorisági listákból több is létezik, ezek használata specifikusan tankönyvi szövegek esetében nem megbízható, mivel ahhoz nem az átlagos magyar ember, hanem célzottan a tanulók által ismert és használt szavak listájára volna szükség.

Olvashatósági formulák az oktatásban

Az olvashatósági formulák legkézzelfoghatóbb felhasználási lehetősége a tankönyvi szövegek elemzésében rejlik. A tanterem az olvashatósági kutatások bölcsőjének tekinthető, hiszen 1923-ban már az első olvashatósági formula megalkotását is az elérhető fizika tankönyvek szövegeinek túlzott komplexitása inspirálta (Lively és Pressey, 1923). A folyamatosan fejlődő formulák pedig azóta is újra és újra rámutatnak az Amerikai Egyesült Államokban használt oktatási szövegek és a tanulók olvasási szintje között rejülő eltérésekre.

Aukerman (1965) felismerte azt a megdöbbentő tény, mely szerint az Egyesült Államok 7–12. osztályos tanulóinak jelentős része nem tudja megfelelő mértékben értelmezni az irodalmi szöveggyűjteményekben szereplő szövegeket. Bár ezen eredmény okait nem lehet egyetlen tényezőre szűkíteni, és a tanulmányban felsorolt lehetséges okok közül nem mindegyik állja ki az idő próbáját, az olvashatósági vizsgálatok segítségével kinyert számadatok arra mutatnak, hogy az összeválogatott szövegek komplexitása és a diákok képességei között lévő távolság lehet az egyik fő oka a sikertelen megértésnek.

Egy további kutatásban Kahn és Pannbacker (2000) harminc, szájpadasadékkal élő gyermeknek és szüleiknek szóló oktatási szöveg SMOG és Fry olvashatósági vizsgálatával mutatott rá arra, hogy nem csupán az olvasandó szöveg és az olvasó szintje közötti eltérések okozhatnak problémát, hanem a rendelkezésre álló különböző oktatási célzatú szövegek közötti nyelvi komplexitásbeli eltérések, valamint ezen eltérések rendszerezetlensége is megnehezíti azt, hogy az olvasó konzisztensen hozzájuthasson a számára megfelelő szövegekhez.

Magyar nyelvű szövegeken is alkalmazható olvashatósági formulák birtokában, hasonló kutatások segítségével lehetőségünk nyílna arra, hogy objektíven felmérjük a magyarországi tanulók olvasási képességei és a számukra olvasásra, megértésre szánt szövegek komplexitása közti távolság mértékét, jellemzőit. Ezt követően pedig az oktatásra szánt szövegek potenciális olvasási nehézségeire rámutató előrejelzéseket tudunk majd tenni, melyek a szakértői véleményezést és más jelenleg is alkalmazott módszereket kiegészítve biztosíthatják azt, hogy a tanulók számukra megfelelő szövegekből tanulhassanak, így hosszú távon javulhasson olvasásértési teljesítményük és iskolai sikerességük.

A formulák rendszerezése

A formulák magyar nyelvű adaptációja vagy önálló magyar nyelvű formula készítése során elengedhetetlen lépés lesz a már elérhető formulák között megfigyelhető trendek összehasonlítása. Mivel különböző szektorok különböző változókat mérő formulákat részesítenek előnyben, a magyar nyelvű oktatás szempontjából leghatékonyabb módszer feltárása érdekében elengedhetetlen lépés a formulák változóik alapján történő kategorizálása, ami a későbbi összehasonlító vizsgálatokat hivatott megkönnyíteni.

A fenti formulákból, a bennük felhasznált változók jellege alapján, két fő csoportot alakíthatunk ki: az egyik csoport a szavak hosszát veszi számítása alapjául, a másik pedig a szófamiliaritást állítja középpontba. Az első csoportunkat további két részre tudjuk bontani a szóhosszúság kiszámításának módja mentén. Az első alcsoport a szavak hosszát szótagszámban mutatja, a második pedig betűszámban állapítja meg azt. Bár a fenti kategóriákat a legnépszerűbb formulák közül csupán néhány tüzetes szemügyre vétele alapján szemléltettük, a legkönnyebben hozzáférhető és legnépszerűbb formulák mindegyike elhelyezhető a kategóriák valamelyikében (1. táblázat).

1. táblázat. A legnépszerűbb olvashatósági formulák mért változók alapján történő csoportosítása

Szófamiliaritás	Szóhosszúság	
	Szótagszám	Betűszám
Dale-Chall Formula Spache Formula	Flesch Reading Ease Gunning Fog Index SMOG Index Powers Sumner Kearsley Formula Fry Readability Graph Lix Readability Formula FORCAST Readability Formula	ARI Coleman Liau Index Raygor Readability Graph Lix Readability Formula Rix Readability Formula

A táblázat az interneten hozzáférhető és népszerű angol nyelvű olvashatósági indexeket tartalmazza. A táblázatban szereplő formulák kiszámításának módja a következő weboldalon érhető el: <https://readable.com/features/readability-formulas/> Utolsó letöltés: 2023. 05. 24.

A magyar nyelvi adaptálhatóság szempontjából a három csoport egyikét sem tudjuk teljes magabiztossággal véglegesen kizárni, de a jelenleg rendelkezésünkre álló információk alapján megállapítható, hogy legkézenfekvőbb kiindulási pontnak a harmadik kategória formuláinak valamelyike bizonyulhat. Az első kategória, bár az angol nyelvű szövegeken legmagasabb korrelációval bíró formulát is tartalmazza, praktikus okokból, a kifejezetten tanulók által ismert szavak felmérését célzó szófamiliaritási lista hiányában szorul háttérbe. A második kategória pedig annak ellenére, hogy a legnépszerűbb formulákat foglalja magába, az angol és a magyar nyelv közötti átlagos szótagszámok eltérései miatt bizonyul nehezen adaptálhatónak.

Az olvashatóság kritikái

Az olvashatósági vizsgálatok körüli magyarországi diskurzus kialakításának érdekében nem szabad megfejeltkeznünk azok hiányosságairól sem. Az olvashatósági formulákat érő leggyakoribb kritika az, hogy nem mindig tükrözik pontosan egy szöveg komplexitását és nehézségét, mivel figyelmen kívül hagynak olyan emberi tényezőket, mint például az olvasási motiváció. Továbbá a szövegalkotás folyamata alatt sem szabad kizárólagosan az olvashatósági formulákra támaszkodni, mivel az a szöveg túlzott leegyszerűsítéséhez vezethet, ami akár a szöveg megértését is megnehezítheti (Selzer, 1981). Ezeknek a látszólagos hibáknak az oka pedig pont az egyszerűségükben rejlik, melynek tengerentúli népszerűségüket is köszönhetik.

A pusztán matematikai módon történő megközelítés vélt elégtelensége azon félreértésre vezethető vissza, hogy az olvashatósági formulák célja nem a szöveg sikeres értő olvasásának mérése, hanem annak előrejelzése. Az olvasás rendkívül összetett folyamatának előrejelzését pedig akár az időjárás megjóslásához is hasonlíthatnánk, hiszen a lehetséges befolyásoló tényezők és változók pusztán számából adódóan az értő olvasás sikerességét sem tudjuk pontosan előre jelezni, ahogy az esőt sem. Ez azonban nem jelenti azt, hogy nem hasznos tudnunk, mikor érdemes magunkkal esernyőt vinni.

Az olvashatósági formulákat nem a szöveg értelmezhetőségének egyetlen meghatározó aransztenderdjeként vagy a szövegalkotás egyértelmű útmutatójaként hozták létre, hanem egy olyan eszközként, ami a szövegalkotási folyamat során vagy a szövegek korosztályi besorolásának felülbírálasakor szolgálhat támpontként (Bogert, 1985). Egy szöveg átható és teljeskörű elemzése több elemzési módszer együttes használatát igényli, ezért az olvashatósági formula nem a klasszikus mérési módszerek helyett, hanem azok mellett használva értelmezendő.

Összegzés

Az olvashatósági formulák magyar nyelvre történő adaptálása fontos és hatékony lépés lehet a hazai olvasási és szövegértési képességek fejlesztésében. Egy adott szöveg értő olvasásának nehézségét a szöveg olvashatóságának vizsgálatával tudjuk előre jelezni. Ezen előrejelzés az olvashatósági formulák használatával válik lehetségessé, melyek a szöveg lexikai és szintaktikai adottságait alapul véve matematikai úton jelzik előre azt, hogy a szöveg megértéséhez milyen fejlettségű olvasási képességre van szükség (Dubay, 2004). Bár többszáz formula számos különböző számítási módon jut el ehhez az értékhez, a legnépszerűbb formulák által vizsgált változókat alapul véve olyan tendenciákat figyelhetünk meg, ami alapján három csoportra oszthatjuk őket.

A (1) szófamiliaritáson alapuló, a szóhosszúságot (2) szótagszámban és (3) betűszámban mérő formulák közül a magyar nyelvi adaptáció szempontjából legegényesebbnek

a harmadik kategória formulái bizonyultak. Az első kategória praktikus okokból, a kifejezetten tanulók által ismert szavak felmérését célzó szófamiliaritási lista hiányában problematikus, a második kategória pedig az angol és a magyar nyelv közötti átlagos szótagszámok eltérései alapján zárható ki.

Természetesen mindezek mellett hangsúlyozni érdemes, hogy az olvashatóságot az olvasáskutatás egyik lehetséges területeként, eszközöként kell tekintenünk, jelentőségét nem szabad túlértékelni sem. Az viszont kétséget kizárólag megállapítható, hogy az olvashatóság vizsgálata az olvasáskutatás, az olvasástanítás és az iskolai, tanulási célú szövegek, tankönyvek alkotásának szempontjából is ígéretesnek tűnő terület, amelynek fejlesztésével a diákok tanulási folyamatának eredményessége, szövegértési képességeinek javulása prognosztizálható, továbbá mindezek által a tanulási célú olvasási folyamatokról is többet tudunk meg.

Bogdán Patrik

*Pécsi Tudományegyetem
Oktatás és Társadalom Neveléstudományi Doktori Iskola
MTA-PTE Olvasási Fluencia és Szövegértés Kutatócsoport*

Köszönetnyilvánítás, támogatás

A tanulmány elkészítését a Magyar Tudományos Akadémia Közoktatás-fejlesztési Kutatási Programja támogatta. Olvasási Fluencia és Szövegértés Kutatócsoport, MTA-PTE.

Irodalom

- Aukerman, R. C. (1965). Readability of Secondary School Literature Textbooks: A First Report. *The English Journal*, 54(6), 533. DOI: [10.2307/811407](https://doi.org/10.2307/811407)
- Bogdán, P. (2022). A gyermekirodalmi szövegek olvashatósági vizsgálata a szöveg grammatikai komplexitásán keresztül. In Bona, J. & Murányi, S. (szerk.), *Nyelvfejlődés folyamata koragyermekkortől kamaszkorig*. Cser. 149–166.
- Bogert, J. (1985). In Defense of the Fog Index. *Business Communication Quarterly*, 48(2), 9–12. DOI: [10.1177/108056998504800203](https://doi.org/10.1177/108056998504800203)
- Burt, H. E. (1949). Typography and Readability. *Elementary English*, 26(4), 212–221. <https://www.jstor.org/stable/41383630?refreqid=excelsior%3A92d-5349f0d06ed80%2017436f080fc2d05e>
- Dale, E. & Chall, J. S. (1948). A Formula for Predicting Readability. *Educational Research Bulletin*, 27(1), 11–20.
- Dale, E. & Chall, J. S. (1949). The Concept of Readability. *Elementary English*, 26(1), 19–26. https://www.jstor.org/stable/41383594?refreqid=excelsior%3Af4b7a45fcd92ab5f62741127c96ac3c#meta-data_info_tab_contents
- Dubay, W. H. (2004). *The Principles of Readability*. <http://www.impact-information.com>
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221–233. DOI: [10.1037/h0057532](https://doi.org/10.1037/h0057532)
- Gunning, R. (1969). The Fog Index After Twenty Years. *Journal of Business Communication*, 6(2), 3–13. DOI: [10.1177/002194366900600202](https://doi.org/10.1177/002194366900600202)
- Kahn, A. & Pannbacker, M. (2000). Readability of Educational Materials for Clients With Cleft Lip/Palate and Their Families. *American Journal of Speech-Language Pathology*, 9(1), 3–9. DOI: [10.1044/1058-0360.0901.03](https://doi.org/10.1044/1058-0360.0901.03)
- Kaposi, J. (2012). A tankönyv szerepének változásai. *Iskolakultúra*, 22(12), 56–70.
- Klare, G. R. (1974). Assessing Readability. *Reading Research Quarterly*, 10(1), 62. DOI: [10.2307/747086](https://doi.org/10.2307/747086)
- Lively, B. A. & Pressey, S. L. (1923). A Method for Measuring the “Vocabulary Burden” of Textbooks. *Educational Administration and Supervision*, 9(7), 389–398.
- M. Nádasi, M. (2011). A tankönyvek szerepe a tanítás-tanulás folyamatában. *Oktatás-Informatika*, 2(3–4), 26–31. https://www.eltereader.hu/media/2013/05/Okt_Inf_2011_3_4_opt.pdf#page=28
- McLaughlin, G. H. (1969). SMOG Grading—a New Readability Formula. *Journal of Reading*, 12(8), 639–646. <https://www.jstor.org/stable/40011226>

- Selzer, J. (1981). Readability Is a Four-Letter Word. *Journal of Business Communication*, 18(4), 23–34. DOI: [10.1177/002194368101800403](https://doi.org/10.1177/002194368101800403)
- Smith, E. A. & Senter, R. J. (1967). *Automated Readability Index*. Wright-Patterson Air Force Base.
- Szemerszki, M. (2018). A tankönyvek szerepe a tanítási-tanulási folyamatban. In Fehérvári, A. & Széll, K. (szerk.), *Új kutatások a neveléstudományokban* / 2018 Kutatási sokszínűség, oktatási gyakorlat és együttműködések. L'Harmattan Kiadó. 98–115.
- Tekfi, C. (1987). Readability formulas: An overview. *Journal of Documentation*, 43(3), 261–273. DOI: [10.1108/eb026811](https://doi.org/10.1108/eb026811)
- Thorndike, E. L. (1921). *Teacher's Word Book*. Bureau of Publications, Teachers College, Columbia University. https://pure.mpg.de/rest/items/item_2395369/component/file_2395368/content

Absztrakt

Az olvashatóság olyan szövegjellemző, amely előrejelzi a szöveg olvasásának nehézségi fokát, megmutatva, milyen nehezen lehet annak információit kinyerni. Az olvashatóság kutatása ezért az oktatás során használt szövegek esetében nélkülözhetetlen. Az olvashatósági formulák célja ezen jellemző matematikai úton történő meghatározása (DuBay 2004). Magyar nyelvű szövegeken alkalmazható formula még nem készült. Ezért jelen tanulmány célja az olvashatóság alapvető elméleti kereteinek áttekintése, a formulák rendszerezése változók alapján. Ezért a szakirodalmi elemzés során először az olvashatóság és az olvasható szöveg meghatározásait járjuk körbe, majd a leggyakrabban használt amerikai formulák közül ötöt bemutatva szemléltetjük azok működési elveit, magyar nyelvű adaptációjuk lehetséges irányait, végül pedig a formulákat érő leggyakoribb kritikákat ismertetjük. Legfontosabb megállapításaink összefoglalása: Bármely szöveg három jellemző kölcsönhatása révén válik olvashatóvá – a szöveg tartalma és szerkezete; az olvasó olvasásértési képessége, motiváltsága és tudása; a vizsgálat célja és jellege –, amelyek közül az olvashatósági formulák az elsöre fókuszálnak, különböző számítási módszerek segítségével (Dale és Chall 1949). A formulák közelebbi vizsgálata során ezen módszereket a mért változók mentén három csoportba rendszereztük, melyek a (1) szófamiliaritáson alapuló, a szóhosszúságot (2) szótagszámban és (3) betűszámban mérő formulák. A magyar nyelvi adaptálhatóság szempontjából a harmadik csoport bizonyul legcélszerűbb kiindulópontnak. Az összefoglalt elméleti keret és a formulák rendszerezése a formulák magyar nyelvi adaptációját, az olvashatóság hazai diskurzusának kibontakoztatását hivatott elősegíteni.

Kulcsszavak: olvashatóság, olvashatósági formula, szövegértés, olvasástanítás