

Doktori (PhD) értekezés tézisei

ASSESSING ENGLISH MAJORS' VOCABULARY
AT THE UNIVERSITY OF PÉCS

WORKING TOWARDS A CORPUS-BASED APPROACH

ANGOL SZAKOSOK SZÓKINCSENEK KORPUSZ-ALAPÚ ESZKÖZÖKKEL
TÖRTÉN MÉRÉSE A PÉCSI TUDOMÁNYEGYETEMEN

Lehmann Magdolna

Témavezető : Dr. Andor József

PÉCSI TUDOMÁNYEGYETEM
Nyelvtudományi Doktori Iskola
Alkalmazott Nyelvészeti Doktori Program

2009

Az értekezés témája és céljai

A fehér papíron rejtelmesen tekergő, kacskaringós fekete betűk látványa már az első találkozások alkalmával lenyűgözött. Emlékszem, szinte könyörögtem a szüleimnek, árulják el, melyik jel mit jelent pontosan és hogyan lesznek belőlük szavak, ők pedig vonakodtak, mert akkoriban az első tanító néniük nemigen kedvelték az iskolában azokat a gyerekeket, akik már tudtak olvasni. De én nem tárgytottam, mert éreztem, ha sikerül megtudnom a betűk értelmét, egy új világ tárul majd fel előttem, és én alig vártam, hogy felfedezhessem. Azóta kiderült számomra, hogy a szavak nem csupán összefüggő betűk és ma már úgy tartom, minél többet tudok meg róluk, annál lehetetlenebbnek látszik, hogy tökéletesen megértssem az általuk megnyitott kapuk mögött rejlő világok minden lehetséges dimenzióját.

Nyelvészként és egyetemi oktatóként a szavak világával foglalkozom és az egyik ilyen mikro-világ, mely megragadta érdeklődésemet, az angol szakos hallgatók szókinccse. Saját és kollégáim tapasztalatai alapján az első éves hallgatók igen nagy nehézségekkel küzdenek a kurzusaikon számukra kötelező olvasmányok megértésével és írásfeladataik elkészítésével. Korábbi kutatások rámutattak, hogy ez a jelenség háttérben szókinccsbeli hiányosságok állhatnak. Kutatásaim tehát arra irányultak, hogy felmérjem, hány angol szót ismernek az első és másodéves, nyelvi alapvizsga előtt álló angol szakos hallgatók a Pécsi Tudományegyetemen, vagyis mekkora szókinccsel rendelkeznek; ez a szókinccs elegendő-e számukra ahhoz, hogy tanulmányaikat sikeresen végezhessék; illetve hogy ismerik-e az egyetemi angol nyelvű szövegekben leggyakrabban előforduló, és ezáltal számukra leghasznosabbnak tekinthető szavakat, és milyen jól ismerik azokat. E kérdések köré csoportosulnak az értekezésben bemutatott empirikus vizsgálatok, meghatározva a dolgozat vázát. A kutatási kérdések megválaszolása előtt azonban számos egyéb kérdést kellett tisztázni, így például azt, hogy mi is pontosan a szó és a szókinccs, mit értünk azon, hogy ismerünk egy szót, ez hogyan mérhető, és milyen szerepet játszik a szavak ismerete az olvasott szöveg megértésében, illetve az önálló szövegalkotásban.

A szó fogalmának meghatározása nem egységes a nyelvészeti szakirodalomban, attól függ, mennyire absztrakt módon, morfológiai, szintaktikai, szemantikai vagy pragmatikai szemszögből közelítjük meg a kérdést (lásd Bloomfield, 1933; Caramazza, 1999; Caron, 1992; Carter, 1998; Kempen & Huijbers, 1983; Martsa, 2007; Nattinger & DeCarrico, 1992; Pawley & Syder, 1983; Singleton, 1999; Wray, 2002). Szókincsen azonban általában az adott nyelvben, szövegben, az anyanyelvi vagy nem-anyanyelvi beszélő fejében (mentális lexikonjában) létező összes szót értjük (Carter, 1998; Ellis, 1994; Gleitman és Landau, 1994; Singleton, 2000). A lexikai kompetencia felépítése Henriksen (1999) szerint háromrétegű és több, egymással kapcsolatban álló egyenesen ábrázolható. Ezek egyike a részleges tudástól a precíz jelentés ismeretig terjed, egy másikon a szóismeret mélysége jelenítheti meg a teljesen ismeretlentől a pontosan ismert jelentésig, míg egy harmadik ábrázolásmód a receptív (szófelismerés) és a produktív (szóhasználat) végpontok között helyezi el a lexikális tudás mértékét.

Az adott nyelvtanuló vagy anyanyelvi beszélő szókincsé adott időben mindig az egyenesek más-más szakaszán található, tehát a lexikai tudás egyáltalán nem tekinthető statikus állapotnak, hanem egy dinamikusan változó jelenségként értelmezhető. Paribakht és Wesche (1999) úgy vélekedik, hogy a szókincs elsajátítás folyamata az adott szóval történő első jelentéssel bíró találkozással kezdődik és apró lépésekkel haladva jut el addig, hogy a szó, jellegzetes elemeinek, vonásainak megismerésével, beépülhessen a mentális lexikonba. Tanulmányában Laufer (1998) egy adott szó ismeretének három komponensét vizsgálja: (1) alap receptív (passzív) tudás, mely a szó leggyakoribb és központi jelentésének ismerete (Laufer, 1998, p. 257); (2) kontrollált produktív ismeret, mely szerint a nyelvhasználó képes egy adott feladat által sugalmazott szavak elhívására és alkalmazására, például egy adott mondatba beírni egy olyan szót, melynek első néhány betűje van csak megadva; a harmadik pedig a (3) szabad produktív ismeret, vagyis amikor a nyelvhasználó képes saját akaratára és választására alkalmazni az adott szót, például egy írásbeli fogalmazási feladatban. Eredményei azt mutatják, hogy a három tudástípus nem azonos mértékben fejlődött az egy tanévet felölelő kísérlet során, tehát valóban meg kell különböztetni egymástól a receptív és produktív lexikális ismeretet. A receptív

szókincs a tapasztalatok szerint mindig nagyobb a produktív szókincsnél (Webb, 2008).

A nyelvtanárok és nyelvtanulók körében leggyakrabban felmerül kérdések többnyire arra irányulnak, hány szóra és pontosan milyen szavakra van szükségük a nyelvtanulóknak. Nation és Waring (1997) rámutat arra, hogy egy nyelvtanuló számára nem minden szó lehet egyformán hasznos, és hogy mi számít hasznos vagy kevésbé hasznos szónak egy adott nyelvben, azt használatának, el fordulásának gyakorisága határozhatja meg. A tanítás és tanulás folyamatának el segítésére a nyelvészek számos gyakorisági szólistát hoztak létre nagy szöveggy jtemények, korpuszok alapján. Közülük az egyik legrégebbi, de ma is legtöbbit hivatkozott lista a *General Service List (GSL)*, melyet Michael West dolgozott ki 1953-ban. A GSL az angol nyelvben el forduló 2.000 leggyakoribb szót tartalmazza egy 5 millió szavas írott korpusz alapján, és Nation és Waring (1997) szerint ez tekinthet a nyelvtanárok és kurrikulum tervez k számára a leggyakoribb angol szavak klasszikus gy jteményének.

A GSL mellett léteznek általános és speciális tanulmányi célokra készült listák is, mint a *University Word List (UWL)*, egyetemi szólista (Read, 2000, p. 159), valamint az ennél újabb kelet , Coxhead által 2000-ben kidolgozott, 570 szócsaládot tartalmazó *Academic Word List (AWL)*, melyet egyetemeken széles körben el forduló szövegekben leggyakoribb szavak alapján állítottak össze. A szakirodalom szerint azok a szavak, melyek e listák egyikében sem található k meg, ritka (low-frequency) szavaknak tekinthet k az angol nyelvben és Nation és Waring (1997) szerint a gyors lexikai fejl dés érdekében a nyelvtanulóknak el ször a kétezer leggyakoribb szót kellene elsajátítani.

Az anyanyelv nyelvhasználó és a nyelvtanuló szókincse közötti legszembet n bb különbség az általuk ismert szavak száma, vagyis szókincsük nagysága. Goulden, Nation és Read (1990) vizsgálata alapján egy egyetemet végzett, angol anyanyelv nyelvhasználó átlagosan 17.000 alapszót ismer. Az olvasás során fejlődik a szókincs (Ellis 1994, Krashen 1989), ezt a folyamatot nevezzük saját fordításomban *járolékos* szótanulásnak [*incidental vocabulary learning*], hiszen az új szavak illetve jelentések megismerése ez esetben az olvasás kísér jelensége. Számos

vizsgálat próbálkozott a folyamat mechanizmusának megvilágításával (Rott, 1999; Paribakht és Wesche, 1999; Wode, 1999; Brown és munkatársai, 1999; Ellis és He, 1999), ám a terminus használata még nem egységes (Huckin és Coady, 1999). Nation (2006) szerint 5-6, míg Webb (2007) szerint tíznél is többször kell találkozni a szövegben egy adott szóval ahhoz, hogy jelentését az olvasás során elsajátítsuk.

Az olvasott szövegértés és a szókincs közötti kapcsolatot sokan vizsgálták (Alderson, 2000; Gass, 1999; Grabe & Stoller, 1997; Hulstijn, Hollander & Greidanus, 1996; Laufer, 1997; Nation, 2006, Paribakht & Wesche, 1997; Quian, 2002; Quian & Schedl, 2004; Zareva, 2005). A szövegek megértéséhez szükséges szókincs nagysága sokat vitatott kérdés. Carter (1998) szerint a GSL első kétezer szava 80 százalékát fedi le bármely vizsgált szövegnek, az első 5.000 leggyakoribb szó ismerete Hirsch és Nation (1992) szerint 90-95%-os lefedettséget biztosít, míg Nation (2001) szerint a GSL és az AWL szavainak ismerete a beszélt nyelvi szókincs 92%-át, a magazinok szókincsének pedig 84%-át fedi le. Egy olvasott szöveg megértéséhez a benne elforduló szavak 95-98%-át kell ismerni (Hirsch & Nation, 1992; Laufer, 2005a, 2005b; Nation & Waring, 1997), tehát minden ötvenedik szó lehet ismeretlen csupán (Schmitt, 2008, p. 330). Hazenberg és Hulstijn (1996) az angolt idegen nyelvként tanuló holland egyetemi hallgatók számára határozta meg azt a minimális szókincset, mely lehetővé teszi az angol nyelv egyetemi tanulmányokhoz szükséges anyagok értő olvasását. Eredményeik alapján ez 10.000 angol alapszó ismeretét jelenti. Korábbi hasonló holland tanulmányok szerint 3.000 és 5.000 alapszó passzív ismerete már elegendő lenne. Nation (2006) számításai szerint a British National Corpus alapján a nyelvtanulóknak 8.000-9.000 szócsaládot kell ismerniük ahhoz, hogy autentikus szövegeket tudjanak olvasni, korábban Laufer (1992) 5.000 szócsalád ismeretét javasolta.

A szóismeret mérésének legfontosabb kérdései, hogy miért mérjük, mely szavak ismeretét mérjük, a szóismeret melyik aspektusát akarjuk mérni, és hogyan mérjük azt. A szókincs mérésére kvantitatív és kvalitatív eszközök is rendelkezésre állnak. A kvalitatív megközelítés lehetőséget nyújt a szóismeret mélységének pontos meghatározására, hátránya azonban, hogy időigényes és ezért korlátozott mértékben alkalmazható. A kvantitatív megközelítés lehetővé teszi, hogy rövid idő alatt

nagyszámú szó ismeretér l alkossunk képet, de nem tudjuk a szóismeret teljes spektrumát vizsgálni. Veszteségek tehát mindkét módszer alkalmazása során fellépnek, ezért mindig az adott kontextus függvénye az, hogy melyik a célravezetőbb. Kvalitatív mérőeszközök a személyes interjúk (Verhallen & Schoonen, 1993, in Read, 2000, p. 94), a szóasszociáción alapuló tesztek (Read, 2000), illetve a széles körben alkalmazott *Vocabulary Knowledge Scale* (saját fordításomban *szóismereti skála*, Paribakht és Wesche, 1997). A kvantitatív mérőeszközök között a receptív szókincs nagyságát méri Goulden, Nation és Read (1990) tesztje, míg az elterjedt *Vocabulary Levels Test* (VLT, Schmitt, Schmitt & Clapham, 2001) öt szógyakorisági szinten határozza meg a receptív szókincs nagyságát.

A produktív szókincs meghatározásakor kvantitatív módon közelítve egy szöveghez megállapítható, hány szóból áll összesen, ez a 'szövegszó' ('token') érték, míg a szövegben elforduló különböző szavak számát 'szóalak' értéknek (type) nevezzük. A kettő hányadosát (T/t) gyakran alkalmazzák a szöveg lexikai változatosságának leírására. Mivel azonban e hányados nagymértékben függ a szövegek hosszától, nem tekinthet elég megbízhatónak, ezért helyette számos matematikai formulát dolgoztak ki e hatás kiküszöbölésére (Guiraud R-, Herdan C-, Zipf Z-, vagy Uber U-index). Megoldásként Laufer és Nation (1995) az írás lexikai minőségének vizsgálatára új, objektívebb mérőeszközt alkotott *Lexical Frequency Profile* (LFP) (Szógyakorisági Profil- saját fordításomban) néven. Elméletük szerint a nyelvtanuló szókincsének nagysága jól tükrözi az írásképességében, a szógyakorisági profil pedig megmutatja, milyen arányban fordulnak el az írott szövegben a különböző szógyakorisági szinteknek megfelelő szavak (Laufer és Nation, 1995, p. 311).

Az LFP legújabb (Paul Nation honlapjáról szabadon letölthető) változata a RANGE nevű szoftver, melynek segítségével könnyen elvégezhető az íráselemzés. A program továbbfejlesztett változata, és sok egyéb hasznos alkalmazás található *Compleat Lexical Tutor* néven Tom Cobb (2000) honlapján is. Mivel azonban jelenleg nem létezik olyan eszköz, amely a szó ismeretének receptív és produktív aspektusát egyetlen tesztformátumban lenne képes mérni, ezért Laufer és munkatársai (2004) több kvantitatív tesztípus egyidejű alkalmazását javasolják.

A kutatás ismertetése és az értekezés felépítése

Vizsgálataim megtervezése során a fent említett kutatási eredményekre támaszkodtam. Kutatási kérdéseim megválaszolása érdekében az értekezés két fő részre és hét fejezetre tagolódik, amint az 1. számú táblázatban látható. Az első rész három fejezete átfogó képet nyújt a második részben bemutatott vizsgálatok elméleti háttéréről és a szókinckutatás három alapvető kérdéséről járja körül: a szókinccs leírását, elsajátítását és mérését. Az első fejezet betekintést nyújt az olvasó számára a szókinckutatás terminológiájába, meghatározza a szókinccs nyelvi kompetenciában betöltött helyét, tárgyalja a szó és a szóismeret értelmezési lehetőségeit, és bemutatja a mentális lexikon felépítésének, a szavak tárolásának és előhívásának főbb modelljeit. A második fejezet a szókinccs elsajátítás szakirodalmát tekinti át, érinti az első- és második nyelvi szókinccs elsajátítási mechanizmusának hasonlóságait és különbségeit, és szót ejt a sokat vitatott tudatos szótanulás és az olvasás során fellépő járulékos szóelsajátítás lehetséges előnyeiről és hátrányairól. A harmadik fejezet fókuszában a szókinccs mérése áll. A fejezet bemutatja a szókinccs értékelésének elméleti háttérét, módszereit és a jelenleg elérhető legelterjedtebb, illetve az empirikus vizsgálatok során alkalmazott szókinccsméréseszközöket.

Az értekezés első részében áttekintett kérdések kiindulópontként szolgáltak a második részben tárgyalt empirikus vizsgálatokhoz, melyeket a Pécsi Tudományegyetem Angol Alkalmazott Nyelvészeti Tanszékén folytattam. A kutatás célja az első éves angol szakos hallgatók szókinccsének vizsgálata annak fényében, milyen lexikai követelményeket támaszt számukra a tanterv, kötelező olvasmányaik és írásfeladataik. A negyedik fejezet ismerteti az empirikus vizsgálatok kontextusát, bemutatja a résztvevőket, az angol szakos hallgatók tantervét és a tanszéken folytatott szókinccsmérések előzményeit és háttérét, valamint rövid módszertani áttekintést nyújt a vizsgálatokról. Az ezt követő empirikus fejezetek felépítése az első éves angol szakos hallgatók szókinccsének feltérképezésére és értékelésére kidolgozott szókinccsméréseszköz kifejlesztésének fázisait tükrözi.

1. sz. táblázat: Az értekezés felépítése

I. A szókincskutatás a második nyelv elsajátítás területén

1. fejezet. Mi a szókincs?
 - A szókincskutatás terminológiája
 - A szókincs és a nyelvi kompetencia
 - A szóismeret összetevői
 - A mentális lexikon
2. fejezet: Hogyan sajátítható el a szókincs?
 - A szókincs elsajátítás menetét befolyásoló tényezők
 - A szókincs elsajátítás mélységi és mennyiségi aspektusai
 - A szókincs elsajátítás és az olvasás kapcsolata
 - A tudatos szótanulás és pedagógiai vonatkozásai
3. fejezet: Hogyan mérhető a szókincs?
 - A szóismeret globális mérésének lehetőségei
 - Kvalitatív mérési eszközök
 - Korpusz-alapú szólisták
 - Kvantitatív mérési eszközök
 - Az értekezésben alkalmazott mérési eszközök bemutatása

II. Empirikus tanulmányok az angol szakos hallgatók szókincsének és lexikális szükségleteinek területén

4. fejezet: Háttér a kutatásokhoz
 - Kutatás kontextusa
 - Résztevők ismertetése
 - Kutatás módszertani ismertetése
5. fejezet: Hány szót? Előzetes tanulmányok az angol szakosok receptív szókincsének feltárására
 - A. Az eljárások
 - B. Eredmények
 - C. Az eredmények tárgyalása
6. fejezet: Milyen jól? Szókincsmérések első éves angol szakos hallgatókkal
 - Az eljárások
 - Eredmények
 - Az eredmények tárgyalása
7. fejezet: Mely szavakat? Korpusz-alapú szókincsmérési eszköz kifejlesztése
 - A korpusz bemutatása
 - A mérési eszköz kifejlesztése
 - Mérési eredmények
 - Az eredmények tárgyalása

Konklúzió és jövőbeni kutatási irányok

- A tanulmányok összefoglalása
 - A vizsgálatok korlátai
 - Korpusznyelvészeti, pedagógiai és nyelvi értékelési vonatkozások
-

Az ötödik fejezetben három olyan el tanulmányról esik szó, melyek fontos információval szolgáltak a kutatás további menetét illetően. Az első tanulmány 93 első - és másodéves angol szakos hallgató receptív szókincsének nagyságát mérte fel annak felderítésére, hogy a nemzetközi szakirodalom alapján a hallgatók szókincsének nagysága alkalmassá teszi-e őket egyetemi tanulmányok folytatására angol nyelven. A kapott eredményeket ezután összevettem az egyetemi végzettség angol anyanyelvűek receptív szókincsének nagyságára, valamint az angolt, mint idegen nyelvet tanuló egyetemi hallgatók receptív szókincsének nagyságára vonatkozó nemzetközi adatokkal. Magyarországon ilyen jellegű mérést korábban nem végeztek. Emellett, a további két, kis mintán végzett vizsgálat elzetes betekintést nyújt a receptív szókincs és az írott szóhasználat, valamint a tudatos szótanulás és az olvasás során fellépő járulékos szóelsajátítás kapcsolatába, melyek tanulságait a következő fejezetekben tárgyalt vizsgálatok hasznosították.

A legfrissebb kutatási eredmények gyakorlatban történő alkalmazása mellett a tesztkészítés egyik fontos jellemzője kell, hogy legyen a változó szükségletekhez és feltételekhez való folyamatos alkalmazkodás és a javító célzatú fejlesztés. Ezt az irányelvet szem előtt tartva, a hatodik fejezet fő célkitűzése annak bemutatása, milyen közbülső fázisokon áthaladva alakítottam ki a hallgatók receptív szókincsének mérési eljárását a hallgatók valós lexikális igényeihez igazítva. Az első lépésben a bevezetőben említett, mások által gyakran használt és elismert szókincsmérő teszt, az úgynevezett *Vocabulary Levels Test* (VLT, Schmitt, Schmitt, & Clapham, 2001) akadémikus szintjét használtam, azt feltételezve, hogy ezek a szavak ismerete meghatározó lehet az angol szakosok tanulmányai során.

Mivel a nyelvi alapvizsgára csak évente két alkalommal kerülhet sor, és a januári vizsgán statisztikai elemzés szempontjából kevés jelentkező áll rendelkezésre, így a vizsgálat folytatására évente egy-egy alkalommal volt lehetőség. A második fázisban ugyanennek a tesztnek az angol nyelvben ritkán előforduló szavak ismeretét mérő szekcióját alkalmaztam, feltételezve, hogy ez a szint jobban képes elkülöníteni a kisebb és nagyobb szókincssel rendelkező hallgatókat, tehát jobban megfelel a nyelvi alapvizsga szűrő funkciójának.

A tanulságok levonása után a harmadik fázisban egy saját kidolgozású mér eszközt használtam, melynek szavai részben a VLT akadémikus szintjét alkotják, részben pedig olyan itemek, melyek saját tapasztalataim és intuitív megítélésem szerint tanulmányaik során fontosak és hasznosak az angol szakos hallgatók számára. Emellett, a szókinccsmér eszközzel nyert receptív szókincre vonatkozó adatokat ebben a fázisban összevettem a nyelvi alapvizsga egyéb komponenseiből nyert adatokkal, azzal a céllal, hogy behatóbban vizsgáljam a szókinccs és az olvasás-, írás-, és hallott szövegértési képesség, az általános nyelvi szint, valamint az angol szakos tanulmányi eredmények összefüggéseit, illetve hogy pontosabb képet kapjak az egyik el tanulóban már vizsgált receptív és produktív szókinccs kapcsolatáról. A vizsgálat során tehát a hallgatók írásainak lexikális megíteléséhez a korpuszelemzés módszereit is alkalmaztam.

A fent bemutatott eljárás a szavak gyakoriságának illetve hasznosságának intuitív megítelésén alapul, ám ez nem feltétlenül esik egybe a szavak valódi gyakoriságával (McCrostie, 2007). Ezért a hetedik fejezet azt vizsgálja, milyen szókincre van valójában szüksége a Pécsi Tudományegyetem angol szakos hallgatóinak ahhoz, hogy sikeresen végezhesék tanulmányaikat angol nyelven. Ehhez a hallgatók számára kurzusaikon kötelező olvasmányokból összeállított reprezentatív szövegmintán végeztem korpuszelemzést. A korpuszt CORES-nak neveztem el (Corpus of Readings in English Studies). A korpusz elemzésével nyert adatok azután forrásként szolgáltak egy olyan szókinccsmér eszköz kidolgozásához, mely jobban tükrözi az angol szakos hallgatók valós lexikai igényeit, mint a nemzetközileg kidolgozott és korábban alkalmazott szókinccsmér eszközök.

Az értekezésben bemutatott vizsgálatok kutatási kérdéseit, adatgyűjtési eszközeit és elemzési módszereit a 2. számú táblázat foglalja össze. A továbbiakban a kutatás négy fázisának eredményeit mutatom be.

2 sz. táblázat: A kutatás fázisai, kutatási kérdései, adatgyjt eszközei, és módszertana

	Kutatási kérdések	Adatgyjt eszközök	Elemzés módszere
Els el tanulmány 93 résztvev	Hány angol szót ismernek az els éves angol szakos hallgatók? Vagyis, mekkora az angol szókincsük és elég nagy-e ahhoz, hogy egyetemi tanulmányokat folytassanak angol nyelven?	A szókincs nagyságát mér teszt (Goulden, Nation & Read, 1990)	Leíró statisztika
Második el tanulmány 33 résztvev	Milyen összefüggés áll fenn a passzív szókincs nagysága és a hallgatók írásában megjelen aktív szókincs között? A nagyobb passzív szókincssel rendelkező hallgatók írásában változatosabb és kifinomultabb aktív szóhasználat jelenik-e meg, és fordítva?	A szókincs nagyságát mér teszt (Goulden, Nation & Read, 1990) A hallgatók írásai	A hallgatók írásainak korpuszelemzése Korreláció analízis
Harmadik el tanulmány 33 résztvev	Hatékonyabb-e hosszú távon a tudatos szótanulás és a szótesztekre való rendszeres felkészülés az olvasás során fellép járulékos szóelsajátításnál?	A szókincs nagyságát mér teszt (Goulden, Nation & Read, 1990) Szókincsfelj dést mér teszt Kvalitatív kérd ív	Korreláció analízis
Els fázis 220 résztvev	Milyen jól ismerik az els éves hallgatók az egyetemi szint szövegek ért olvasásához elengedhetetlen (akadémikus) szókincset? Hogyan diszkriminál a hallgatók között a Schmitt, Schmitt & Clapham (2001) által kidolgozott szóteszt akadémikus szókincset mér része?	Szókincsmér teszt (VLT, Schmitt, Schmitt & Clapham, 2001) Szókincsmér teszt (VLT, Schmitt, Schmitt & Clapham, 2001)	Leíró statisztika Korreláció analízis Item analízis
Második fázis 211 résztvev	Az akadémikus vagy a ritka szavak receptív ismeretének mérése alkalmasabb-e a nem megfelel szókincssel rendelkező hallgatók sz résére? Melyik diszkriminál jobban?	Szókincsmér teszt (VLT, Schmitt, Schmitt & Clapham, 2001)	Korreláció analízis Item analízis
Harmadik fázis 135 résztvev	Milyen kapcsolat áll fenn a passzív illetve az aktív szókincs és az írott- és hallott szövegértést, az íráskészséget, és a nyelvhelyességet mér teszteken elért eredmények között? Milyen kapcsolat mutatható ki a passzív és az aktív szókincs illetve az általános nyelvi szint között? Milyen összefüggés áll fenn a passzív és az aktív szókincs között?	Nyelvtudásmér teszt Az általam angol szakosok számára hasznosnak vélt szavakból kidolgozott szókincsmér eszköz Nyelvtudásmér teszt Az általam angol szakosok számára hasznosnak vélt szavakból kidolgozott szókincsmér eszköz Nyelvtudásmér teszt Az általam angol szakosok számára hasznosnak vélt szavakból kidolgozott szókincsmér eszköz	A hallgatók írásainak korpuszelemzése Korreláció analízis Korreláció analízis Lineáris regresszió analízis

	Kutatási kérdések	Adatgyűjtési eszközök	Elemzési módszerek
	A tanulmányban vizsgált 9 változó milyen összefüggést mutat az angol szakosok tanulmányi eredményeivel?	Nyelvtudásmérési teszt Az általam angol szakosok számára hasznosnak vélt szavakból kidolgozott szókinccsmérési eszköz	Lineáris regresszióanalízis
Negyedik fázis 134 résztvevő	Mely szavak ismeretére lehet szüksége egy angol szakos hallgatónak a PTE-n, hogy képes legyen megérteni a kötelező olvasmányként feladott tanulmányokat?	Az általam összeállított CORES korpusz	Korpuszelemzés Szókinccsprofil
	Korpuszelemzés alapján milyen mértékben segítheti a Coxhead (2000) által összeállított akadémikus szólista ismerete az angol szakosok olvasott szövegértését?	Az általam összeállított CORES korpusz	Korpuszelemzés Szókinccsprofil
	Milyen ritka szavakkal találkozhatnak az angol szakos hallgatók kötelező olvasmányaikban?	Az általam összeállított CORES korpusz	Korpuszelemzés Szókinccsprofil
	Mennyire ismerik a hallgatók a kötelező olvasmányaik korpuszában leggyakrabban előforduló akadémikus és ritka szavakat?	CORES korpusz alapú szókinccsmérési eszköz	Leíró statisztika Item analízis

A kutatás eredményei

Az előző tanulmányok fontos adatokkal szolgáltak a további kutatás megtervezéséhez, ezért fontos ezekről is szót ejteni. Az első előző tanulmány azt vizsgálta, mekkora az első éves, nyelvi alapvizsga előtt álló hallgatók szókinccsének nagysága és ez elegendő-e ahhoz, hogy angol nyelven folytathassanak egyetemi tanulmányokat. A mérés eredménye szerint nagyok az egyéni eltérések, de szókinccsük nagysága 4.000-11.000 szó között mozog, az átlag 7.123,66 (SD= 1777,97). Amennyiben azt vesszük alapul, hogy minimálisan 5.000 szó receptív ismerete szükséges az írott autentikus szövegek megértéséhez, akkor hallgatóink 15 százalékának kell nehézségekkel szembenéznie az olvasás során. Ha viszont a holland kutatások által javasolt 10.000 szó ismeretét tekintjük szükségesnek, akkor a hallgatók 94 százaléka nem éri el a megfelelő szintet. Még a Nation (2006) által javasolt, mértékletesebb, 8.000 szóra vonatkozó adatok szerint is hallgatóink közel 60 százalékának kell lexikális nehézségekkel számolnia tanulmányai során.

Mivel a hallgatóknak a kurzusok teljesítése során írásfeladatokat is teljesítenie kell, a második tanulmányban igyekeztem elzáró képet kapni arról is, milyen kapcsolat állhat fenn a receptív szókincs nagysága és a produktív szóhasználat között. Bár a vizsgálat nem talált szignifikáns összefüggést a két tényező között, a szókincsprofil kimutatta, hogy a nagyobb szókinccsel rendelkező hallgatók írásaikban több nem gyakori szót használtak, mint kisebb szókincsű társaik. Ez összhangban áll Laufer és munkatársainak (2004) eredményeivel, ám egyben arra is utal, hogy az íráskészségben nem csupán a szókincs játszik szerepet, ezért további vizsgálatokra van szükség ezen a téren.

A harmadik eltanulmány a szódolgozatokra való rendszeres és tudatos felkészülés és az olvasás során fellépő járulékos szóelsajátítás hatékonyságát vizsgálta. Az eltesztelés során a teszt- és a kontrollcsoport szókincsének nagysága hasonló volt. Mindkét csoport ugyanazokat az olvasmányokat olvasta a félév során azzal a különbséggel, hogy a tesztcsoport az órákon szókincsre fókuszált feladatokat végzett és rendszeres szódolgozatokat írt, míg a kontrollcsoport csupán az olvasmányok tartalmi vonatkozásaira koncentrált. A 14 héten át tartó kezelést követően utóteszten nem volt szignifikáns különbség a két csoport eredménye között az olvasmányok szavainak kvantitatív és kvalitatív módon is mért ismerete tekintetében. Ez arra utalhat, hogy az olvasás során is sajátíthatunk el szavakat és a hallgatók kérdésben adott válaszaiból az is kiderült, hogy a szódolgozatokat hasznosnak és fontosnak tartják ugyan, de a megtanult szavak többsége nem rögzült bennük hosszú távon. Fontos következtetés tehát, hogy a szavak tartós rögzüléséhez többször és többféleképpen kell találkozni az adott szóval. Ezt a nyelvfejlesztés során lehetővé kell tenni, különösen a nem gyakori szavak esetében.

Láttuk tehát, hogy az első éves angol szakos hallgatók többségének szókincsa a nemzetközi adatok tükrében nem elég nagy ahhoz, hogy a tanuló probléma nélkül képes legyen autentikus írott szövegek megértésére, továbbá azt is megfigyelhettük, hogy a receptív szókincs nagysága összefüggésben van az írott szóhasználattal is. A hallgatóknak kurzusaik során pedig szükséges egyetemi szintű szakirodalmat olvasni és írásfeladatokat készíteni angol nyelven, ebben a körben tehát a szókincsbeli hiányosságok a tanulmányi sikerekre is hatással lehetnek. További kutatásaim

fókuszában ezért az állt, mely szavak ismerete lehet hasznára a hallgatóknak és ezeket milyen jól ismerik.

Korábbi kutatások igazolták, hogy magasabb nyelvi szinteken a nyelvtanuló szókincsér l pontosabb képet alkothatunk receptív szókincsének nagysága és produktív szóhasználata alapján, mint kvalitatív vizsgálatokkal (Zareva, 2005). Ezt figyelembe véve, valamint a praktikalitás, objektivitás, az eredmények számítógépes értékelhet sége, és a mérés megvalósítása lehetséges szempontjainak mérlegelésével, a hallgatók szókincsét a nyelvi alapvizsga nyelvhasználatot mér komponensének receptív szókincsét mér tesztrel való kiegészítése mellett döntöttem. Erre a célra el ször a VLT (Schmitt, Schmitt & Clapham, 2001) akadémikus szókincsét mér itemeit választottam ki. A vizsgálat els fázisában megállapítottam, hogy a hallgatók a tanulmányaik szempontjából hasznosnak tekinthet szavakat többségében ismerik, és a teszt nem diszkriminál megfelel en. Ugyanakkor arra is fény derült, hogy a receptív szókincs és a nyelvtani teszt eredményei között szignifikáns összefüggés áll fenn, ami azt jelentheti, hogy a két konstruktum nem független egymástól, ahogyan arra Wray (2002) is utal. Az enyhe korreláció miatt azonban az is nyilvánvaló, hogy a két teszt eltér dolgot mér, szándékaink szerint.

A következ fázisban, egy évvel kés bb, arra kerestem a választ, hogy a nem gyakori szavak ismerete jobban diszkriminál-e a kisebb illetve nagyobb receptív szókincs hallgatók között az akadémikus szókincsét mér tesztnél. Az eredmények igazolták e feltevést, ugyanakkor felvet dött a kérdés, mennyire tekinthet k hasznosnak a VLT tízezres gyakorisági szintjének szavai a mi hallgatóink számára. Következésképpen, felmerül az a kérdés, hogy valóban ezeket a szavakat kell-e szókincsmérés céljára alkalmaznunk, vagy talán létezik-e ennél jobb megoldás is az adott kontextusban.

A megoldást az jelentette, hogy az els fázisban itemelemzés alapján megfelel en diszkrimináló néhány akadémikus item megtartása mellett olyan új itemeket dolgoztam ki, melyek intuitív megítélésem szerint jobban tükrözik az angol szakos hallgatók lexikális igényeit. A statisztikai adatok azt mutatták, hogy az új itemek többsége megfelel en diszkriminál (átlag $DI = 0,41$ és $0,3$ a két tesztváltozaton), néhány azonban revízióra szorul majd. Vizsgálataim ebben a fázisban

kiterjedtek a hallgatók írásfeladataiban megfigyelhető produktív szókincsre is. Az adatok arra utalnak, hogy a produktív szókincs mérésére alkalmazott három mérőszám közül a T/t arány és az U index nem tökéletesen alkalmas, míg a szókincsprofil megfelelőbbnek bizonyult. A receptív és produktív szókincs összevetésében a Pearson korrelációs táblázatok alapján az látszik, hogy a várt erős korreláció helyett csupán a szókincsprofil alapján meghatározott, a szövegben ritkának tekinthető szavak aránya (a GSL kétezer szaván kívüli szavak használatának mértéke = 'Beyond 2.000') mutat gyenge, de szignifikáns kapcsolatot a receptív szókinccsel. Ennek egyik magyarázata az lehet, hogy a receptív és a produktív szókincs fejlődése nem párhuzamosan zajlik, ahogy arra Laufer (1998) is rámutatott, illetve, hogy a hallgatók receptív szókincsének nagysága jóval nagyobb, mint amit aktívan használni képesek az írás folyamán. Másrészt, Jarvis (2002) szerint az írás lexikai változatosságát csupán a nyelvelsajátítás kezdeti szakaszaiban befolyásolja a passzív szókincs nagysága, ami hallgatóink pedig magasabb szintű nyelvtudással rendelkeznek. Emellett valószínűsíthető az is, hogy az írásokban megfigyelhető lexikai változatosság sokkal összetettebb jelenség, mint a benne előforduló szavak száma és gyakorisága, és nem írható le csupán matematikai összefüggésekkel.

Kutatási kérdésként megfogalmaztam azt is, milyen összefüggés lehet a receptív illetve a produktív szókincs, és a nyelvi alapvizsga tesztkomponensein (olvasott és hallott szövegértés, írás, nyelvtan és nyelvhasználat) elért eredmények között. A receptív szókincs szignifikáns korrelációt mutat mind a négy komponenssel, míg a produktív szókincs csupán a nyelvtani teszt eredményeivel mutat összefüggést. Ez az eredmény a receptív szókincs tekintetében nem meglepő, hiszen az olvasott és hallott szövegértés szintén receptív készségeket feltételez. Az, hogy a produktív szókincs mérőszámai esetünkben nem mutatnak összefüggést az írásfeladaton elért pontszámmal, abból adódhat, hogy az írás megítélése során az értékelők nem csupán a szókinccset vették figyelembe, hanem fontos szerepet játszott a kohézió, koherencia, a nyelvhelyesség, valamint a feladatban meghatározott követelmények teljesítésének mértéke is.

Vizsgáltam továbbá a teljes alapvizsgán elért eredmény (mely a nyelvtudás szintjét jellemzi) és a szókincs kapcsolatát is. Korreláció és regresszió analízis alapján

megállapítható, hogy a receptív szókincs és a nyelvtudás szintje nem független egymástól, míg a fent említett mér számokkal kifejezett produktív szókincs nem mutat összefüggést. Végül, kiszámítottam a résztvevők első évben elért tanulmányi eredményeinek átlagát. Ez azt mutatja, mennyire sikeres az adott hallgató tanulmányai során. Pearson korreláció analízis alkalmazásával ezeket az értékeket összevettem a vizsgálatban szereplő kilenc változóval (receptív szókincs, a produktív szókincs három mér száma, az olvasott- és hallott szövegértés, írás, nyelvtan és nyelvhasználat tesztkomponenseken, valamint a nyelvi alapvizsga egészén elért pontszámok).

A két tesztváltozat közül az elsőben gyenge, de szignifikáns összefüggés mutatható ki a tanulmányi eredmények és a hallott szövegértés, valamint az íráskomponensen elért eredmények között, míg a második változat nem mutat összefüggést. A jelenség egyik magyarázata lehet, hogy a hallgatók a kreditrendszernek megfelelően különféle kurzusokat teljesítenek, melyeken az oktatók nem egységes értékelési szempontokat alkalmaznak. Ugyanakkor a kurzusok teljesítésének kritériuma többnyire egy esszé feladat elkészítése, várható tehát, hogy aki jól ír, sikeresebb is lehet a tanulmányai során. Érdekes aspektus emellett a hallott szövegértés összefüggése is, hiszen a tanulmányok során fontos a tanárok, előadók beszédének megértése. Nagy (2008) kimutatta, hogy a PTE angol szakos hallgatói a tanórákon nem szívesen kommunikálnak, véleményem és adataim szerint talán ezért a magyar nyelvben az egyetemistákra használatos „hallgató” elnevezés.

Összefoglalva, a harmadik fázisban kidolgozott és alkalmazott szókinccsmér eszköz alkalmazása hatékonyabbnak bizonyult a korábbi, nemzetközileg használatos eszköznél a mi kontextusunkban, és az idegen nyelvi szókincs számos fontos aspektusára világított rá. Ugyanakkor, a szavak gyakoriságának, vagy hasznosságának intuitív megítélése megkérdőjelezhető (McCrostie, 2007), melyet a teszt szókincsprofilja is bizonyított. A következő lépésben szükségessé vált tehát annak pontos meghatározása, milyen valós lexikai igényei vannak a PTE angol szakos hallgatóinak.

Ennek megválaszolására egy, a hallgatók számára kötelező olvasmányokból álló reprezentatív korpusz (CORES) létrehozására volt szükség. A CORES lexikai elemzését szógyakorisági profiljának elkészítésével végeztem. A korpuszt három

alkorpusz alkotja (nyelvészeti, alkalmazott nyelvészeti, és irodalom/kultúra szövegek) és összesen 118.808 szóból áll, leszámítva a tulajdonneveket és számokat. Ez az általános korpuszokhoz képest kicsinek tekinthető, ám specializált korpuszként alkalmas az adott diszciplína szóhasználatának feltérképezésére. Bár a lexikai változatosságot jellemző T/t érték nem magas (0,10), a korpusz lexikai sűrűsége azt mutatja, hogy minden három szóból kettő különböző, s ezzel meg lehet a lexikai kihívás elé állítani az olvasókat. A szókincsprofil alapján látszik továbbá, hogy a teljes szöveg szókészletének 77 %-át, vagyis több mint kétharmadát, a GSL első 2.000 szava alkotja, igazolható tehát a feltevés, miszerint ezek a szavak ismerete jelentős szerepet játszik a szöveg megértésében. Korábbi vizsgálatok szerint ez az érték általában 80% körül mozog (O'Keefe és munkatársai, 2007). További 7,54%-ot tesz ki az AWL szólistában megtalálható szavak aránya, melyek ismerete tovább növelheti a hallgatók szövegértési képességét. Összeadva mindez a teljes korpusz szókészletének közel 85%-át jelenti, amely azonban még nem elegendő arány, figyelembe véve, hogy a szavak 95-98%-ának ismerete szükséges a szöveg megértéséhez (Hirsch & Nation, 1992; Laufer, 2005a, 2005b; Nation & Waring, 1997). Ebből kifolyólag, a fennmaradó 15 százalék arra utal, hogy a hallgatóknak ismerniük kell emellett olyan, egyébként a korábban említett kritériumok alapján ritkának nevezhető szavakat is, melyek viszonylag gyakran fordulnak el ezen a tudományterületen. Az elemzés ezen eredményeit az értekezés melléklete tartalmazza.

A CORES korpusz elemzése alapján a tesztfejlesztés negyedik fázisában tehát olyan szókincsmérő eszközt dolgoztam ki, mely formátumában megegyezik a korábbiakkal, ám itemei a korpuszban előforduló szavak gyakorisága alapján kerültek kiválasztásra. A 134 résztvevő vizsgálat eredményei szerint a harminc itemes teszten az átlag változatonként 18,314 és 16,833 ($SD_1=4,686$ és $SD_2= 3,918$), az itemek többsége pedig kielégítően diszkriminál ($DI_1= 0,33$, $SD_1= 0,14$; és $DI_2= 0,31$, $SD_2 = 0,18$). Az eredmények biztatóak, ám a tesztfejlesztési folyamat ezzel még nem zárult le. Fontos lesz a továbbiakban a kevésbé jól működő itemek revíziója, újabb itemek, esetleg újabb tesztváltozatok készítése, valamint egy nagyobb mintán végzett, gondos validálási eljárás lefolytatása, mielőtt az eszköz széles körben is alkalmazható lesz a jövőben.

A kutatás korlátai

A kutatás eredményeinek bemutatása mellett szólni kell a korlátairól is. Az értekezésben bemutatott vizsgálatok fókuszában az írott szókinccs állt. Mivel a kutatás elsődleges célja az volt, hogy az első éves hallgatók számára mérőföldként elírt nyelvi alapvizsga részeként olyan szókinccsmérő eszközt fejlesszen ki, mely alkalmas lehet a diszciplína valós lexikai igényeinek figyelembe vétele alapján a nem megfelelő szókinccsel rendelkező hallgatók számára, vizsgálataim a hallgatók írott szóhasználatára irányultak. Ennek oka az, hogy az alapvizsgának akkor még nem volt szóbeli része, ám a jövőben hasznos lehet a beszédprodukción hasonló vizsgálata is a bolognai alapképzés (BA) és a mesterképzés (MA) kapcsán.

A vizsgálatokba bevont hallgatók száma a kutatáshoz adott volt, hiszen minden első éves számára kötelező a nyelvi alapvizsga, így a minta a kérdéses kontextusra reprezentatívnek tekinthető. Mivel azonban a statisztikai elemzésekben a résztvevők számának növekedésével bizonyítottan jobb megbízhatóság érhető el (Szabó, 2008), így a jövőben a dolgozatban bemutatott eredmények alapján átdolgozott korpusz-alapú szókinccsmérő eszközt nagyobb populáción szükséges kipróbálni, mivel a szélesebb közönség számára is elérhetővé tehető.

A hallgatók írásainak elemzése, illetve a CORES korpusz szókinccsének vizsgálata során alkalmazott szógyakorisági profil képes megállapítani, melyek az adott szövegekben leggyakrabban előforduló szavak. Ugyanakkor a profil nem alkalmas a több szóból álló kifejezések vagy idiómák vizsgálatára, melyek előfordulása a szövegben Muncie (2002) szerint éppen a választékos nyelvhasználat egyik bizonyítéka. A szóismeret kvalitatív aspektusának vizsgálatát az értekezésben az a megfontolás írta felül, hogy bár a kvalitatív megközelítés során a szóismeret mélységének pontosabb meghatározására van lehetőség, jelentősen kisebb mértékben alkalmazható csupán, mint amire esetünkben szükség volt. Mivel jelenleg nem létezik olyan szókinccsmérő eszköz, mely a szóismeret minőségi és mennyiségi jellemzőit egyetlen tesztformátumban lenne képes mérni (ennek egy lehetséges megoldásaként tettem kísérletet a harmadik előtanulmányban), Laufer és munkatársai (2004) a

szókincs mérését több kvantitatív mérő eszköz együttes alkalmazásával javasolják. Az általuk említett feltételeknek az értekezésben bemutatott kutatás eleget tesz, hiszen a nyelvi alapvizsga tartalmaz egy passzív szókincset mérő, korpusz alapú, valós lexikai igényelemzéssel kidolgozott eszközt; a hallgatók írásainak lexikális elemzése szógyakorisági profil készítésével történt; a hallott és írott szövegértést mérő tesztkomponenseken pedig jelentős szerepet kap a szókincs és a szövegkörnyezet is. E kombináció együttesen átfogóbb képet nyújt a hallgatók lexikális ismereteiről, mint egyetlen mérési módszer alkalmazása.

A mérő eszköz kifejlesztése során lehetséges problémaforrásként merül fel az előtesztelés során nyert információ korlátozott volta. A mérő eszközök előtesztelése igen fontos, ugyanakkor ellentmondásos fázisa a tesztfejlesztésnek (Szabó, 2008; Alderson, Clapham, & Wall, 1995; Hughes, 1991). Az értekezésben tárgyalt kutatás során az előtesztelés azért okozott nehézséget, mert ideális esetben az előtesztelt csoport nagysága és tulajdonságai a lehető legnagyobb egyezést mutatnak a célcsoport megfelelő paramétereivel. Sajnos, mivel a mérő eszközöket csupán saját kollégáimon volt alkalmam előtesztelni - mert biztosítani kell az alapvizsga titkosságát -, nem állíthatom, hogy a tesztfejlesztés során ennek a kritériumnak tökéletesen eleget tettem. Azonban az előtesztelés során gyengének vagy félreértelmezhetőnek bizonyuló itemeket cseréltem és újabb előtesztelésnek vettem alá. Sőt, ilyen esetben az egész klaszter cseréjére sor került, mert azon belül az itemek bizonyítottan nem függetlenek egymástól (Schmitt, Schmitt & Clapham, 2001). A CORES-korpusz alapú mérő eszköz jelenleg tehát nem tekinthető még tökéletesen megbízhatónak, és gondos validálási eljárásnak szükséges a jövőben alávétetni nagyobb mintán, mielőtt más magyarországi egyetemeken angol programjai számára is ajánlható lenne.

Végül, a jövőben a korpusz-alapú szókincsmérő eszköz revíziója és további itemek írása során egy lehetséges újabb szempontként merül fel a szavak korpuszbeli gyakorisága mellett a tesztelésre kiválasztandó szavak kulcsszó státusza. A kulcsszavak Scott (1999) szerint nem feltétlenül a leggyakoribb szavak a szövegben, hanem a „szokatlanul” gyakoriak egy adott normához képest. Egyes szoftverek, például a Wordsmith Tools, mely vizsgálatom során még nem állt rendelkezésemre, nagy referencia korpuszokhoz hasonlítva képesek meghatározni és listázni azokat a

szavakat, amelyek a beadott szövegben szokatlanul gyakran fordulnak el, vagyis kulcsszavai annak. A CORES korpuszra vonatkoztatva, egy ilyen kulcsszó lista segítségével még pontosabb képet kaphatunk arról, mely szavakat ismeretét érdemes mérni az angol szakos hallgatóink esetében.

Összegzés és a kutatás jövőbeli kiterjesztése

Mindemellett, kutatásom eredményei számos tekintetben hozzájárultak az angol szakos hallgatók angol, mint idegen nyelvi szókincsének leírásához, pedagógiájához és értékeléséhez egyaránt, s egyben új irányvonalakat jelöltek meg a további kutatások számára.

Korpusznyelvészeti vonatkozások

- A hallgatók írásaiból összeállított korpusz, mely bár más korpuszokhoz képest kicsinek tekinthető, betekintést nyújthat a magyar anyanyelv, angol szakos egyetemi hallgatók írott nyelvhasználatába és hozzájárulhat az angolt, mint idegen nyelvet tanulók íráskészségének tanulmányozásához. A hallgatói korpuszt a jövőben tovább lehet majd vizsgálni más tanulói korpuszok viszonylatában is, illetve a szókincsen kívül egyéb nyelvi elemekre is kiterjesztve.
- A CORES, mint az angol szakos hallgatók számára kötelezően olvasandó tanulmányokból összeállított szöveggyűjtemény, hozzájárulhat a diszciplínára jellemző nyelvhasználati sajátosságok feltárásához.
- A kötelező olvasmányok korpuszát a jövőben folyamatosan bővíteni kell és idővel elérhetővé lehet tenni más magyarországi felsőoktatási intézmények hallgatói, oktatói és kutatói számára is.

Nyelvpedagógiai vonatkozások

- A tanulói korpuszok hatékonyan alkalmazhatók az írásoktatásban, ahogy azt korábban Horváth (2001) már bizonyította, tehát hasznos lehet a jövőben az

írást oktató kollégák és nyelvtanárok számára hozzáférhetővé tenni a hallgatók írásainak korpuszát.

- A CORES korpusz alapján összeállított gyakorisági- és kulcsszó listák fontos információval szolgálhatnak a nyelvfejlesztéssel foglalkozó kollégáknak arról, mely szavakra fókuszáljanak tudományterületünkön a hallgatók egyetemi szintű, angol nyelvű tanulmányok olvasására való felkészítése során. Eredményeim arra utalnak, hogy az angol nyelvben leggyakoribb és az egyetemi szövegekben leggyakrabban előforduló szavak ismerete mellett jelentősen segítheti a hallgatók olvasott szövegértési képességét azon szavak ismerete, melyek általában az angol nyelvben ritkán, tudományterületünkön azonban viszonylag gyakran fordulnak elő. Ebből adódóan, az első évek nyelvfejlesztésére irányuló kurzusok oktatói sokféle módon hasznosíthatják a CORES korpuszt a tanórán és azon kívül is a Pécsi Tudományegyetemen, de tágabb vonatkozásban más, hasonló angol szakos képzésben országwide.
- A CORES korpusz jövőbeli alkalmazási lehetőségei nagyban elősegíthetik a hallgatók tanulási autonómiáját és a vizsgákra való önálló felkészülést. A korpusz alapján összeállított gyakorisági- és kulcsszó listák, amennyiben a hallgatók szabadon hozzáférhetnek az Anglisztikai Intézet honlapján, a tanórán kívüli tudatos szókincsfejlesztés fontos elemeivé válhatnak.
- A korpuszelemzés egyes eszközeinek alkalmazása, például konkordanciák futtatása a CORES szövegein, lehetőséget nyújt a hallgatók számára arra, hogy azonosítsák, és egyéni megoldást találjanak nyelvi problémákra az írott nyelvhasználat minden szintjén. A korpusz referenciaként szolgálhat a hallgatók számára az írásfolyamat különböző fázisaiban, például azáltal, hogy ellenőrizhetik benne a kérdéses szavak területükre jellemző kollokációit, melyre a korábban kifejlesztett egyszerű gyakorisági szólisták nem adnak lehetőséget. Ezáltal a CORES elősegítheti a választékosabb, helyes szóhasználatot és a tudatos szókincsfejlesztést.
- Jövőbeli terveim között szerepel egy olyan szoftver kifejlesztése a CORES korpuszra, mely további gyakorlási és szókincsfejlesztési lehetőséget jelenthet

az angol szakos hallgatók számára a tanórán és azon kívül egyaránt. A számítógépes program elképzelésem szerint a felhasználó választása szerinti nehézségi- vagy szógyakorisági szinteken C- vagy *cloze* tesztek generál majd a CORES korpusz és/vagy bármely alkorpusza alapján, szintén a felhasználó igényei szerint. Az eszköz azután minden PTE angol szakos hallgatónak, s t esetleg országos érdeklődőknek is, szabadon elérhető lehetne, bár ehhez valószínűleg további anyagi és technikai forrásokat kell majd találni.

Értékelési vonatkozások

- A kvantitatív és kvalitatív szókinccsmérési módszerek egy tesztformátumban történő ötvözéséről ad számot a bemutatott két előtanulmány, mely egyben kísérlet a szakirodalomban ezen téren mutatkozó hiány pótlására.
- Az empirikus vizsgálatok eredményei azt látszanak alátámasztani, hogy a szókinccs és a grammatikai ismeret egymással szorosan összefügg és a nyelvi mérés terén egy konstrukciónak kezelhető.
- A kutatás eredményei fontos információval szolgáltattak a szókinccsmérés területén arról, mely szavak ismeretének mérésére kell hangsúlyt fektetni, ha a mérés célja az angol szakon sikeres tanulmányok folytatásához nem megfelelő receptív szókinccsel rendelkező hallgatók kiszűrése.
- Az angol szakos hallgatók speciális lexikális igényeinek figyelembe vételével a CORES-alapú szókinccsmérési eszköz hatékonyabb, egyedi kontextusra szabott szókinccsmérést tesz lehetővé.

A jövő titka még, hogy a célközönség mennyire tartja majd érdekesnek és hasznosnak a fenti javaslatokat. Én azt remélem, hogy az értekezésben felvetett gondolatok és kérdések hozzájárulnak az angol szakos hallgatók lexikális ismereteinek és igényeinek mélyebb megértéséhez, és pótolják a szókinccskutatás ezen téren mutatkozó hiányt.

Hivatkozások

- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Alderson, J.C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart and Winston.
- Brown, C., Sagers, S.L., & LaPorte, C. (1999). Incidental vocabulary acquisition from oral and written dialogue journals. *Studies in Second Language Acquisition*, 21 (2), 259-283.
- Caramazza, A. (1999). Hány feldolgozási szint van a lexikai hozzáférésben? In Z. Bánréti (Ed.), *Nyelvi struktúra és az agy. Neurolingvisztikai tanulmányok* (pp. 249-286). Budapest: Corvina.
- Caron, J. (1992). *An introduction to psycholinguistics*. New York: Harvester and Wheatsheaf.
- Carter, R. (1998). *Vocabulary. Applied linguistic perspectives*. London: Routledge.
- Cobb, T. (2000). Compleat Lexical Tutor. [site Web] <http://www.lextutor.ca>.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition*, 21 (2), 285-301.
- Gleitman, L., & Landau, B. (Eds). (1994). *The acquisition of the lexicon*. Cambridge, Massachusetts: The MIT Press.
- Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, 11 (4), 341-363.
- Grabe, W., & Stoller, F. L. (1997). Reading and vocabulary development in a second language: A case study. In J. Coady, & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 98-122). Cambridge: Cambridge University Press.
- Hazenbergh, S., & Hulstijn, J. H. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics*, 17 (2), 145-163.
- Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21 (2), 303-317.
- Hirsch, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8 (2), 689-96.
- Horváth, J. (2001). *Advanced writing in English as a foreign language. A corpus-based study of processes and products*. Pécs: Lingua Franca Csoport.
- Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, 21 (2), 181-193.
- Hughes, A. (1991). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Hulstijn, J.H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80, 327-339.
- Jarvis, S. (2002). Short texts, best fitting curves and new measures of lexical diversity.

- Language Testing*, 19 (1), 57-84.
- Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, 14, 185-209.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis. *The Modern Language Journal*, 73, 440-464.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). London: Macmillan.
- Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20-34). Cambridge: Cambridge University Press.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? *Applied Linguistics*, 19 (2), 255-271.
- Laufer, B. (2005a). Focus on form in second language vocabulary learning. *EUROSLA Yearbook*, 5, 223-250.
- Laufer, B. (2005b). Instructed second language vocabulary learning: The fault in the 'default hypothesis'. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language learning* (pp. 311-329). Berlin: Walter De Gruyter.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics*, 16 (3), 307-322.
- Laufer, B., Elder, C., Hill, C., & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing*, 21 (2), 202-226.
- Martsa, S. (2007). *English morphology. An introduction*. Budapest: Nemzeti Tankönyvkiadó.
- McCrostie, J. (2007). Investigating the accuracy of teachers' word frequency intuitions. *RELC Journal*, 38(1), 53-66.
- Muncie, J. (2002). Process writing and vocabulary development: comparing lexical frequency profiles across drafts. *System*, 30, 225-235.
- Nagy, B.C. (2008). "To will or not to will" *Exploring advanced EFL learners' willingness to communicate in English*. Unpublished PhD thesis. University of Pécs.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 1-12 .
- Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge: Cambridge University Press.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- O'Keefe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: language use and language teaching*. Cambridge: Cambridge University Press.
- Paribakht, T. S., & Wesche, M. (1999). Reading and "incidental" L2 vocabulary acquisition. An introspective study of lexical inferencing. *Studies in Second*

- Language Acquisition*, 21 (2), 195-224.
- Pawley, A., & Syder, F.H. (1983). Two puzzles for linguistic theory: native-like selection and native-like fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191-226). London: Longman.
- Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: an assessment perspective. *Language Learning*, 52 (3), 513-536.
- Qian, D. D., & Schedl, M. (2004). Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing*, 21 (1), 28-52.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Acquisition*, 21, 589-619.
- Schmitt, N. (2008). Instructed second language vocabulary learning. Review article. *Language Teaching Research*, 12 (3), 329-364.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18 (1), 55-88.
- Scott, M. (1999). *Wordsmith Tools*. Software. Oxford: Oxford University Press.
- Singleton, D. (1999). *Exploring the second language mental lexicon*. Cambridge, Cambridge University Press.
- Singleton, D. (2000). *Language and the lexicon. An introduction*. London: Arnold.
- Szabó, G. (2008). *Applying item response theory in language test item bank building*. Frankfurt am Main: Peter Lang.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28 (1), 46-65.
- Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, 30 (1), 79-95.
- West, M. (1953). *A general service list of English words*. London: Longman.
- Wode, H. (1999). Incidental vocabulary acquisition in the foreign language classroom. *Studies in Second Language Acquisition*, 21 (2), 243-258.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Zareva, A. (2005). Models of lexical knowledge assessment of second language learners of English at higher levels of language proficiency. *System*, 33, 547-562.

A témához kapcsolódó saját publikációk

- Lehmann, M. (2007a). Is intentional or incidental vocabulary learning more effective? *International Journal of Foreign Language Teaching*, 3 (1), 23-28.
<http://www.tprstories.com/ijflt>
- Lehmann, M. (2007b). The lexical diversity of short texts: Exploring the productive-receptive continuum of lexical knowledge. In J. Horváth & M. Nikolov (Eds.), *UPRT 2007: Empirical studies in English applied linguistics* (pp. 293-305). Pécs: Lingua Franca Csoport.

- Lehmann, M. (2007c). The size of an English receptive vocabulary. An exploratory study of the EFL vocabulary size of English majors at the University of Pécs. In R. Sheorey, & J. Kiss-Gulyás (Eds.), *Studies in Applied Theoretical and Applied Linguistics* (pp. 25-38).
- Sifakis, N., Oder, T., Lehmann, M. & Bluma, D. (2007). Aspects of learner autonomy in the national curricula of four European countries. In A-B. Fenner, & D. Newby (Eds.), *Coherence of principles, cohesion of competences: Exploring theories and designing materials for teacher education* (pp. 132-151). Council of Europe Publishing: Strasbourg.
- Lehmann, M. (2006a). Vocabulary as a filter with first-year English majors. In M. Nikolov, & J. Horváth, *UPRT 2006. Empirical studies in English Applied Linguistics*, (pp. 121-137). Pécs: Lingua Franca Csoport.
- Lehmann, M. (2006b). Magoljunk vagy olvassunk?: El tanulmány a tudatos és járulékos szótanulás összehasonlítására. *Modern nyelvoktatás*, 12 (1), 44-58.
- Lehmann, M. (2005). Bárdos Jen : Él nyelvtanítás-történet. Könyvismertet . *Magyar Pedagógia*, 14 (3), 377-378.
- Lehmann, M. (2004). Az idegennyelvi szókincs elsajátítás és mérés innovatív irányzatai. *Alkalmazott nyelvtudomány*, 4 (1), 41-57.
- Lehmann, M. (2003a). Range. Szoftver-kritika. *Modern nyelvoktatás*, 9 (4), 88-89.
- Lehmann, M. (2003b). The lexis of writing and vocabulary size: The relationship between receptive knowledge and productive use. In J. Andor, J. Horváth, & M. Nikolov (Eds.), *Studies in English Theoretical and Applied Linguistics*, (172-182). Pécs: Lingua Franca Csoport.
- J. G. Gebhard, M. Fodor & M. Lehmann. (2003c). Teacher development through exploration: Principles, processes and issues in Hungary. In J. Andor, J. Horváth, & M. Nikolov (Eds.), *Studies in English Theoretical and Applied Linguistics*, (250-262). Pécs: Lingua Franca Csoport.

Konferencia előadások

- Lehmann, M. (2009a). Korpusz-alapú szókincsmérések angol szakos hallgatókkal. 2009. október 26-27. Új módszerek az alkalmazott nyelvészeti kutatásban című konferencia. Budapest, ELTE-Kodolányi János F iskola.
- Lehmann, M. (2009b). Free productive vocabulary in student essays. 2009. június 12. University of Pécs Round Table conference.
- Lehmann, M. (2009c). Issues of student vocabulary: quantity, quality, purpose. 2009. január 22-24. Pécs, HUSSE konferencia.
- Lehmann, M. (2007a). Lexical diversity of short texts: Exploring the receptive/productive continuum of lexical knowledge. 2007. június 2. Pécs. University of Pécs Round Table Conference.
- Lehmann, M. (2007b) Angol szakos hallgatók produktív szókincse. 2007. április 19-21. Kodolányi János F iskola, Siófok. XVII. Magyar Alkalmazott Nyelvészeti Kongresszus.
- Lehmann, M. (2006a). Az idegennyelvi szókincs és az olvasott szövegértés kapcsolata. Mérések angol szakos hallgatókkal. XVI. MANYE Kongresszus, Szent István Egyetem, Gödöll . 2006. április 10-12.

- Lehmann, M. (2006b). Is intentional or incidental vocabulary learning more effective in the long-term retention of words? University of Pécs Round Table conference. 2006. január 20. Pécs.
- Lehmann, M. (2005). EFL vocabulary as a predictor of success in academic reading. 2005. szeptember 14-17. EUROSLA 15 Conference, Dubrovnik.
- Lehmann, M. (2004). Az idegen nyelvi receptív szókincs nagysága. IV. Országos Neveléstudományi Konferencia 2004. október 20-22. Budapest, MTA.
- Lehmann, M. (2003). Korpusz alapú szógyakorisági listák alkalmazása a szókincsmérésben. III. Országos Neveléstudományi Konferencia. 2003. október 9-11. Budapest, MTA.
- Lehmann, M. (2002). The relationship between the lexis of L2 learners' writing and the size of their receptive vocabulary. XII. MANYE konferencia, 2002. március 27-29. Szeged, Juhász Gyula Tanárképző Iskola.