

Joachim László¹

A MAGYAR TANULÓI SZÓTÁRAK CÍMSZÓKIVÁLASZTÁSÁNAK FŐBB KÉRDÉSEI (és néhány szempont a magyar alapszókincs meghatározásához)

Abstract

The article addresses the main questions relating to the headword selection process and suggests some theoretical and methodological factors to consider. It briefly discusses some preliminary decisions to make concerning the size of the dictionary (and the financial and time constraints of the dictionary making process, all based on the aims of the dictionary and the target group) and also the question of what to include in a learner's dictionary as headwords. The paper presents different possible sources of headwords (regarding them as possibly complementary): corpora, collecting words "manually", using lists of previously selected words (e.g. language books, the Hungarian version of *The Threshold Level*, dictionaries, translation glossaries) and words suggested by users of online dictionaries. The article also discusses different approaches and methods for selecting the headwords: frequency, communicative value, drawing up a list of possible domains to avoid glaring omissions and also considering the influence of the other language and culture (in case of a bilingual dictionary). The paper also states the overlap of the discussed aspects with those of designing a core vocabulary and highlights the practical significance of such a project.

Keywords: *headword selection, building the headword list, Hungarian learner's dictionary, core vocabulary, basic vocabulary*

Kulcsszavak: *szógyűjtés, címszókiválasztás, címszókiválogatás, címszavak, tanulói szótár, alapszókincs*

¹ Joachim László, magyar mint idegen nyelv tanár, Református Missziói Központ – Menekültmisszió (Iskolai Integrációs Program), laszlo.joachim.hu@gmail.com

1. Bevezetés: a magyar (mint idegen nyelvi) tanulói szótárak helyzete

Magyar nyelvkönyvek és segédanyagok jelentős bővülésének lehetünk szemtanúi az elmúlt néhány évben (kezdő szinten különösen), a magyar tanulói szótárak helyzete azonban ennél kisebb mértékben javult². Pedig egyre több különféle típusú magyar tanulói szótárra lenne szükség, hogy segíteni tudja az egyre számosabb, nagyobb és sokfélébb (mind az anyanyelv, mind például a nyelvi szint vagy a tanulási cél szempontjából) tanulói csoportokat. Az eltérő szerkesztési szempontok miatt sem a magyar anyanyelvűek számára írt kétnyelvű tanulói szótárak (pl. magyar–angol szótár magyaroknak), sem pedig az ugyancsak alapvetően magyar anyanyelvűek számára készített értelmező szótárak nem nyújtanak számukra elegendő segítséget (Magay 2006: 83–84, Szöllősy 2006: 176, ill. Magay 2006: 60–63). A magyar tanulói szótárak szerkesztésének egyik lényeges kérdését jelentik a címszókiválasztás elvei, jelen tanulmány ennek az alapvető és nehéz munkafolyamatnak a főbb szempontjait tekinti át.

2. A magyar tanulói szótárak koncepciójának és terjedelmének kérdése

Egy szótár terve az adott szótár iránti használói igényből indul ki (érdemes minél inkább tudatában lenni a célcsoport jellemzőinek és igényeinek, vö. Atkins–Rundell 2008: 27–35, Gyöngyösi 2012), ez és a költségvetés határozza meg az erőforrásokat, a szótár terjedelmét és a határidőket (Atkins–Rundell 2008: 18, vö. Magay 2004: 58). A tartalommal kapcsolatos részletek eldöntése pedig alapvetően mindezek ismeretében történik. A használói igények és az elérhető költségvetés könnyen ellentmondásba is kerülhetnek egymással, többek között mivel a szótárhasználók többnyire azt szeretnék, hogy az általuk használt szótár szinte minden helyzetre kínáljon megoldást. Természetesen egyetlen szótár sem képes arra, hogy minden hasznos szót tartalmazzon, a terjedelmi lehetőségekhez képest azonban minél több, a célcsoport(ok) számára fontos szituációt kell lefednie. Ebből az is következik, hogy tanulói szótárak készítésekor alapvető, de ugyanakkor nehéz feladat is a legmegfelelőbb címszavak kiválogatása (Atkins–Rundell 2008: 178, P. Márkus 2011: 19).

Szigorú terjedelmi vagy időbeli korlátok esetén fontos tudatában lenni, hogy minél nagyobb a címszóállomány, általában annál kevesebb információt tud megadni a szótár egy-egy szóról. Tanulói szótárak esetében ezzel kapcsolatban inkább az a kívánatos,

² Ha a (szinte biztosan a legbővebb kínálatlal rendelkező) magyar és angol nyelvpárt nézzük, akkor azt láthatjuk, hogy létezik közöttük például képes szótár (a nemzetközi sorozat részeként megjelent Davies–Szabó 2009), tematikus szótár (Pelcz–Szita 2011), gyermekeknek készített szótár (Bíró–Kontra–Radnai 1989), elektronikus szótár (konkrét tananyaghoz kapcsolódó tíznyelvű szótár, amelynek egyik nyelve az angol, Nádor 2009), sőt az újtiszótárak között is van, amelyik talán itt is említhető (Magay – Mentlné Láng – Skripecz – Rátz – Végh 1991 és későbbi kiadásai). Ezek többsége azonban viszonylag kis méretű, és nem is mindegyik ad meg elegendő, a tanulók számára fontos információt (vö. például a grammatikai információk és a példák, példamondatok) a felvett címszavakról. Más nyelvet használó magyarul tanulók ennél sokkal kevésbé vannak ellátva tanulói szótárakkal (ha egyáltalán), általános egynyelvű magyar tanulói szótárról pedig nincs is tudomásom (most csak néhány alapvető tanulói szótártípust említve).

hogyan „az alapszókinccset mennél intenzívebben, mélyrehatóbban dolgozza föl mind szemantikailag, mind szintaktikailag és szintagmatikailag” (a kétnyelvű tanulói szótárak kapcsán Magay 2006: 83).

3. Mi számítson címszónak?

Erre a kérdésre érdemes egyfajta előzetes, általános „technikai” definícióval³ válaszolni, amely a későbbi döntések alapelveül szolgálhat (a válogatás elveiről ld. az 5. részt). Fontos a valamilyen szempontból **határhelyzeteknek minősülő esetek** megfelelő kezelése (a szókészlet hagyományosan periférikusnak tekintett bizonyos elemeit ld. az 5.3. részben).

A **kifejezések** (például *hogyan vagy?, egészségedre! valahány óra múlt*) nyomtatott tanulói kézisztárakban gyakran nem önálló címszóként, hanem vezérszó alatt szerepelnek (P. Márkus 2011: 23), elektronikus szótáraknál pedig a kifejezés bármely szavára (vagy akár egészére) történő keresésnél megjeleníthetők⁴ (Prószéky 2004: 84). Az utóbbi esetben így hangsúlyos kérdés lehet, hogy minden kifejezést valamelyik szócikk részeként érdemes-e kezelni, vagy egy részüket esetleg szerencsésebb lenne önálló címszóvá emelni (ld. az önállóan soha elő nem forduló szavakat tartalmazó kifejezéseket, pl. *szabadlábra helyez*, Prószéky 2004: 83).

Akár **toldalékok** címszóként való szerepeltetésének kérdése is felvetődhet, hiszen ez segítheti képzett, illetve todalékolt szavak megértését. A magyar anyanyelvűek számára készített *Értelmező kézisztárban* (Pusztai 2003) és az angol *Oxford Advanced Learner's Dictionary* egy nyelvű tanulói szótárban (Hornby – Wehmeier 2010) is szerepelnek todalékok címszóként (pl. *-ban/-ben*, ill. *-ed, -ion*), ennek a gyakorlatnak az esetleges átvétele, illetve egy akár nem feltétlenül nagy terjedelmű magyar tanulói szótárban való alkalmazásának részletei azonban a kifejezetten gazdag alaktanú magyar nyelv esetében alapos végiggondolást igényelnek.

A **képzett szavak** közül Pálffy (2004: 76) szerint azokat kell felvenni a szótárba, „amelyeknek a jelentése nem következtethető ki az alkotóelemekéből”, és „amelyeknek az idegen nyelvű megfelelőit az átlagos szótárhasználó nem tudja az ismeretei alapján generálni”. Ugyanez az elv a **szösszetételekkel** kapcsolatban is követhető (Pálffy 2004: 77).

Atkins és Rundell (2008: 179) szerint érdemes számba venni, hogy a szótárban **mely szófaji típusok és altípusok szerepelnek majd címszóként** (a zárt szófajok szavai között könnyebb a válogatás, P. Márkus 2011: 20), illetve kiegészíthetjük a kérdést **egyes szófaji (al)típusok bizonyos todalékolt alakjaival** is: címszavak legyenek-e például az *enyém*, a *kéne* vagy a *neki* alakok? Az ezzel kapcsolatos döntés során is szem előtt tarthatjuk Pálffy képzett szavakra vonatkozó elvét, a válasz pedig ezért többek között a

3 A címszó-definíció megalkotásakor meg kell nézni azt is, hogy a felhasználni kívánt korpusz(ok) és más források a szó milyen definíciója alapján működnek (Carter 1998: 234), és hogyan lehet szükség esetén összehangolni a különböző felfogásokat, vagy reagálni a különbségekre.

4 Érdemes is élni a lehetőséggel, hiszen a felhasználók gyakran nem szavakat, hanem kifejezéseket, sőt akár egész mondatokat írnak a keresőmezőbe.

másik nyelvtől is függhet (ld. az 5.4. részt). A döntés a gyakorisági elv érvényesítésének részleteire is kihathat (ld. a szótöves és a toldalékolt gyakorisági listák, 5.1. rész), valamint az utaló címszavak kérdéséhez⁵ is kapcsolódhat.

Mindez nem feltétlenül csak a szótárba felveendő formák gyűjtésének a kérdése, hanem az információ elrendezését is (**egységnyi információ hány szócikkre oszlik**): például hogyan kezeli a szótár a homonímia és a poliszémia kérdését, irányítószóként vagy a címszó részeként adja-e meg a vonzatokat, címszóként szerepeltet-e egyes kifejezéseket vagy toldalékolt alakokat.

4. A szótár szavainak forrásai

4.1. Egy tanulói szótár alapjául szolgáló korpusz(ok)

Egy tanulói szótár készítését megalapozó **korpusznak többféle szövegtípust, regisztert, írott és beszélt nyelvi adatokat** is kell tartalmaznia, még hozzá **a céloknak megfelelő arányban** – nehéz is ennél sokkal többet mondani az arányokról, mivel azok súlyozására tudományos módszer nem létezik (Atkins–Rundell 2008: 66). Gyakorlat és minták persze vannak, a források és az arányok eldöntése pedig (mint a döntések általában) számtalan kérdést vet fel, többek között a különféle nyelvváltozatokkal kapcsolatban (Lee 2001: 252-254). Ezek egyike a beszélt nyelvváltozatok reprezentálásának kérdése: egy tanulói szótár számára ideális korpusznak többek között hétköznapi, köznyelvi, spontán beszélgetésekből, telefonbeszélgetésekből, iskolai / egyetemi órákról, értekezletekről, interjúkból, beszédekből, híradókból is kell beszélt nyelvi adatokat tartalmaznia (Atkins–Rundell 2008: 68, Fóris 2002: 89). A magyar korpuszok egy része internetes fórumokról vagy filmfeliratokból nyert nyelvi adatokkal reprezentálja a beszélt nyelvet⁶, de emellett ma már kisebb méretű, kimondottan beszélt nyelvi adatokat feldolgozó korpuszok is léteznek, ilyen a *Budapesti Szociolingvisztikai Interjú* (BUSZI) és a *Beszélt nyelvi adatbázis* (BEA)⁷.

Az általános nyelvhasználatot tükröző korpuszok mellett adott esetben **speciális (rész)korpusz** is felmerülhet mint szótári címszavak forrása (akár a szótár alapjául szolgáló korpusz részeként; a korpuszkészítésről ld. Atkins–Rundell 2008: 53-96, valamint Szirmai 2005). Ide kapcsolódó minta lehet (még ha nem is szótár, hanem egy adott terület legfontosabb szavait feltérképező lista) az angol nyelv esetében az ötmillió szavas korpusz alapján készült, az iskolai tananyag szókincsére koncentráló *The American*

⁵ Az utaló címszavak elsősorban nyomtatott szótárak esetében fontosak, online szótáraknál az átirányíthatóság és a morfológiai komponens át tudja venni szerepüket, hiszen ez utóbbi a címszótól nagymértékben eltérő toldalékolt alakokat (pl. *velem, tegyél*) is képes kezelni (vö. Prószéky 2004: 82, ill. uo. 84-85, Joachim 2013: 51-54).

⁶ A *Magyar Nemzeti Szövegtár* személyes regiszterének forrását például internetes fórumok jelentik, erről a vonatkozó oldalon az olvasható, hogy „bizonyos esetekben nagyon hasonlít a beszélt, élő kommunikációhoz”. A filmfeliratokat is használó *Hunglish* korpuszról, valamint annak angol nyelvoktatásban történő hasznosíthatóságáról pedig ld. Sass 2007.

⁷ A BUSZI és a BEA létezésére Sass Bálint hívta fel a figyelmemet.

Heritage Word Frequency Book (Carroll–Davies–Richman 1971) vagy az *Academic Word List* (Coxhead 2000). Ez utóbbi különféle tudományterületek szövegeiből összeállított, egyfajta „általános tudományos szaknyelvi” korpusz gyakorisági listájának azon része, amely nem szerepel az angol nyelv egyfajta alapszókincsének tekintett *General Service Listben* (West 1953)⁸.

4.2. „Kézi” gyűjtés, illetve mások által korábban végzett gyűjtések, listák felhasználása

Bizonyos esetekben a klasszikus **manuális címszógyűjtés** is jól működhet, **a legkülönfélébb források**, például könyvek, híradó vagy az internet felhasználásával (B. Papp 2012: 75). Kézi gyűjtés esetén fokozottan kell figyelni arra, hogy ne maradjon ki fontos terület (ld. a 5.3. részt). Mások által kiválogatott szókincset (szövegek és/vagy listák formájában) tartalmazó munkák: **szótárak, nyelv- és szakkönyvek, az egyes nyelvi szintek leírásai, fordítók számára készített internetes (szak)szójegyzékek, adatbázisok / tudásbázisok** is lehetnek címszavak forrásai (vö. Fóris 2002: 64–65). Nem reprezentatív korpusz alapján készült munka forrásként való használata esetén ugyanakkor (a kézi gyűjtéshez hasonlóan) fokozottabban jelen van a címszavak válogatásában a szubjektivitás, a többé vagy kevésbé véletlenszerű kiválasztás.⁹

A címszavak kiválasztását segítheti például a *Küszöbszint* kommunikatív szemléletű leírása (Aradi–Erdős–Sturcz szerk. 2000), még ha önmagában nem is feltétlenül elegendő (különösen nagyobb szótárakhoz). Kardos (2009: 204) szerint ez a munka „a magyar nyelv alapszókincsének egyetlen, többé-kevésbé reprezentatívnak tekinthető gyűjteménye”, és saját meglátásom szerint is jó (egyik) forrás lehet. Nem szabad azonban elfeledkezni arról, hogy a *Küszöbszint* céljának megfelelően egy keret¹⁰.

4.3. Online szótáraknál: a felhasználók bővítési javaslatai

Online szótárak esetében a kezdeti címszólista további bővítésének jó forrása lehet, ha **a felhasználók is javasolhatnak újabb szavakat**, azt is megadva, hogy hol találkoztak az adott szóval (a javaslatokat azonban érdemes lehet szűrni és a megfelelő információkkal kiegészíteni, vö. Szöllősy 2006: 178). Szintén sokat segít, ha **a rend-**

⁸ Vizsgálatok szerint a GSL általános 80% körüli lefedettségi mutatójához további nagyjából 8–10 százaléket tesz hozzá tudományos szövegekben ez a nagyjából 570 szócsalád (Nation–Waring 1997: 16–17), ami jóval nagyobb növekedés, mint amit az általános gyakorisági lista következő ugyanennyi elemével el lehetne érni.

⁹ Olasz tankönyvek szókincsvizsgálata legalábbis ezt az eredményt hozta: „lexikális segédeszközök rendszeres használata nélkül, bármennyire tapasztaltak legyenek is a szerzők, lehetetlen olyan döntéseket hozni, amelyek koherensen tükrözik a tényleges nyelvhasználatot” Kardos (2009: 209).

¹⁰ Az egyes szintek leírásai számos fogalomkört (néhány jellemző példa megadása után) nyitva hagynak (van Ek, J. A. – J. L. M. Trim 1998: 177), elismervén, hogy azok azonos tartalommal való megtöltése nem lenne praktikus minden tanuló számára. Ezen kívül szómutatójukat az egyes kommunikációs céloknál megadott, időnként meglehetősen általános szerkezetekkel kapcsolatos példamondatok is befolyásolják, ami pedig ugyancsak magában hordozhatja az ad hoc veszélyét. Ezen kívül többek között arról is vannak viták, hogy mennyire határozta meg a Küszöbszint készítésének folyamatát az eredendően más szerkezetű nyelvekre hangolt általános keret és esetleg a más nyelvű minták, ill. mennyire lehetett a magyar nyelv(használat) sajátosságait is szem előtt tartani, figyelembe venni a leírás során.

szer naplózza a kereséseket, feltüntetve, hogy eredményeztek-e találatot – ahogy ez a magyar–szomáli tanulói kisszótár online változatánál is történik (Joachim – Tukale 2013). Mindkét esetben tulajdonképpen „házhoz jön” a szerkesztőhöz a bővítés egy részének anyaga, és ezekből az adatokból ráadásul érdekes következtetéseket lehet levonni arról, hogy milyen szavak hiányoztak a szótárból, ill. milyen szavakra van szükségük a nyelvtanulóknak. (Az eredmények pedig még úgy is hasznosak, hogy a kontextus hiányával, illetve rosszul írt alakokkal is számolni kell.) A bővítés során szótárba kerülő címszavak (és szócikkek) forrásnyelvi és célnyelvi alakokból egyaránt kiindulhatnak (vö. az 5.4. rész).

5. A címszavak kiválogatásának elvei, módszerei

5.1. *A(z általános) gyakoriság mint szempont*

Egy nyelv tanítási céllal kiválasztott legfontosabb szavai között kétségkívül alapvető a (nyelvtanulók számára legrelevánsabb kontextusokban, illetve az ezeket reprezentáló korpuszban vagy korpuszokban) **leggyakrabban előforduló szavak**¹¹ hangsúlyos jelenléte. Az olasz nyelv Tullio De Mauro és csoportja által meghatározott alapszókincsben a mintegy 7000 szó közül 5000 a gyakoriság szempontja alapján szerepel (Fóris 2012: 285, Kardos 2009: 203).

A nyers gyakorisági adatokat persze nem érdemes automatikusan kezelni¹². Kiválasztási kritérium lehet, hogy a szavak a legkülönbébb szövegekben is előforduljanak¹³ (még ha nem is egyforma gyakorisággal, vö. Nation és Waring 1997: 18). Kérdés az is, hogy egy-egy szó milyen jelentésében mennyire gyakori (Atkins–Rundell 2008: 61, Nation–Waring 1997: 13), illetve, hogy milyen alakjában mennyire az.¹⁴ Hogy egy-egy szó mennyire fontos a nyelvtanulás szempontjából (gyakorisága vagy bármely más elv alapján), azt akár magában a tanulói szótárban is hasznos lehet jelezni valamilyen módon (Szirmai 2005: 135).

¹¹ Szó alatt a szótárszerkesztő címszó-definícióját értem. Nation és Waring (1997: 18) szerint például ha egy idioma vagy kifejezés felveszi a versenyt gyakoriságban a szavakkal, és jelentése nem következtethető ki egyszerűen alkotóelemei jelentéséből, akkor érdemes lehet szerepeltetni egy gyakorisági listában. (És ha egy szótárban ezeket nem kezeljük címszóként, a fenti kritériumoknak a példák válogatásakor való alkalmazása akkor is megfontolandó, vö. Sass–Várad–Pajzs–Kiss 2010.)

¹² A bekezdés további részében tárgyalt szakmai okok mellett ennek technikai oldala is lehet: egy ilyen listában könnyen felbukkanhatnak tulajdonnevek, számok stb. (ha a korpusz nincs már eleve megtisztítva az ilyen elemektől, vö. Atkins–Rundell 2008: 84–86).

¹³ Lee (2001: 255) a szóródás (dispersion) kritériumát összekapcsolja a használati értékkel, az adott szavak sok szituációban való használhatóságával és a „rendelkezésre állással” (vö. a francia „disponible”, ld. az 5.2. részt).

¹⁴ A 3. részben tárgyalt alapvető döntések meghozatala után azt is érdemes megvizsgálni, hogy a felhasználni kívánt korpusz(ok) a szó milyen definíciója alapján működnek (Carter 1998: 234), és szükség esetén összehangolni az eltérő felfogásokat, vagy reagálni a különbségekre. Ez hatással lehet arra is, hogy szótóves vagy toldalékolt gyakorisági listát használunk, netán mindkettőt (ld. pl. az MNSZ személyes regiszterének vele, kéne stb. formáit), hiszen a korpusz szó-felfogásából következik, hogy mely formák gyakorisági adatait kapcsolja egy-egy szótóvéhez.

5.2. Kommunikatív / használati érték

Egy tanulói szótár szerkesztésekor a gyakorisági szempont mellett annál szubjektívebb, nehezebben mérhető kritériumokat is kell alkalmaznunk: olyan, jellemzően **a hétköznapi élet tárgyaival, fogalmaival kapcsolatos szavakat** is fel kell venni a címszavak közé, **amelyek nem szerepelnek a(z általános) gyakorisági listák élén, ismeretük ugyanakkor fontos a nyelvtanulók számára** (vö. Carter 1998: 225-226, Chiari–De Mauro 2010: 23, Kardos 2009: 202-203, Magay 2006: 81). Kardos (2009: 209) rendelkezésre álló szavaknak¹⁵ nevezi őket (vocabolario di alta disponibilità, vocabulaire de haute disponibilité, high availability vocabulary): az elnevezés eredendően a szavak mentális lexikonból való könnyű előhívhatóságára utal (Carter 1998: 235), és egy-egy téma leginkább hozzáférhető szavait jelenti¹⁶ – ezek természetesen egyénenként, illetve nyelvenként, nyelvközségenként is mutathatnak eltéréseket egymástól. A gyakoriság mellett ez volt az olasz alapszókincs újabb meghatározásának másik fontos elve francia minta alapján (Chiari–De Mauro 2010: 25-26, Kardos 2009: 203), de az angol *General Service List*nél sem a gyakoriság az egyetlen szempont¹⁷.

5.3. A szókészlet csoportokra bontása és súlyozása

Amellett, hogy a tanulói szótáraknak érdemes felvenniük az alapszókincset címszavaik közé (vagy kisebb szótárak esetében az alapszókincs egy részét), a magyar szókészletet érdemes már a konkrét címszókiválasztás előtt **többféle szempont alapján is különféle csoportokra osztani**¹⁸, és eldönteni, hogy az egyes csoportokba sorolható szavakat felvegye-e a szótár, és ha igen, milyen **súlyozással**, vagy ne kerüljön bele (Atkins–Rundell 2008: 184, valamint vö. Pomázi 2004: 158). Milyen mértékben szerepeltessünk (ha szerepeltetünk) például szlenget vagy diáknyelvi szavakat, informális szavakat vagy kifejezéseket, csúnya szavakat, hivatalos elnevezéseket, a nyelvből eltűnőfélben lévő vagy frissen megjelent szavakat, különféle területi nyelvváltozatok szavait, egyes speciális tárgykörökhöz, szituációkhoz, beszédszándékokhoz tartozó szókincs bizonyos elemeit? És még a leghétköznapibb kategóriák esetében is: milyen mennyiségben szerepeltessünk szavakat az adott esetben akár több tízezer lehetséges szó közül

¹⁵ Fóris (2012: 285) nagyon fontos szavaknak hívja őket, és a kategória típusait is megadja: „alapvető biológiai cselekedeteket, mint evés, alvás, szülés, halál stb. megjelölő szavak; a főbb testrészek nevei (szem, száj, fej stb.); a külső természet jelenségei (tűz, víz, Nap, Hold stb.); általános viszonyfogalmakat kifejező szavak (személyes névmások, mutatószók, tagadószók, mértékek stb.); kultúraspecifikus szavak”.

¹⁶ Ezt eredetileg kérdőívvel mérték (Michéa 1953: 341).

¹⁷ Nation és Waring (1997: 13) az egyes formákhoz tartozó különböző jelentések gyakorisági adatainak feltüntetése mellett ezért is tartják az angol nyelv legjobb szógyakorisági listájának. West gyakoriságon kívüli szempontjai között szerepelt, hogy érdemes előnyben részesíteni például az érzelmi többletjelentést hordozó szavakat, illetve azokat, amelyek más szavakkal vissza nem adható gondolatokat fejtenek ki, de ezen kívül a tanulás nehézsége is szempont volt: könnyebb egy ismert szó újabb jelentését megtanulni, mint egy teljesen új szót (Nation–Waring 1997: 19). Mai szemmel így is felvethető, hogy West akár nagyobb nyomtatékot is adhatott volna a hasonló szempontoknak (Carter 1998: 208).

¹⁸ Az adott szótárban szerepeltetendő területek (domain) meghatározásakor egy hierarchikus fogalomkörü rendszer felállítása sokat segíthet abban, hogy ne maradjon ki fontos terület (Atkins–Rundell 2008: 184). A csoportokra bontás nemcsak szótárszerkesztéskor, az alapszókincs adott célú kiegészítésekor hasznos, hanem akár az alapszókincs meghatározásakor is, hiszen például a rendelkezésre álló szókincset is konkrét témák mentén lehet felmérni.

(például állatfajták, betegségnevek, vö. Magay 2004: 58)? Vagy törekedjünk inkább az általánosra és a (proto)tipikusra (vö. Lee 2001: 254-255)? Hasonló kérdések merülnek fel a **szakszavakkal**¹⁹, például bizonyos hobbik jellemző szókészletével vagy egyes foglalkozások, szakterületek szakszargonával, szókincsével kapcsolatban is (vö. P. Márkus 2011: 24). Természetesen ezeknek a döntéseknek is összhangban kell lenniük az adott szótár profiljával, címszövegfogásával, a címszókiválasztás e dolgozat más részeiben ismertetett elveivel, valamint a területi lehetőségekkel, adottságokkal is.

A csoportok meghatározásakor nem szabad elfeledkezni a **szókészlet hagyományosan periférikusnak tekintett bizonyos elemeiről** (Magay 2004: 59), így például a tulajdonnevekről sem: felvetődhet például a nagyobb magyarországi városok szerepeltetése (a helyhatározóragok miatt is: például *Budapesten*, de *Debrecenben*), a fontosabb, teljesen vagy részlegesen magyar nyelvű földrajzi nevek (pl. *Bécs*, *Pozsony*, *Földközi-tenger*, *Duna*, *Alföld*) és más tulajdonnevek felvétele is, például az *ENSZ*, a *Kúria*, a *Hold* stb. (Atkins–Rundell 2008: 188). Fontos elemek lehetnek gyakoribb, **sajátosan a beszélt nyelvre jellemző fordulatok, felkiáltások** (Fóris 2002: 90), vagy az önálló szavak határán többek között becézett alakok (*zsepi*, *kösz*, *Laci* stb.), **rövidítések** (pl. *u.*, *krt.*, *pl.*) és mozaikszók is (pl. *ÁFA*, *MTA*, *BKV*, *kb.* – persze ezek egy része igen könnyen el is évülhet).

Bizonyos, a **célnyelvi kultúrával kapcsolatos enciklopédikus, kulturális információk** szerepeltetése²⁰ is gyakran szükséges a tanulói szótárakban, hiszen számtalan szó, amelyet az anyanyelviek (kultúrájuk alapvető ismeretének birtokában) a legnagyobb természetességgel, utalásszerűen használnak, a külföldiek számára magyarázatot igényel – ezek kiválogatása igen nehéz feladat (Magay 2004: 59), és igencsak szubjektív (Lukács 2010: 41, B. Papp 2012: 75-76). Az egyes csoportok között lehetnek átfedések, amint az az előző bekezdésben szereplő példák némelyikénél is látszik. Az egyes területek súlyozásakor különleges figyelmet kell fordítani a **speciális, konkrét célcsoportnak releváns szavakra** is. Sok angol egynyelvű tanulói szótár például angoltanár használói miatt tartalmaz számtalan nyelvészeti szakszót (vö. Atkins–Rundell 2008: 190), de ugyanígy a tanulók miatt például a tanulási környezet szavait vagy a nyelvórai feladatok jellegzetes instrukcióinak szókincsét is szerepeltetni kell.

¹⁹ A köznyelv és a szaknyelvek között ugyanakkor nem is lehet éles határt húzni: a modern technológia és a tudomány számos eredménye a hétköznapi élet részévé vált, és ez természetesen a szótárakban is tükröződik (Magay 2004: 59).

²⁰ A kultúraspecifikus, „lefordíthatatlan” elemek jelentésmegadását a szótárak körülírással szokták megoldani (Márkus–Szöllősy 2006: 102). Atkins és Rundell (2008: 187) vonatkozó példáihoz hasonló magyar megfelelő lehet pl. a Mikulás (mint a magyar kultúra egy olyan eleme, amely számos magyarnyelv-tanuló kultúrájának nem része), de sok más területre, például eltérő oktatási rendszerekre vonatkozó szavak, fogalmak megadásakor is felmerülhet a rövid magyarázat szükségességének kérdése. A kultúraspecifikus szavakat nem (vagy nem elsősorban) önálló fogalomkörü csoportként érdemes felfogni, hanem inkább a szókészlet különféle csoportjaiban előforduló, bővebb magyarázatot, illetve különleges figyelmet igénylő elemekként.

5.4. Kétnyelvű szótáraknál: a másik nyelv és kultúra hatása

Egy tanulói szótár készítésekor nem csak a magyar nyelv **fontos kiindulópont**, hanem **a másik nyelv is**²¹, hiszen ahogy Kardos (2009: 204) fogalmaz: a magyar nyelv és a kultúra „exportálása” mellett a beszélő „importálása” is segíti a szótárat használó nyelvtanulók kommunikációját. Vagyis a magyar nyelv autentikus szerkezetei mellett „a nyelvi elemeknek a tanuló országában elfoglalt használati értékét” sem szabad figyelmen kívül hagyni (vö. a másik nyelv alapszókinccse, legfontosabb szavai, amennyiben elérhető információ erről). A címszókiválasztás szempontja ezen kívül a fordíthatóság is: ha egy szó (valamilyik vagy mindkét irányban történő) fordítása nehézséget okoz a nyelvtanulók számára²², az a szótárba történő felvétele mellett szól (Atkins–Rundell 2008: 190). Az is ide tartozik, hogy szavaknak egy másik nyelvben nem feltétlenül szavak felelnek meg: például ha magyar toldalékok másik nyelvű ekvivalense önálló szó, akkor az a magyar alak esetében is érv lehet a címszóként szerepeltetés mellett (az angol és a magyar nyelv viszonylatában ilyenek lehetnek a viszonyragok, például *-ban/-ben, -tól/-től, -ig*, P. Márkus 2011: 21, vö. a 3. rész). Emellett pedig érdemes azon szavak közül is szerepeltetni némelyeket, amelyek **a tanulók saját kultúrájának fontos részei**, és létezik valamilyen szintű megfelelőjük magyar nyelven (ezek az elemek persze nem feltétlenül azonosak a magyar nyelv leggyakoribb szavaival, vö. Kardos 2009: 208, valamint Magay 2006: 83).

6. A válogatási elvek, módszerek egy másik alkalmazási területe: a magyar (mint idegen nyelv) alapszókinccsének meghatározása

Egy tanulói szótár címszavainak kiválogatása és a pedagógiai értelemben vett alapszókinccs meghatározása két olyan feladat, amely mutat közös vonásokat, így az eddigi utalások mellett érdemesnek látszik az utóbbi kérdéssel önállóan is, legalább a témafelvetés szintjén foglalkozni. Nyelvi minimumok meghatározásakor a korábban tárgyalt elvek közül elsősorban az első kettőnek van jelentősége, „az utóbbi időben a két elképzelést egy **gyakorisági-kommunikatív keret**ben próbálják meg integrálni” (Kardos 2009: 202).

Természetesen többféle (részben egymást kiegészítő, részben egymásnak ellentmondó) koncepció alapján is meg lehet határozni egy nyelv pedagógiai szempontú alapszókinccsét (Carter 1998: 36–46, Lee 2001: 251–256). Szavak listáinak²³ készítése mint eszköz, illetve mint módszer is természetesen vitatható: az alapszókinccshez tartozás

²¹ Ez a szempont egynyelvű szótárak szerkesztésekor is ott lehet a háttérben, még ha ilyenkor nehezebb is a kérdés kezelése a forrásnyelv és az ahhoz kapcsolódó kultúra vagy kultúrák sokfélesége miatt.

²² Ennek speciális esete lehet a hamis barátok jelensége (ide értve azokat a szópárokat is, amelyeknél az azonos vagy hasonló forma nem elsősorban fogalmi, hanem stílusbeli különbségekkel jár együtt), illetve az, ha egy adott szónak nem is létezik megfelelője a másik nyelven.

²³ A listákat nem csupán szavak felsorolásaként érdemes felfogni, hiszen ezek akkor lesznek jól felhasználhatók (például tananyagkészítés vagy más szótárak készítése során), ha sok információt tartalmaznak az egyes szavakról (a gyakorisági listákkal kapcsolatban vö. Nation–Waring 1997: 19). A General Service List például tartalmazza a szófajt, az adott szó különféle jelentéseit és az azokhoz tartozó gyakorisági adatokat, valamint kifejezéseket is (Carter 1998: 208).

fokozat kérdése is (Carter 1998: 238), illetve nem is feltétlenül van egyetlen, minden tanulói csoportnak egyformán hasznos lista (Kardos 2009: 204, van Ek, J. A. – J. L. M. Trim 1998: 177). A magyar (mint idegen) nyelv **valamilyen szintű konszenzus alapján meghatározott alapszókincse** mindezek ellenére azonban a gyakorlatban a **tanulói szótárak egyik forrásaként szolgálhatnak, és hivatkozási alapot adhatnak emellett nyelvkönyveknek, gyakorlókönyveknek és vizsgafeladatsoroknak** is.²⁴ Megfelelő szoftverrel pedig azután bármilyen szöveg elemezhető abból a szempontból, hogy a megadott listában (vagy akár a lista vagy listák egyes szegmenseiben) szereplő formák milyen arányban fedik le (az angol nyelvre több ilyen ingyenesen elérhető szoftver, ún. vocabulary profiler is létezik).

7. Összegzés

Tanulmányomban a magyar tanulói szótárak készítésének egy lényegbevágó munkafázisáról, a címszavak kiválasztásáról, elsősorban annak forrásairól és elveiről, módszereiről volt szó. (A konkrét helyzetekben való döntés azután természetesen a lexicográfusok joga és feladata.) Egy szótárt lehetőség szerint elkészülte után is kell gondozni, ennek részeként pedig a címszólista karbantartása is fontos, hiszen a legjobb szótár is elavul idővel (Magay 2004: 58), így előbb vagy utóbb további kiadására, illetve frissítésére lesz szükség.

Emellett pedig – bár alapvető témám a tanulói szótárak címszókiválasztása volt – a magyar alapszókincs meghatározásának szótárakat és a magyar mint idegen nyelv oktatását általánosabban is érintő kérdését is felvettem, mivel látok közös pontokat a két téma között. Remélem, e kérdések áttekintése a gyakorlati munka segítségére lesz.

Irodalom

- Atkins, B.T. Sue – Michael Rundell 2008. *The Oxford guide to practical lexicography*. Oxford: Oxford University Press.
- Carter, Ronald 1998. *Vocabulary: Applied Linguistic Perspectives*. London, New York: Routledge.
- Chiari, Isabella – Tullio De Mauro 2010. The new basic vocabulary of Italian: problems and methods. *Statistica Applicata - Italian Journal of Applied Statistics Vol. 22 (1)*. 23–36. http://www.alphabit.net/PDF/2012_Chiari_DeMauro_NVDB.pdf
- Fóris Ágota 2002. *Szótár és oktatás*. Pécs: Iskolakultúra. <http://mek.oszk.hu/01800/01890/>
- Fóris Ágota 2012. A lexikon megközelítései. In: Alberti Gábor – Kleiber Judit – Farkas Judit (szerk.) *Vonzásban és változásban*. (A nyelvészeti Doktorandusz füzetek különkiadása) Pécs: Pécsi Tudományegyetem Nyelvtudományi Doktori Iskola. 277–307.

²⁴ Úgy tűnik, hogy az elérhető munkák – például az egyes nyelvi szintek leírásai – inkább keretet és általános szempontokat adnak ezekhez a feladatokhoz, mint többé-kevésbé konkrétan meghatározott, leírt szókincset.

- Gyöngyösi Livia 2012. A hasznosság elve mint motiváció – a magyar mint idegen nyelv szókinccének elsajátításában. *Hungarológiai Évkönyv 13*: 29–42. http://epa.oszk.hu/02200/02287/00013/pdf/Hungarologiai_Evkonyv_13_029-042.pdf
- Joachim László 2013. Morfológiai és szintaktikai információk a magyar mint idegen nyelv tanulói számára készített szótárakban és szöszedetekben. (Egy tanulói szótár előmunkálatainak szempontjai, tapasztalatai.) *Hungarológiai Évkönyv 14*: 40–62. http://www.epa.oszk.hu/02200/02287/00014/pdf/EPA02287_hungarologiai_evkonyv_2013_14_040-062.pdf
- Kardos Orsolya Hedvig 2009. Az olasz mint idegen nyelv alapszókincse kiválasztásának elméleti és gyakorlati kérdései. In: Fábíán Zsuzsanna (szerk.): *Szótárírás és szótárírók. Lexikográfiai füzetek 4*. Budapest: Akadémiai Kiadó. 201–209.
- Lee, David Y. W. 2001. Defining core vocabulary and tracking its distribution across spoken and written genres. Evidence of a gradient of variation from the British National Corpus. *Journal of English Linguistics*, Vol. 29. 250–278. <http://www.corpus4u.org/forum/upload/forum/2005060601513713.pdf>
- Lukács András 2010. A kulturális szótár mint műfaj. *Fordítástudomány XII (1)*, 38–49. http://www.elteft.hu/f/File/Forditastudomany/2010._XII._vfolyam_1._szm.pdf
- Magay Tamás 2004. Szó, ami szó: Lexikológia és lexikográfia angol nyelvtudományi megközelítésben. In: Fóris Ágota–Pálffy Miklós (szerk.): *A lexikográfia Magyarországon*. Budapest: Tinta Könyvkiadó. 53–71.
- Magay Tamás 2006. Tanulói szótárak: múlt és jelen. In: Magay Tamás (szerk.): *Szótárak és használóik*. Lexikográfiai füzetek 2. Budapest: Akadémiai Kiadó. 57–93.
- Márkus Katalin–Szöllősy Éva 2006. Angolul tanuló középiskolásaink szótárhasználati szokásairól. In: Magay Tamás (szerk.): *Szótárak és használóik*. Lexikográfiai füzetek 2. Budapest: Akadémiai Kiadó. 95–116.
- Michéa, René 1953. Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage. *Les langues modernes*. 47: 338–344.
- Nation, Paul–Robert Waring 1997. Vocabulary size, text coverage and word lists. In: Schmitt, Norbert–Michael McCarthy (szerk.): *Vocabulary: description, acquisition and pedagogy* Cambridge: CUP. 6–19. <http://catdir.loc.gov/catdir/samples/cam031/97042211.pdf>
- Pálffy Miklós 2004. Szótári típus és címszóállomány. A címszókiválasztás szempontjai. In: Fóris Ágota – Pálffy Miklós (szerk.): *A lexikográfia Magyarországon*. Budapest: Tinta Könyvkiadó. 72–80.
- B. Papp Eszter 2012. About a Hungarian dictionary of culture. *Magyar Terminológia 5 1*: 72–78.
- P. Márkus Katalin 2011. *Kéziszótárak az ezredfordulón (Angol–magyar, magyar–angol kéziszo-tárak 1981 után)*. Doktori disszertáció. Budapest: ELTE.
- Pomázi Gyöngyi 2004. Szótárszerkesztési módszerek. (Német zsebszótárak). In: Tóth Szergej – Földes Csaba – Fóris Ágota: *Lexikológiai és lexikográfiai látkép: Problémák, paradigmák, perspektívák*. Szeged: Generalia. 157–163. <http://mek.oszk.hu/04400/04468/>

- Prószték Gábor 2004. Az elektronikus papírszótártól az „igazi” elektronikus szótárak felé. In: Fóris Ágota – Pálffy Miklós (szerk.): *A lexikográfia Magyarországon (Segéd-könyvek a nyelvészet tanulmányozásához XXXV)*. Budapest: Tinta Könyvkiadó. 81–87.
- Sass Bálint 2007. A Hunglish korpusz mint oktatási segédeszköz. In: Heltai Pál: *Nyelvi modernizáció – szaknyelv, fordítás, terminológia*. A MANYE 2006. évi XVI. kongresszusának kötete. Pécs, Gödöllő: MANYE – Szent István Egyetem. http://www.nytud.hu/oszt/korpusz/resources/sb_hunglish.rtf
- Szirmai Mónika 2005. *Bevezetés a korpusznyelvészetbe. A korpusznyelvészet alkalmazása az anyanyelv és az idegen nyelv tanulásában és tanításában*. Budapest: Tinta Könyvkiadó. http://korpusz.com/Monika/A_szerzorol_files/SzirmaiMonikaKorpusznyelvészet.pdf
- Szöllősy Éva 2006. Az MTA SZTAKI és a Szegedi Tudományegyetem on-line angol-magyar-angol szótárai: erények, hibák, fejlesztési lehetőségek. In: Magay Tamás (szerk.): *Szótárak és használók*. Lexikográfiai füzetek 2. Budapest: Akadémiai Kiadó. 165–179.
- van Ek, J. A. – J. L. M. Trim, 1998. *Threshold 1990*. (A revised and corrected edition.) Cambridge: Cambridge University Press. http://www.coe.int/t/dg4/linguistic/Threshold-Level_CUP.pdf

Források

- Aradi András – Erdős József – Sturcz Zoltán (szerk.) 2000. *Küszöbszint. Magyar mint idegen nyelv*. Budapest: Műegyetemi Távközpont. <http://bme-tk.bme.hu/other/kuszob/>
- Bíró, Ruth G. – Kontra Miklós – Radnai T. Zsófia 1989. *Hungarian Picture Dictionary for Young Americans*. Budapest: Tankönyvkiadó.
- Carroll John B. – Davies Peter – Richman Barry 1971. *The American Heritage word frequency book*. New York: American Heritage Publishing Co.
- Coxhead, Averil 2000. A New Academic Word List. *TESOL Quarterly* 34/2. 213–238.
- Davies, Helen – Szabó Helga 2009. *Beginner's Hungarian Dictionary*. Budapest: Holnap Kiadó.
- Hornby, Albert Sydney. – Sally Wehmeier 2010. *Oxford Advanced Learner's Dictionary*. Oxford: OUP. <http://oald8.oxfordlearnersdictionaries.com/>
- Joachim László – Muhyadin Hussein Tukale 2013. *Hungariyan – Soomaali qaamuus, tusiyaha qaybaha guud. / Magyar – szomáli szótár (szomálokknak). Fogalomkörü mutatóval*. (Online szótár.) Budapest: RMK-Menekültmisszió. <http://qaamuus.rmk.hu/>
- Magay Tamás – Mentlné Láng Ilona, Skripecz Sándor – Rátz Ottó – Végh Béla 1991. *Útiszótár / Tourist's Dictionary. (Angol–magyar, magyar–angol szótár)*. Budapest: Akadémiai Kiadó.
- MNSZ = Várad Tamás et al. 2005. *Magyar nemzeti szövegtár*. Budapest. <http://http://mnsz.nytud.hu/>

- Nádor Orsolya (főszerk.) 2009. *eMagyarul. Online tíznyelvű szótár*. Budapest: Balassi Intézet. <http://e-magyarul.balassi-intezet.hu/new/1.Szotar/>
- Pelcz Katalin – Szita Szilvia 2011. *Egy szó mint száz. (Magyar-angol tematikus szókinctár)*. Budapest: Akadémiai Kiadó.
- Pusztai Ferenc (főszerk.) 2003. *Magyar értelmező kéziszótár*. Budapest: Akadémiai Kiadó.
- Sass Bálint – Váradi Tamás – Pajzs Júlia – Kiss Margit 2010. *Magyar igei szerkezetek – A leggyakoribb vonzatok és szókapcsolatok szótára*. Budapest: Tinta Könyvkiadó.
- West, Michael 1953. *A General Service List of English Words*. London: Longman, Green and Co.