



Gennem kejserdømme, kommunisme og republik – akademiordbogen lever!

Om den ungarske modersmålsleksikografi fra begyndelsen og til nyeste tid med særlig hensyntagen til det Ungarske Videnskabernes Selskabs nye storordbog

Af Loránd-Levente Pálfi og Júlia Pajzs

Med denne artikel vil vi give dels en oversigt over ungarsk modersmålsleksikografi fra begyndelsen til nyeste tid, dels en redegørelse for Det Ungarske Videnskabernes Selskabs nye flagskib: *A magyar nyelv nagyszótára I-IIff*¹ (Ittész et al., 2006ff). En indledende kort ordbogshistorisk oversigt vil – ud over at have et formål i sig selv – have det særlige formål at belyse betydningen af den nye storordbog, som for nylig, nemlig i november 2006, er udkommet med de to første af planlagte 18 bind.² Da den ene af forfatterne til denne artikel er medredaktør på storordbogen, kan der her ikke gives nogen kritisk analyse af værket.³ Vi vil i stedet forsøge at placere det i en (ungarsk) ordbogshistorisk kontekst, samt redegøre for dets tilblivelseshistorie, for de leksikografiske metoder, der ligger til grund, og ikke mindst for problemerne under arbejdet.

Det nye storværk, som primært vil blive behandlet i kapitlet "Den nye akademiordbog", vil herefter blive omtalt som "akademiordbogen". For ordens og rigtigheds skyld om end med fare for at skabe unødigt forvirring, skal det her imidlertid bemærkes, at der i Ungarn allerede findes to ældre akademiordbøger, såfremt man ved akademiordbøger i overensstemmelse med Wiegand (2003: XI) forstår ordbøger, som er blevet udarbejdet inden for akademiers eller akademilignende institutioners myndhedsområde. Disse to ældre akademiordbøger vil vi dog så vidt muligt udelukkende henvise til med deres titler og/eller ophavsmænd efterfulgt af angivelse af udgivelsesår.

I sammenligning med anden europæisk – f.eks. dansk – ordbogsskrivning (Pálfi/Bergenholtz, 2007; Bergenholtz/Pálfi, 2008) kan ungarsk modersmålsleksikografi påberåbe sig en forholdsvis lang tradition, idet den første ungarsk-ungarske ordbog udkom så tidligt som 1784 (Baróti, 1784). Endnu mere imponerer den ungarske modersmålsleksikografi imidlertid med sin rigdom, dvs. det væld af afsluttede for ikke at tale om uafsluttede ordbogsarbejder, som den kan fremvise (se neden-

for). Desuden er dens tre højdepunkter især ud fra et vesteuropæisk synspunkt interessante mht. den tid og det samfund, de er opstået i: Den første storordbog, *A magyar nyelv szótára I-VI*⁴ (Czuczor/Fogarasi, 1862-1874), blev udgivet, mens Ungarn var en del af det Habsburgske Monarki; dens efterfølger, *A magyar nyelv értelmző szótára I-VII*⁵ (Bárcei et al., 1959-1962), udkom i en tid, hvor Ungarn var under russisk besættelse⁶, mens den tredje, seneste og nyeste storordbog, *A magyar nyelv nagyszótára I-IIff* (Ittész et al., 2006ff), som endnu er under udgivelse, udkommer i det moderne Ungarn, som er en demokratisk republik, der filmed for kort tid siden blev medlem af Den Europæiske Union.

Da det ikke kan forventes, at vore vel primært skandinaviske læsere er fortrolige med det ungarske sprogs historie, slægtskabsforhold mm., og da det vil være relevant senere i artiklen – især i forbindelse med den ordbogshistoriske oversigt – at kende lidt til disse ting, vil vi nedenfor give en meget kort indføring.

Det ungarske sprog

Det ungarske sprog hører som bekendt til den uraliske sprogæt – ikke til den indoeuropæiske, hvortil hører så forskellige sprog som dansk, fransk, græsk, russisk, walisisk, iransk, albansk, armensk mm. Betegnelsen "uralsk" kommer af, at den oprindelige hjemstavn for de pågældende folk menes at være Uralbjergkæden. Den uraliske sprogæt, som i dag tæller omkring 18 levende sprog⁷, deles traditionelt op i den finsk-ugriske

⁴ Da.: Ordbog over det ungarske sprog.

⁵ Da.: Betydningsforklarende ordbog over det ungarske sprog.

⁶ I tiden 1945-1990 var Ungarn besat af USSR. I 1947 overtog kommunisterne magten i landet; det kommunistiske regime bragtes først til ophør i 1989, og der blev afholdt frie valg året efter.

⁷ Her tænkes på følgende 18 sprog: Vogulsk, ostjakisk, komi-syrjænsk, komi-permjakisk, votjakisk, tjeremissisk, mordvinsk, karelsk, vepsisk, ingrisk, nenetsisk, enetsisk, nganasan, selkup, ungarsk, finsk, estisk og samisk. Der findes muligvis flere, muligvis færre – det er uklart og stort set umuligt at hitte rede i. På den ene side er der delte meninger blandt forskerne om, hvorvidt nogle af sprogene bør regnes for sproggrupper (f.eks. samisk), hvorfor man muligvis kunne tale om mange flere enkeltprog, på den anden side har flere af de små sprog i længere tid været døende. Således er f.eks. votisk, som taltes i området omkring Roja nordvest for Riga i Letland, og livisk, som taltes i området nord for Narva og vest for Sankt Petersborg ved den estisk-russiske grænse, begge uddøde i løbet af sidste halvdel af det 20. århundrede. Desuden har det 20. århundredes russificeringsbestrebelse og ikke mindst sovjet-tidens upålidelige folketællinger (for ikke at tale om folketæl-

¹ Da.: Storordbog over det ungarske sprog.

² Vi skal her takke seniorforsker Anna Braasch fra Center for Sprogteknologi for kritik af en tidligere version af nærværende artikel.

³ Júlia Pajzs har været tilknyttet akademiordbogsprojektet de seneste 23 år; hun blev i 1985 ansat på ordbogen som IT-fagkyndig med nogen lingvistisk baggrund. Ud over at hun har haft og stadig har ansvaret for den informationsteknologiske hhv. computerleksikografiske del af arbejdet, har hun været involveret i dets ledelse.



sprogfamilie på den ene side og den samojediske på den anden. Med undtagelse af ungarsk, finsk og til dels samisk tales de finsk-ugriske sprog, som i øvrigt deles op i sproggrupperne finsk⁸ og ugrisk, i dag udelukkede i det forhenværende Sovjetunionen. Set ud fra antallet af modersmålstalende er ungarsk det største blandt de finsk-ugriske sprog, idet det tales af over 14 millioner⁹, mens finsk tales af lidt over fem millioner og estisk af kun en million (Sinor, 1988; Csepregi, 1998).

I de historiske kilder optræder Ungarn første gang i det 5. århundrede e.Kr. (Forgács, 2007: 20), men sproget er væsentligt ældre: Det udskilles fra den ugriske sproggruppe omkring 1000 f.Kr. og er således 3000 år gammelt.

De tidligste sprogmindesmærker er de såkaldte strømindesmærker bestående af enkelte spredte gloser eller i sjældne tilfælde hele syntagmer (men ikke sætninger!) i nogle arabisk-persiske håndskrifter fra det 9. og 10. århundrede e.Kr. (Forgács, 2007: 20). Afhængigt af den givne forskers temperament er det imidlertid enten *Tihanyi apátság alapító levele* [Tihany abbediets stiftelsesbrev] fra 1055 eller en nekrolog, *Halotti beszéd* [døds-messetale], fra omkring 1195, der traditionelt regnes for det tidligste sprogmindesmærke. Førstnævnte dokument er en tekst forfattet på latin, men den indeholder 58 ungarske ord og 33 endelser (suffikser); sidstnævnte dokument er en tekst skrevet alene på ungarsk, der indeholder 190 ord.

Om end den dokumenterede skriftlighed kan forekomme forholdsvis sen i sammenligning med flere andre af de europæiske sprog, er ungarsk det sprog iblandt de uralske, hvor den er ældst. Såvel de uralske sprogs høje ælde som deres sene skriftlighed er i øvrigt alment kendte forhold: Langt de fleste uralske sprog har overhovedet først fået en skrift i det 20. århundredes Sovjetunionen.

Fra det 15. århundrede haves i Ungarn en decideret vernakulær bogproduktion, og den første komplette bibeloversættelse foreligger i 1590 (en delvis oversæt-

lingerne forud for den tid) på den ene side, og de forskellige ikke-slaviske folks selvvalgte opgivelse af nationalsproget ud fra ønsket om fuld integration i det russiske samfund på den anden side gjort det særdeles vanskeligt for forskerne at verificere forholdene i dag.

⁸ Det skal bemærkes, at den danske benævnelse "finsk(-)" i denne sammenhæng ikke er særligt præcis. På engelsk taler man om "Finno-Ugric", når man vil benævne sprogfamilien (på dansk "finsk-ugrisk"), og "Fennic" eller "Finnic" i modsætning til "Ugric", når man vil benævne den pågældende sproggruppe (da. "finsk" og "ugrisk"), for først dernæst at bruge betegnelserne "Finnish", "Hungarian" osv. om de enkelte sprog (da. "finsk", "ungarsk" osv.).

⁹ Heri er medregnet den ungarske diaspora på ca. fire millioner fordelt over især Transsylvanien (som siden Trianontraktaten, der blev underskrevet 4. juni 1920, har været en del af Rumænien), det nordlige Serbien, Slovakiet, det sydvestlige Ukraine, Østrig, Tyskland, Frankrig, Storbritannien, USA og Canada.

telse foreligger dog allerede i 1533 udført af Benedek Komjáti). Litteratursprogets egentlige opkomst og opblomstring sker imidlertid først fra udgangen af det 18. til og med den første halvdel af det 19. århundrede. Bevægelsen for sprogfornyelse spillede i denne sammenhæng en vigtig rolle. Det var hovedsagelig en politisk bevægelse, som tilstræbte ensartethed i modersmålet, en deraf affødt udligning af dialekterne og ikke mindst en kraftig modstand mod tysk, som på daværende tidspunkt var det dominerende sprog i især øvrighedsanliggender. Det var først i 1844, at ungarsk opnåede status som officielt sprog i Ungarn, men denne status gik tabt allerede få år efter – nemlig 1849, dvs. efter revolutionen – til fordel for tysk. Den blev genvundet i 1867.¹⁰

Den sproghistoriske inddeling af ungarsk ser således ud (Kis/Pusztai, 2005: 16-17):

- *urungarsk* (1000 f.Kr.-896 e.Kr.; ved en finere inddeling taler man om en *ældre* og en *ynge urungarsk* tid, hvor der skelnes mellem tiden fra den ugriske udskillelse frem til folkevandringen omkring det 5. årh. e.Kr. og tiden herefter frem til landnamstiden i 896)
- *oldungarsk* (896-1526, dvs. fra landnamstiden til slaget ved Mohács 1526, som markerer begyndelsen til tyrkervældet)
- *middelungarsk* (1526-1772, dvs. fra Osmannernes besættelse til oplysningstidens indtræffen)
- *nyungarsk* (1772-1920, dvs. fra oplysningstiden til Trianontraktaten)
- *nutidsungarsk* (1920-).

Ungarsk modersmålsleksikografi fra begyndelsen til i dag

Kendetegnende for den nationale selvbevidstheds fødsel i Ungarn var blandt andet nationalsprogets begyndende opblomstring og bevidste brug. Det officielle sprog blandt gejstligheden såvel som blandt de øvrige herskende klasser i samfundet var igennem hele middelalderen latin, og fra det 18. århundrede blev tillige tysk dominerende, idet Ungarn blev en del af det Habsburgske Monarki. Men hen imod århundredskiftet kom det på mode i de lærde adelige kredse at bruge ungarsk i tale og skrift, og det er i denne tid, at en egentlig ungarsk litteratur fødes og begynder at udbredes. Tidsalderens begyndelse sættes traditionelt til 1772 (se ovenfor om den sproghistoriske inddeling), som er udgivelsesåret for Bessenyei Györgys digterværk *Ágias*

¹⁰ Sprogsituationen i Ungarn var i grunden mere sammensat, end det er muligt at redegøre for i denne korte indledning. Det er måske oversimplificeret at hævde, at tysk var det officielle sprog. Der var en ihærdig kamp imod tysk som officielt sprog, og latin blev aldrig helt opgivet. Før 1844 var det latin, der var det officielle sprog, dvs. det sprog, som myndighederne brugte. I 1844 fik ungarsk for første gang denne status. Efter revolutionen blev tysk det officielle sprog, mens latin stadig havde en slags officiel status, da det i kirken brugtes til registrering af fødsler, ægteskaber osv. Kirken begyndte først meget sent at bruge ungarsk, nemlig fra udgangen af det 19. århundrede.



tragédiája¹¹. Her begynder den ungarske sprogfornyelse, og det er dette år, der er blevet valgt som begyndelsesår for kildesamlingen til den nye akademiordbog.

Opbrudstidens højdepunkt nåedes 1825, da greve Széchenyi István (1791-1860) grundlagde Det Ungarske Videnskabernes Selskab, eller som det på ungarsk hedder: Magyar Tudományos Akadémia (MTA), dengang kaldet Magyar Tudós Társaság, iblandt hvis vigtigste målsætninger det var at udarbejde en omfangsrig betydningforklarende, historisk modersmålsordbog, der kunne måle sig med de øvrige europæiske akademiordbøger, som på det tidspunkt enten var udkommet eller under udgivelse. Der blev udkastet en plan, men man indså snart, at det var umuligt inden for en forsvarlig tid og for en rimelig betaling at udarbejde en ordbog så stor som foreskrevet af planen.

Som en slags midlertidig løsning lod man derfor udarbejde seksbindsværket *A magyar nyelv szótára I-VI*¹² (Czuczor/Fogarasi, 1862-1874). Ordbogen tilstræber en om muligt fuldstændig leksikografisk kodificering af de navnkundige skønlitterære forfatteres ordforråd fra især første halvdel af det 19. århundrede. Sprognet i værket er normativt, men ordbogsredaktørerne så igennem fingre med forkert sprogbrug, om end det blev markeret, såfremt der var tale om gængse forsyndelser, som også fandt hjemmel i kendte skribenters værker. Dialektale og forældede ord blev kun medtaget, såfremt ordbogsredaktørerne fandt dem interessante ud fra et rent sproghistorisk synspunkt. Om end værket også er puristisk i sit sprogsyn, indeholder det mange sådanne fremmedord, som der dengang endnu ikke fandtes udbredte og alment kendte rent ungarske ord for. Ud over betydningforklaringerne bringes – dog noget spredt – levende eksempler (citater fra den anerkendte skønlitteratur), der optegnes kendte ordsprog, og ved udgangen af artiklerne anføres undertiden enkelte talemåder.

Næppe stod værket færdigt, før man fra forskellig side anså det for nødvendigt med yderligere supplerende leksikografiske tiltag. I anden halvdel af det 19. og første halvdel af det 20. århundrede udkom således adskillige – delvis ufuldendte – betydningforklarende historiske ordbøger.

Målet med *Magyar nyelvtörténeti szótár* [...] *Lexicon linguae Hungaricae aevi antiquioris I-III*¹³ (Szarvas/Simonyi, 1890-1893), som er en ordbog over old- og middelungarsk, var en udtømmende leksikografisk optegnelse og beskrivelse af det i skriftsproget forefundne ordforråd fra de tidligste sprogminde-mærker til og med udgangen af det 18. århundrede. Desværre er det anvendte kildemateriale dels langtfra udtømmende, dels er der opstået talrige filologiske unøjagtigheder under ind-

samlingen og bearbejdelsen af stoffet (Kiss, 1994: 13-32). Trods disse omstændigheder er værket til stadighed den eneste fuldførte ordbog over old- og middelungarsk.

I første halvdel af det 20. århundrede søgte man ihærdigt at få igangsat udarbejdelsen af den store akademiordbog, som oprindeligt var ment som en fortsættelse til Szarvas/Simonyi (1890-1893). De fleste forsøg bar imidlertid ikke frugt. I stedet udkom der andre, mindre ordbøger, f.eks. den af den senere præsident for den store akademiordbogs kommission Kálmán Szily redigerede *Magyar nyelvújítás szótára*¹⁴ (Szily, 1902-1908), hvori kun det ordgods, som var resultatet af det 19. århundredes sprogfornyelse, søgtes optegnet og beskrevet. Fuldstændighed har ikke været tilstræbt i dette værk, idet der fortrinsvis blev medtaget orddannelser, som endnu var i brug ved udgangen af det 19. århundrede. Ord, som i mellemtiden var uddøde, blev kun medtaget, for så vidt deres historie i nogen grad kunne have videnskabelig interesse.

1950'erne og 1960'erne regnes i nyere ungarsk ordbogshistorie for en aldeles frugtbar periode. På den ene side havde den kommunistiske statsmagt en stor interesse i – såvel deskriptiv som normativ – leksikografi, som den mente var et spejlbillede af samfundsforholdene, hvorfor den sikrede ordbogsskrivningens vilkår inden for Det Ungarske Videnskabernes Selskabs Institut for Sprogvidenskab, som på dette tidspunkt nyligt var blevet oprettet ud af en række andre institutter efter det sovjetiske akademis mønster. På den anden side skulle der sikres stillinger for talrige sprogvidenskabsfolk, som tidligere havde bestridd professorater på de forskellige universiteter landet over, men som nu – grundet deres borgerlige herkomst – skønnedes at være for farlige som undervisere af den studerende ungdom.

Under disse omstændigheder skete det, at et forskerhold bestående af lutter fremtrædende skikkelser udarbejdede syvbindsværket *A magyar nyelv értelmező szótára I-VII*¹⁵ (Bárczi et al., 1959-1962), hvis kildemateriale bl.a. udgjordes af det seddelkartotek, som oprindeligt var blevet tilvejebragt som grundlag for udarbejdelsen af den store akademiordbog og som i 1962, da ex-cerperingen (indsamlingen af ord- og citatmateriale) afsluttedes, rummede seks millioner sedler. Om end tanken om den store akademiordbog stadig ikke var blevet opgivet, blev den heller ikke med dette værk søgt virkeliggjort. Ikke desto mindre nyder ordbogen den dag i dag særstatus i det leksikografiske landskab i Ungarn, da den er enestående i sin bredde mht. anvendt kildemateriale. Selvom den på sin vis er blevet vakt til live igen, idet den for nogle år siden blev udgivet på cd-rom (Bárczi et al., 2004), er den naturligvis i vid udstrækning forældet i dag. Redaktørerne har aflagt beretning om de gjorde erfaringer i talrige ledsageartikler; f.eks. giver Országh (1960) på engelsk en ud-

¹¹ Da.: Tragedien om Agis.

¹² Da.: Ordbog over det ungarske sprog.

¹³ Da.: Sproghistorisk ordbog over ungarsk. Fra de ældste sprogminde-mærker til sprogfornyelsen.

¹⁴ Da.: Den ungarske sprogfornyelses ordbog.

¹⁵ Da.: Betydningsforklarende ordbog over det ungarske sprog.



førlig udredning af de vigtigste redaktionelle principper.

Ordbogen indeholder 58.000 opslagsord, af hvilke mange illustreres med levende eksempler. Set ud fra sideantallet (7383 sider) er den væsentligt større end *Duden. Das große Wörterbuch der deutschen Sprache in zehn Bänden* (1999) (4800 sider, 200.000 opslagsord), men set ud fra antallet af opslagsord er den lidt mindre end *Den Danske Ordbog 1-6* (Hjorth et al., 2003-2005) (4422 sider, 63.000 opslagsord). Den indeholder imidlertid næsten 150.000 angivelser af komposita (første- og sidsteled), hvorved den – om end for størstedelens vedkommende nok så perifert – behandler et meget større antal ord. Af de ord, som var opstået i tiden umiddelbart forud for værket udkomst, blev de mest betydningsfulde medtaget, dvs. dem som allerede før kommunismens fremmarch havde fået borgerret i sproget (f.eks. *tózsde* [børs], *bordély* [bordel]).

Ordbogens største fejl er dens ideologiske tilsnit: at den er gennemsyret af det marxistisk-leninistiske verdenssyn, uden hvilket den naturligvis aldrig kunne være blevet udgivet.¹⁶ Der var nedsat et særligt redaktionshold, som udelukkende havde den arbejdsopgave at lade de ellers kyndigt og korrekt formulerede betydningsforklaringer undergå en ideologisk prøve. Det er takket være denne arbejdsindsats, at ordbogen fremviser adskillige betydningsforklaringer, som må forekomme underholdende for den nutidige bruger; se f.eks. definitionen af *jótékony* [velgørende], et ord som blev præget før den socialistiske tid:

<Et eller flere i velstand levende mennesker> som <i det udbytteriske samfund> ved brug af menneskevenlige slagord med religiøse undertoner til at skjule det egentlige budskab om den fundamentale samfundsomvæltningens nødvendighed lader de udvalgte "fattige" blive understøttet på ydmygende vis. (*A magyar nyelv értelmészóótára I-VII.* (Bárczi et al., 1959-1962.) Vores oversættelse).

På trods af de nævnte mangler er der tale om den hidtil mest omfangsrige betydningsforklarende modersmålsordbog på ungarsk, som også bringer ikke-konstruerede, såkaldt levende eksempler som belæg.

En sammentrængt udgave af værket udkom med håndordbogen *Magyar értelmező kéziszótár*¹⁷ (Juhász et al., 1972), som trods dens mindre størrelse bringer mange flere opslagsord (72.000). Til dette værk har også de dengang nyeste ord fundet vej, mens den ideologiske orientering er aftaget kraftigt. Med ønsket om trods de strenge udvortes rammer at optage flest mulige ord, – værket foreligger både i en etbinds- (på 1550 sider) og i en tobindsudgave – blev løsningen med opsætningen af ord med samme rod noget ejendommelig: Den stærkt

nichealfabetiske opstilling¹⁸ er lidet brugervenlig, idet den er hindrende for tilgangsstrukturen¹⁹. Omfangsbegrænsningen til trods gives et overvældende udvalg af leksikaliserede flerordsenheder med en tydelig skelnen mellem idiommer, ordsprog mm. Selvom man også ved dette værks udarbejdelse har benyttet sig af akademiordbogens seddelsamling som kildegrundlag, gives der pga. pladsmangel kun en sjælden gang levende eksempler. I reglen er det såkaldt døde eksempler (dvs. af leksikograferne konstruerede eksempler) og betegnende ordforbindelser, der skal sikre den rette forståelse og den rigtige brug af det givne ord. Det er også omfangsbegrænsningen, der har medført, at betydningsforklaringerne ofte blot er angivelser af et enkelt synonym: I reglen gives der et mere kendt ord for et mindre kendt – noget, som måske ikke altid imødekommer brugernes behov. For ikke-modersmålstalende er ordbogen svært brugbar til sprogindlæring, selv den almindelige modersmålstalende kan undertiden blive skræmt væk. Trods disse omstændigheder er værket i sin nye, lettere reviderede udgave (Pusztai, 2003) til dags dato den mest brugte og bredest dækkende almensproglige betydningsforklarende modersmålsordbog. I den nye udgave andrager antallet af opslagsord 75.000.

I cd-rom-udgaven er søgningen blevet gjort væsentligt lettere, bl.a. volder de nichealfabetisk ordnede opslagsord her ingen problemer. Den nye udgave er desuden ikke alene blevet udvidet med de forgangne årtiers vigtigste nye ord, men tillige med nye betydninger, og som noget helt nyt har man medtaget ord og betydninger fra de ungarske sprogvarieteter, som findes uden for landets nuværende grænser. Et voldsomt krævende arbejde har det desuden været at frigøre betydningsforklaringerne fra deres ideologiske præg. I den nye udgave er betydningsforklaringerne tilstræbt udformet så neutrale som muligt. Man har bestræbt sig på at vige uden om enhver form for politisk stillingtagen og undlade at bringe almene værdidomme.²⁰ Et eksempel på definitionernes forskellighed i den gamle og den nye

¹⁸ Her er tale om en "alfabetisk makrostruktur", hvor der forekommer nicher, dvs. strengt alfabetisk ordnede mængder af artikler, som er blevet samlet i artikelblokke.

¹⁹ Med *tilgangsstruktur* menes strukturen af de indikatorer, som fører ordbogsbrugeren frem til de oplysninger, han/hun ønsker. Man skelner mellem *ydre* og *indre tilgangsstruktur*. Den ydre tilgangsstruktur er strukturen af de indikatorer, som begynder allerede på ordbogens omslag, og som fører brugeren frem til den enkelte ordbogsartikel, mens den indre tilgangsstruktur er strukturen af de indikatorer, som fører brugeren videre ind i artiklen og frem til den søgte information (Bergenholtz et al., 1997: 256).

²⁰ I virkeligheden er alle valg ved udarbejdelsen af en ordbog sprogpoltiske, således lyder tesen i Bergenholtz/Gouws (2006): "Every single lexicographical decision has a language policy relevance and therefore in the end a political dimension"; jf. også Bergenholtz/Pálfi (2008): "Problemet er, at sådanne sprogpoltiske afgørelser i de fleste ordbøger gøres implicit, en eksplicit tilkendegivelse af redaktionens politiske holdning, f.eks. efter et særligt symbol eller en forkortelse (f.eks. ANM), ville gøre det mere klart, hvor en redaktion står."

¹⁶ Noget, som vi også kender fra f.eks. østtysk leksikografi (Ludwig, 1998; Malige-Klappenbach, 1989), og i det hele taget fra Østblokken (Wierzbicka, 1995).

¹⁷ Da.: Betydningsforklarende håndordbog over ungarsk.



udgave (med begge artikler oversat til dansk) ses i det følgende:

middelklasse subst Det velhavende småborgerskab og intelligentsiaen, som i det kapitalistiske samfund befinder sig imellem de to grundlæggende samfundsklasser. (*Magyar értelmező kéziszótár* (Juhász et al., 1972.) Vores oversættelse).

middelklasse subst Borger eller samfundsgruppe, som er i stand til på værdig vis og af egen kraft og eget arbejde at forsørge sig selv og sin familie. (*Magyar értelmező kéziszótár* (Pusztai, 2003.) Vores oversættelse).

Under arbejdet med 2003-udgaven har man desværre ikke brugt de elektroniske tekstkorpora, som var blevet udarbejdet på MTA's Institut for Sprogvidenskab, og som på det tidspunkt var tilgængelige (<http://www.nytud.hu>), til andet end bestemmelsen af opslagsordenes hyppighed.

Den ved systemskiftet udgivne *Magyar Larousse. Encyklopédikus szótár I-III*²¹ (Bakos et al., 1991-1994) er en interessant sammensætning af en – om end nok så lille – encyklopædi, en håndordbog og en fremmedordbog. For gennemsnitsbrugeren må værket utvivlsomt være et særdeles nyttigt og praktisk værktøj; kender brugeren ikke betydningen af et ord, eller er han/hun usikker på dets brug, skal vedkommende ikke rådspørge en række forskellige kilder, men vil tit kunne få såvel saglig som sproglig oplysning i et og samme værk. I betragtning af, at ungarske ordbøger sædvanligvis ikke bringer encyklopædiske definitioner, er dette værk et særdeles velkomment tiltag på det ungarske bogmarked. Dette være i øvrigt sagt, velvidende at en skarp skelnen mellem sprog- og sagleksikografi er problematisk, jf. hertil på den ene side Bergenholtz (1995b; 1996; 1999) og Pálfi (2007) og på den anden side Pedersen (1994) og Lundbladh (1999). I den foreliggende sammenhæng og i artiklen i det hele taget er det således ikke vores forståelse, men netop de pågældende ordbøgers selvforståelse, der kommer til udtryk.

Som noget helt nyt er udkommet tobindsværket *Értelmező szótár*²² (Eöry, 2007), som med sine 16.000 opslagsord har et væsentligt lavere antal af opslagsord end *Magyar értelmező kéziszótár* (Pusztai, 2003), men som formodentlig alligevel vil vinde langt større udbredelse og popularitet end denne. Der er tale om en udpræget brugervenlig ordbog udtænkt og udformet til brug for unge mennesker og ordbogsuvante voksne. Værket er illustreret, har en luftig sideopsætning, og opslagsordene såvel som de øvrige elementer i artiklerne er lette at finde. Desværre støtter betydningsforklaringerne sig for meget op ad definitionerne i *Magyar értelmező kéziszótár* (Pusztai, 2003), hvorved ikke få af dennes fejl gentages (se f.eks. definitionen af *áram* [strøm]). Der bringes rigeligt med synonymer, og hvor det er muligt, oplistes tillige antonymer; til sidstnævnte har *Magyar*

*szókinctár*²³ (Kiss, 1998) været brugt som kilde. Eksempelsætningerne er ikke ægte belæg, men af leksikograferne forfattede eksempler, hvilket dog ingenlunde betyder, at de er uanvendelige for brugeren. Ordbogen giver talrige råd om sprogbrug og trodser dermed den i dagens Ungarn fremherskende og af lingvisterne forkyndte opfattelse, at den præskriptive lingvistik udelukkende er skadelig og i øvrigt hører fortiden til. Der medtages også gængse talemåder og ordsprog indsamlet fra tre ældre idiom- og ordsprogsordbøger, nemlig Bárdosi (2003), Forgács (2003) og T. Litovkina (2005). Ved udgangen af artiklerne gives tillige etymologiske oplysninger udarbejdet på baggrund af den nyligt udkomne og helt nyskrevne etymologiske ordbog af Zaicz (2006). Det er en nyskabelse i ungarsk leksikografi, at der ved udgangen af artiklerne også gives en slags emneklynge, idet der henvises til en række andre begrebsligt beslægtede opslagsord. *Értelmező szótár*+ egner sig især til brug som etsprogsordbog for ikke-modersmålstalende lærere på mellemniveau.

I det ungarske ordbogslandskab af i dag er især korpusleksikografien underrepræsenteret; den nye akademiordbog vil delvis kunne imødekomme denne mangel. Den er delvis udarbejdet på grundlag af et elektronisk lagret tekstkorpus, dvs. en digital tekstmasse, delvis på grundlag af en traditionel fastformsseddelsamling, nemlig belægsamlingen, den bringer levende eksempler samt frekvensangivelser, og den henter hjælp og støtte i den nyeste lingvistiske forskning, når det gælder ordklassernes bestemmelse. Alene på grund af dens planlagte omfang (18 bind) og den tid, arbejdet forventes at tage (ifølgende planen skal sidste bind udkomme i 2031), egner ordbogen sig dog næppe til den almindelige lægmands hverdagsbehov. Værket er heller ikke konciperet til den brede offentlighed, men til akademikere; det er at opfatte som en slags sprogmuseum, som dokumenterer de forgangne 230 års ungarske ordforråd. Som potentielle primærmålgrupper har især følgende været tænkt: Biblioteker, universiteter og ungarske kulturinstitutioner (også i udlandet).

Den nye akademiordbog

Tilblivelseshistorie

At dømme ud fra arbejdets størrelse og varighed, værkets størrelse og grundighed, de økonomiske omkostninger gennem årene samt medarbejdernes antal (som andrager over 100 ansatte hhv. tilknyttede) er *A magyar nyelv nagyszótára I-IIff* (Ittész et al., 2006ff) utvivlsomt ungarsk ordbogsskrivnings hidindtil størst anlagte og mest ærgerrige tiltag, i hvert fald såfremt vi ser på den traditionelle sprogleksikografi og ikke på sagleksikografien, som bl.a. fremviser storværker som *Magyar nagylexikon I-XIX*²⁴ (Élesztős et al., 1993-2004), dvs. nationalencyklopædien, og dennes forgænger *Révai nagy lexikona I-XXI*²⁵ (1911-1935).

²¹ Da.: Ungarsk Larousse. Encyklopædisk ordbog.

²² Da.: Betydningsforklarende ordbog+.

²³ Da.: Ungarsk begrebsordbog.

²⁴ Da.: Den store ungarske encyklopædi.

²⁵ Da.: Révais konversationsleksikon ("Révai" er et efternavn).



Allerede i 1817 blev idéen om en stor akademiordbog fostret i de lærde kredse i Ungarn. Det var greve József Teleki, som senere blev Det Ungarske Videnskabernes Selskabs første præsident, der som den første fremlagde en plan om "udarbejdelsen af en fuldstændig ungarsk ordbog" (oversat efter Gerstner, 2006: 10). Man indså dog snart, at planen var urealistisk og lod derfor som en midlertidig løsning udarbejde *A magyar nyelv szótára I-VI* ved Gergely Czuczor og János Fogarasi (se ovenfor). I 1898 blev der endelig nedsat en ordbogskommission, som holdt sit første officielle møde den 17. maj samme år, men der gik ikke lang tid, før arbejdet pga. dårlig økonomi og senere pga. Første Verdenskrigs udbrud gik i stå. Det blev genoptaget igen i 1922. I 1924 rummede belægsamlingen 2,5 millioner citatsedler, og i 1933 var den vokset til 3,16 millioner (Gerstner, 2006: 12), der heldigvis kom uskadt igennem Anden Verdenskrig. Ved midten af århundredet var samlingen vokset til fire millioner sedler, og det forventedes, at ordbogen, som skulle indeholde et sted mellem 300.000 og 400.000(!) opslagsord, var færdig i 1973! I 1960 ses for første gang konkrete tal på tryk. Her hedder det, at ordbogen skal fylde 25-30 bind à 1300-1400 sider og 10.000-12.000 opslagsord eller 300.000 opslagsord i alt (Gáldi, 1960: 185). Da kræfterne i disse år fortrinsvis blev brugt på *A magyar nyelv értelmező szótára I-VII* (Bárczi et al., 1959-1962) og *Magyar értelmező kéziszótár* (Juhász et al., 1972), (se ovenfor), fik den store akademiordbog imidlertid ikke megen opmærksomhed, og belægsamlingen blev indstillet i 1962. På dette tidspunkt rummede den over seks millioner citatsedler! Om end en sporadisk supplerende citatsamling blev opretholdt nogle år endnu (Szabó, 2006: 53), løb man snart tør for pekuniære midler, og arbejdet indstilledes helt i 1974.

Seddelsamlingen lå dog ikke ubrugt hen, men fandt anvendelse i talrige andre forskningsprojekter. Foruden i nogle af de ovenfor nævnte værker blev den f.eks. brugt i arbejdet med den store etymologiordbog, *A magyar nyelv történeti-etimológiai szótára I-IV*²⁶ (Benkő, 1967-1984), som siden er udkommet i en omarbejdet udgave på tysk med titlen *Etymologisches Wörterbuch des Ungarischen I-III* (Benkő, 1993-1997), og det stort anlagte, men endnu uafsluttede ordbogsværk over ungarske dialekter, *Új magyar tájszótár*²⁷ (B. Lőrinczy, 1979-2002), som blev påbegyndt i 1950'erne, samt i en række sproghistoriske monografier.

Senere hen i 1970'erne blev arbejdet med akademiordbogen vakt til live igen, og det blev nu understreget, at Det Ungarske Videnskabernes Selskabs Institut for Sprogvidenskab hverken kunne eller skulle stå alene med ansvaret, men at ordbogen var af største interesse ikke blot for den nationale sprogvidenskab, men også for samfundsvidenskaberne – ja, "i grunden drejer det sig om et fællesakademisk ansvar!", hed det i et opråb i slutningen af 70'erne (frit oversat efter Imre 1977: 286).

Således lykkedes det endelig i begyndelsen af 1980'erne – og efter at projektet var kuldsejlet utallige gange, hhv. havde resulteret i nogle midlertidige løsninger bl.a. i form af de to mellemstore akademiordbøger – at få fast grund under fødderne. Ifølge den nye plan skulle ordbogen, som nu bar titlen *A magyar irodalmi és köznyelv nagyszótára*²⁸ og havde *Trésor de la langue française 1-16* (Imbs et al., 1971-1994) som forbillede, omfatte det ungarske sprog fra 1533 til 1990. Begyndelsestidspunktet er udgivelsesåret for Benedek Komjátis ufuldstændige ungarske bibeloversættelse (jf. hertil Kiss/Pajzs 1989). Den skulle fylde 20 bind med i alt omtrent 200.000-250.000 opslagsord. Efter en del stridigheder om omfanget og efter at have indset, hvor frygtelig omkostningsfuldt det ville blive at skabe et værk så omfattende, blev konceptet ændret adskillige gange, (den ene af gangene hed det, at ordbogen grundet økonomiske hindringer ikke kunne omfatte mere end 10-12 bind indeholdende højst 100.000 opslagsord i alt!), indtil der omkring år 2002 nåedes enighed om den nuværende plan: Ordbogen skal fylde 18 bind, rumme 110.000 opslagsord, begrænse sig til at optegne og beskrive ungarsk i tiden 1772-2000 (dvs. i stedet for 1533-1990), og den forventes at udkomme med sidste bind i 2031, dvs. ca. 200 år efter idéens fødsel og 50 år efter redaktionsarbejdets påbegyndelse.

Kildemateriale

Hovedforskellen mellem den første, den anden og den tredje akademiordbog – dvs. *A magyar nyelv szótára I-VI* (Czuczor/Fogarasi, 1862-1874), *A magyar nyelv értelmező szótára I-VII* (Bárczi et al., 1959-1962) og *A magyar nyelv nagyszótára I-IIff* (Ittész et al., 2006ff) – er ikke alene størrelsen, men især den omstændighed, at man i sidstnævnte værk dokumenterer hver eneste betydning med et autentisk eksempel, et belæg.

Den leksikografiske beskrivelse af de enkelte ord i den nye akademiordbog bygger hovedsageligt på følgende tre kilder: 1) belægsamling; 2) korpus; 3) supplerende tekstsamling (bestående af flere hundrede overvejende skønlitterære værker udgivet på cd-rom). Disse har tilsammen affødt 30.000 bibliografiske poster i kildefortegnelsen. Akademiordbogen bygger således dels på nogle af den moderne leksikografis landvindinger, korpusmetoden, dels på den traditionelle excerperingsmetode med en deraf affødt belægsamling (om belægsamling vs. tekstkorpus se Bergenholtz (1988: 17) og Dyrberg et al. (1988: 211)). Belægsamlingen indeholder excerperter (udpluk) fra skøn- og faglitteratur, religiøs litteratur, populærvidenskabelig litteratur, aviser, tidsskrifter mm. fra tidsrummet 1772-2000.

Da man i 1985 påbegyndte ordbogsarbejdet på ny, forestillede man sig, at den via traditionelle metoder tilvebragte belægsamling, som var opstået i de tidligere faser af ordbogsarbejdet, ikke skulle bruges til udarbejdelsen af værket, men at man udelukkende skulle

²⁶ Da.: Historisk-etymologisk ordbog over det ungarske sprog.

²⁷ Da.: Ny ungarsk dialektordbog.

²⁸ Da.: Storordbog over det ungarske litterære og almene sprog.



bruge en digitaliseret tekstsamling, et korpus, som materialegrundlag. For at virkeliggøre denne plan oprettede Det Ungarske Videnskabernes Selskabs Institut for Sprogvidenskab i 1985 to nye stillinger: En data-nomstilling og en operatørstilling. Begge disse ansatte er den dag i dag tilknyttet projektet. Det første år blev prøvebearbejdningen igangsat på Commodore 64-maskiner, alt imens litteraturhistorikere i samarbejde med lingvister var beskæftigede med udvælgelsen af de tekstsekvenser, som skulle indgå i den elektroniske tekstsamling (korpus).

Til at begynde med var tanken at lave et korpus på 10-13 millioner løbende ord, men som arbejdet skred frem, blev det åbenlyst, at der ville blive brug for en langt større tekstmasse. Resultatet blev det korpus, som sammen med belægsamlingen og den supplerende tekstsamling (cd-rom-udgivelserne) danner grundlag for værket og bærer navnet *Magyar szótörténeti szöveggyűjtés*²⁹. Det er blevet udarbejdet i tidsrummet 1985-2005 af et forskerhold bestående af litteraturhistorikere og sprogvidenskabsfolk, det er tidsmæssigt afgrænset nøjagtigt som siddelsamlingen, dvs. det er sammensat af tekster fra tidsrummet 1772-2000, og det indeholder 27,5 millioner løbende ord (Szabó, 2006: 53).

Målet har været at indfange hverdags sproget såvel som det litterære og det akademiske hhv. det i videnskabelige kredse brugte sprog – i det hele taget sproget i så vidt muligt alle dets forekommende former. Også aviser (fra 1990'erne frem til 2000) og exteriorica, dvs. ungarsk litteratur skrevet uden for Ungarn, er blevet indlemmet. For at sikre en så alsidig repræsentation af sproget som muligt er også mindre kendte og glemte forfatters skrifter medtaget i korpus, ligesom man som noget – i hvert fald i sammenligning med danske forhold – ganske enestående³⁰ har medtaget oversat tekst, dvs. ungarske oversættelser af udenlandske værker. Oversat tekst er også blevet excerperet til belægsamlingen; se f.eks. artiklen *árté* [blande sig], hvor der i linje seks henvises til kilden for et forudgående citat (det drejer sig om et værk af den engelske lyriker Robert Dodsley), idet det samtidig oplyses, at værket er oversat af Ferenc Faludi. Desuden er oversat tekst medtaget i den supplerende cd-rom-tekstsamling; jf. f.eks. artiklen *áزالékállatka* [infusionsdyr], hvor der i linje 21 henvises til kilden for et forudgående citat (det drejer sig om Alfred Edmund Brehms værk om dyreriget), idet det samtidig oplyses (dog uden angivelse af oversætterens

navn), at værket er en oversættelse. Der er i alt blevet brugt ca. 22.000 tekstsekvenser hidrørende fra ca. 2500 skribenter. I de tilfælde, hvor der er tale om oversat tekst, er det oversætteren, der regnes for ophavsmanden (om end også den oprindelige ophavsmand kan angives).

Den digitale tekstmasse kan nedhentes vederlagsfrit fra internettet på <http://www.nytud.hu/hhc/> Den vil i løbet af arbejdet fremover blive opdateret, og kildefortegnelsen vil derfor blive udvidet i form af mindre, supplerende fortegnelser bagest i de fremtidige ordbogsbind (Szabó, 2006: 54). Størstedelen af tekstmassen, af hvilken 82 % er skøn- og faglitteratur, er fra det 20. århundrede (nemlig omkring 16 millioner løbende ord), mens der er ca. halvt så meget fra det 19. århundrede og kun to millioner løbende ord fra udgangen af det 18. århundrede. Om korpussammensætningen og de overvejelser, der er blevet gjort i forbindelse med den, se Csengery (2006: 18-19).

Kildefortegnelsen, som fylder 732 sider(!), tjener to formål (Szabó, 2006: 53): For det første at gøre det muligt at genfinde eksempelsætningerne, dvs. belæggene, ikke mindst for at efterse rigtigheden af dem. For det andet at dokumentere de kilder, hvorfra belæggene hidrører. Med henblik på at optimere muligheden for genfindning – ikke mindst til glæde for metaleksikografer, (sprog)historikere og andre forskere – er samtlige kilder blevet anført med omfattende og præcise bibliografiske oplysninger i nøglebindet, hvor kildefortegnelsen findes på siderne 65-796. Til sammenligning fylder kildefortegnelsen i *Den Danske Ordbog* (Hjorth et al. 2003-2005) 42 sider (s. 628-669 i sjette bind) og fremviser desuden en aldeles utilstrækkelig kildehenvisningspraksis (jf. kritikken hos især Bergenholtz/Vrang, 2006: 188). Interessant i sammenligning med *Den Danske Ordbog* er det endvidere, at den ungarske akademiordbog, som rigtigt nok bruger digitale kilder (f.eks. udgivelser på cd-rom) ligesom *Den Danske Ordbog*, i modsætning til denne ikke bruger online-kilder, da redaktionens definition af "publikation" ikke indbefatter disse (sml. hertil Bergenholtz/Vrang 2006: 188, som også kritiserer, at *Den Danske Ordbog* overhovedet bruger online-kilder).

I forbindelse med kilderne og deres gengivelse er man fra akademiordbogsredaktionens side desuden gået tekstkritisk til værks. Om de nøje overvejelser mht. benyttede udgaver, gengivelse af sted- og personnavne, ortografi mm. se Csengery (2006) og Szabó (2006).

Databearbejdning

Siden databearbejdningen begyndte på Commodore 64-maskiner, var det med de givne tekstredigeringsmuligheder allerede fra begyndelsen klart, at de i ungarsk brugte diakritiske tegn og historiske bogstaver – såsom *ű*, dvs. *ü* med accent aigu, mm. – ikke var mulige at lagre digitalt. Ved hjælp af et til formålet udformet særegent kodesystem bestående af dels arabertal og dels det engelske alfabet blev disse tegn og bogstaver

²⁹ Da.: Historisk tekstkorpus over ungarsk.

³⁰ Der findes os bekendt ingen danske hverken helt eller delvis korpusbaserede ordbøger, hvor man har ladet oversat tekst indgå i korpus; således har praksis været i hvert fald i *Den Danske Ordbog* (Hjorth et al., 2003-2005), Bergenholtz (1992) og Maegaard/Ruus (1981a; 1981b). Denne fremgangsmåde forsvares typisk med argumentet om, at oversat sprog ikke er autentisk sprog. Vi er stærkt uenige i denne opfattelse, idet vi mener, at man netop, når man har med oversættelser at gøre, tvinges ud i de af sprogets afkroge, hvor man ellers aldrig kommer. Om det problematiske ved begrebet "autentisk sprog" se desuden Pálfi (2005: 144).



markeret (a1 = á; e1 = é; o2 = ö; u2 = ü; o3 = ő osv.). Således kunne teksterne altid konverteres entydigt og let til den på det givne tidspunkt brugte computer. Siden 1980'erne har de informationsteknologiske ressourcer naturligvis gennemgået en gradvis fornyelse.

Ved projektets begyndelse var forestillingerne om, hvilken form for databearbejdning der foruden selve tekstbearbejdningen ville være brug for under ordbogsredigeringen, uheldigvis meget uklare. Dette skulle senere vise sig at volde alvorlige problemer ved udarbejdelsen af det bibliografiske apparat. Ud over den rene tekstbearbejdning blev et særligt, enkelt tegnsæt taget i anvendelse til markering af afsnit, versgrænser, notaters udsondring mm. Da redaktionen i slutningen af 1980'erne blev bekendt med SGML-kodning og senere med XML efter anbefalinger fra TEI (Text Encoding Initiative), prøvede man ved hjælp af konverteringsprogrammer at omforme og supplere de data, som på det tidspunkt forelå i bearbejdet stand.

Morfologisk analyse

Redaktionen fandt det fra begyndelsen vigtigt, at man i de bearbejdede tekster ikke kun skulle kunne søge tekstord, men også leksemer. Et sådant ønske er imidlertid ikke let at imødekomme, når man har at gøre med et sprog som ungarsk, hvor de grammatikalske forhold normalt udtrykkes via sammensætningen af en lang række endelser (noget, som er kendetegnende for de såkaldt agglutinerende sprog under hvilke ungarsk henhører). En betydelig del af de hyppigt brugte rødder ændrer form foran bestemte affikser (f.eks. *ló* – *lovak* [hest (sing.) – hest (plur.)], *alszik* – *aludtam*, *alvó* [sove (3. pers. sing., præes.) – sove (1. pers. sing., præet.), sove (præs. part.)], *bokor* – *bokrot* [busk (nominativ) – busk (akkusativ)], *fa* – *fát* [træ (nominativ) – træ (akkusativ)] osv.). Hvis man således søger i de ikke-analyserede korpusstekster, skal man ved bestemte ord finde samtlige rodvarianter enkeltvis, og hvis man søger ord, som begynder med bestemte bogstavrækker, får man undertiden mange forekomster, som ikke passer på det søgte (f.eks. får man ved ord som begynder med *áll* [stå (3. pers. sing., præes.)] tillige ordene *állam* [drøm], *állat* [dyr] osv.). Derfor udvikledes som noget af det første et program til analyse af ungarsk morfologi. Ikke alene viste programmet sig at være aldeles fremmede for ordbogsarbejdet, men det kom også til at udgøre fundamentet i den stavekontrol, som senere indgik og stadig indgår i den ungarske version af Winword (Pajzs, 1991; Prószéky, 1996; Prószéky/Kis, 1999). Programmet er i stand til korrekt at genkende ganske udviklede suffiks-sammensætninger, det bruger en rodbase og en suffiksbase, kender de rigtige sammensætninger og betragter kun disse som værende grammatiske. Ved udgangen af 1990'erne videreudvikledes programmet i en retning, så det kunne genkende flere og flere af de gamle ordformer korrekt.

Et andet selvstændigt program undersøgte de regelmæssige afvigelser i ortografien i de gamle tekster og omformede de ellers ikke analyserbare ord til analyser-

bare former, som efterfølgende skulle analyseres. Hvis programmet fandt ordformer, som forekom at være korrekte, opfattedes disse som genkendelige former. Et eksempel skal tydeliggøre forholdet: Ordet *büszke* [stolt] blev i kodesystemet noteret *bu2szke*. I de gamle tekster forekommer imidlertid ofte en form af *ü*-bogstavet, som man ikke længere bruger i ungarsk, og denne form markeredes under tekstbearbejdningen som *bu20szke*. Men af denne markering fremgår det ikke entydigt, om nutidens retskrivning i dette tilfælde ville fordre kort eller lang vokal (*ü* eller *üi*), derfor fremviser konverteringsprogrammet i første omgang begge mulige varianter: *bu2szke* og *bu3szke*. De to varianter bliver efterfølgende analyseret af det morfologiske analyseprogram, og i tilfælde af korrekt analyse vil programmet foreslå den første variant som værende rigtig, eftersom det er den, der svarer til nutidens ortografi. I en betydelig del af tilfældene har løsningen dog ikke været nær så enkel, eftersom mange af ordene modsvarer talrige varianter i nutidssproget, hvorfor det ikke har været muligt automatisk at bestemme, hvilken løsning der var den mest hensigtsmæssige. Under alle omstændigheder har denne fremgangsmåde øget antallet af analyserbare og genkendelige tekstord betydeligt.

Entydighed

En meget stor del af tekstordene (omkring 30 %) kan af programmet analyseres som værende korrekt på mindst to måder: F.eks. kan ordet *várna* være flertalsformen i tredje person af verbet *vár* [vente] eller dativformen af substantivet *vár* [borg]; ordet *volt* kan være entalsformen i tredje person af verbet *van* [være] i datid eller substantivet *volt* [væren] eller evt. dannet som verbaladjektiv af verbet *van*. For at få en om muligt automatisk opløsning af disse flertydigheder udvikledes et program, som ved hjælp af såkaldt lokale regler prøvede at bestemme, hvad der i de givne tekstsammenhænge sandsynligvis var det mest rigtige. Hvis f.eks. de ovennævnte flertydige ord havde det bestemte kendeord *a* foran [svarer til de danske substantivendelser *-en/-et*], regnedes de af programmet for substantiver, men hvis der ikke var andre verber i den pågældende ledsætning, regnedes de for verber. Ved brug af talrige lignende lokale regler lykkedes det at reducere antallet af flertydige ord betydeligt.

Dette arbejde har til en vis grad tjent som forberedelse til de korpuslingvistiske og sprogteknologiske projekter, som søsattes i anden halvdel af 1990'erne i Ungarn.

Udvælgelsen af opslagsord

På grundlag af det analyserede og entydiggjorte korpus fremlagdes en liste, som indeholdt alle de i korpus forekommende og (potentielt) genkendelige ord med angivelse af deres hyppighed samt den korpusinterne kronologisk set tidligste og seneste forekomst forsynet med årstalsangivelse. Da den første liste var færdig i 1997, indeholdt den lidt mere end 160.000 (vordende) opslagsord. Det er ikke lang tid siden, at den aktuelle liste er blevet færdiggjort; denne indeholder 900.000 kandidater til opslagsord. Heraf skal leksikograferne



udvælge de ca. 110.000 opslagsord, som skal indgå i ordbogen.

Den gamle liste har under alle omstændigheder været til stor hjælp, idet den gav en vis forestilling om tilstrækkeligheden af det hidindtil sammensatte korpus. Bl.a. er det ad denne vej, man nåede til erkendelsen af, at den planlagte akademiordbog ikke kunne udarbejdes alene på grundlag af det på det tidspunkt forhåndenværende korpus, da det indeholdt alt for få data. Der var i alt kun 60.000 ord, om hvilke der fandtes mindst fem forskellige oplysninger i korpus. Ved udformningen af prøveartiklerne har man derfor været nødt til at støtte sig til belægsamlingen. I 2002, da udarbejdelsen af ordbogsartiklerne begyndte at tage fart, blev kilde-materialet udvidet med en supplerende tekstsamling i digital form, nemlig de ovenfor nævnte cd-rom-udgivelser.

Redigeringsprogrammet

Allerede de i 1990'erne udarbejdede prøveartikler blev skrevet i XML, men siden har redaktionen afprøvet flere programmer, og i dag bruges redigeringsprogrammet XMetal, som af de hidtil afprøvede programmer har vist sig at være det mest pålidelige og mest hensigtsmæssige. Jævnside med udarbejdelsen af retningslinierne for ordbogsredigeringen formuleredes en dokumenttypedefinition, som bestemmer, hvilke elementer en ordbogsartikel skal indeholde og i hvilken rækkefølge, hvilke der er nødvendige, og hvilke der er valgfrie eller betinget af tilstedeværelsen af bestemte andre elementer. Dokumenttypedefinitionen er blevet udformet således, at den i givet fald løbende kan modificeres.

I løbet af artikelskrivningen har det været muligt – ligesom det også nu er og fremover vil være det – at se flere udgaver af den givne artikel og at udskrive den på papir. Efter den første bearbejdning kan man f.eks. udskrive et udkast, som kun indeholder f.eks. betydningsforklaringerne uden eksempelsætninger. Ordbogsredaktøren kan på denne måde lettere få et overblik over artiklens skelet, og han/hun kan foreslå ændringer i det endnu ikke endeligt fastlagte udvalg af eksempelsætninger.

Fordelene ved XML-formatet er blevet udnyttet til fulde i forbindelse med udarbejdelsen af det første egentlige ordbogsbind: Det hørte til det daglige arbejde at gennemgå databasen med henblik på kontrol og om nødvendigt foretage konverteringer. Ud af det færdige XML-dokument kunne der i løbet af et par uger fremstilles et tryklart, ombrudt manuskript, som naturligvis også indeholdt de i citaterne forekommende historiske bogstaver.

Artikelskrivning

Artikelskribenten gennemser alle de kilder, i hvilke de afmærkede vordende opslagsord eller den pågældende ordfamilie findes: 1) De vigtigste etsprogsordbøger; 2) belægsamlingen; 3) det elektroniske tekstkorporus; og 4)

den supplerende tekstsamling på cd-rom. På baggrund af de herved fremfundne oplysninger identificeres og defineres de enkelte betydninger, hvorefter det tilstræbes at illustrere hver betydning med to til tre passende eksempelsætninger fra de nævnte kilder – og helst sådan at der opnås størst mulig spredning og variation mht. tid og genre. Samtidig benytter skribenten XMetal-redigeringsprogrammet og korpusprogrammerne. De eksempelsætninger, som er blevet fundet i den digitale tekstbestand, bliver sammen med de bibliografiske oplysninger kopieret direkte over i ordbogsartiklen, mens de eksempelsætninger, der er hentet fra seditalsamlingen, indskrives manuelt. Det første udkast til ordbogsartiklen udskrives og gives til ordbogsredaktøren. Denne gennemser artikeludkastet, efterser de oprindelige kilder, foreslår eventuelle ændringer – det være sig angående selve artikelopbygningen, udformningen af definitionen/definitionerne, valget af eksempelsætninger osv. – og giver det efterfølgende tilbage til artikelskribenten, som foretager de nødvendige foreslåede modificeringer, udskriver artikeludkastet på ny og giver det tilbage til redaktøren. Ved erfarne artikelskribenter er det som regel nok med to til tre rettelser hhv. ændringer, ved nybegyndere kan arbejdsgangen gentage sig adskillige gange, inden den endelige udgave af artiklen foreligger. Herefter gennemses hver artikel af tre sproglektorer, og de artikler, som vedrører et bestemt fagområde (f.eks. et håndværk, en videnskab, en sportsgren osv.), bliver tillige eftersat af nogle af de mange fagkonsulenter, som er tilknyttet ordbogsarbejdet. Til sidst gennemlæses hver eneste artikel af hovedredaktøren. Af XML-versionen udarbejdes ved hjælp af et program den ombrudte version, som inden udgivelsen naturligvis kommer igennem yderligere korrektureer.

Obygning og indhold

Der foreligger foreløbig kun et enkelt ordbogsbind, nemlig bind II, idet bind I er et nøglebind, i hvilket bl.a. kildefortegnelsen og ordbogsgrammatikken findes. Bind II indeholder 4.443 opslagsord i 4.500 selvstændige ordbogsartikler (at disse to tal ikke stemmer overens kommer af, at en del artikler er "affiksartikler"), 1.087 underopslagsord og 388 henvisningsord/-artikler. Desuden gives betydningsforklaringer til 537 ordforbindelser. I alt gives 14.346 betydningsangivelser og 47.533 eksempelsætninger. Hertil kommer to forkortelseslister – én over de redaktionelle og én over de i eksempelsætningerne forekommende forkortelser – samt anden omtækst.

Ud over det almindelige rigssprog medtages dialektalt ordgods, fagord og fagudtryk (dog i begrænset omfang), diastratiske varieteter (f.eks. slang), fremmedord, orddannelseselementer og sammensætningsformer og sidsteled, jf. Ittzés (2006: 26-27). Ikke medtaget er i reglen egennavne (med undtagelse af de tilfælde, hvor ordene har fået en sådan status i sproget, at de fleste sprogbrugere ikke længere tænker over dem som egennavne (f.eks. – her oversat til dansk – *sol, måne, bibel, parlament, forfatningsdomstol* og enkelte andre i talemå-



der fastfrosne navneformer) og fremmedord bestående af flere enkeltord (med undtagelse af ord, som har en vis udbredelse eller en bestemt betydning og/eller form (f.eks. *alma mater*, *a cappella*)), jf. hertil Ittzés (2006: 27).

Der findes tre slags artikler: 1) selvstændige artikler/opslagsord; 2) nicheartikler (såvel førsteleds- som sidsteledsartikler; jf. f.eks. *acél*- [stål-], hvor man finder følgende 20 grundigt behandlede underopslagsord (her oversat til dansk og oplistet i alfabetisk rækkefølge): *-afstøbning*, *-blå*, *-fjeder*, *-grøn*, *-grå*, *-hestesko*, *-højovn*, *-hånd*, *-industri*, *-klinge*, *-muskel*, *-plade*, *-rør*, *-skelet*, *-spejl*, *-streng*, *-støberi*, *-tråd*, *-valseværk*, *-værk*; og *-arány* [-forhold], hvor der ligeledes gives en række underopslagsord, som her dog skrives helt ud: *indkomstforhold*, *talforhold* mm.); og 3) henvisningsord/-artikler (som både kan være selvstændige eller gives ved udgangen af en anden artikel); henvisningsstrukturen³¹ kommer vi kort tilbage til nedenfor.

Ordbogens første bind indeholder endvidere en ordbogsgrammatik (s. 803-1114) forsynet med en kort brugervejledning/indledning (s. 799-801). I den ordbogsgrammatiske del bringes flere hundrede tabeller over verbernes, adjektivernes og substantivernes bøjningsmønstre, dvs. tabeller over såvel konjugation, komparation som deklination. Således henvises brugeren f.eks. fra artiklen *alszik* [sove] med koden "19a1" til side 962-963 i ordbogsgrammatikken, hvor det pågældende verbers fuldstændige bøjningsparadigme vises.

Ud fra antallet af forekomster i det 27,5 millioner løbende ord omfattende elektroniske tekstkorpus angives samtlige opslagsords hyppighed. Der skelnes imellem syv grader af hyppighed illustreret ved et måneformet piktogram: Tom måne betyder nul forekomst i korpus, dvs. at ordet er sjældent brugt i ungarsk; fuldmåne betyder høj forekomst, nemlig mindst 100.000 forekomster i korpus, dvs. at ordet er hyppigt brugt.

I en artikel angives ud over opslagsord (og evt. underopslagsord): 1) homonymopløsning; 2) ordklasse; 3) bøjningsparadigme (i form af eksplicit og direkte henvisning til den pågældende tabel i ordbogsgrammatikken); 4) diasystematisk markering³²; 5) hyppighed; 6) opslagsordsvariant (se *azsúr*² [ajour], hvor varianten *à jour* angives); 7) betydningsinddeling og -forklaring; 8) latinsk betegnelse (ved navne på dyr, planter og insekter); 9) valens hhv. konstruktionsoplysninger (dog ikke så konsekvent og ikke i samme udstrækning som f.eks. i *Den Danske Ordbog* (Hjorth et al. 2003-2005)); 10) talemåde(r) forsynet med betydningsforklaring(er); 11) citater (med angivelse af det givne dokument side,

³¹ Med *henvisningsstruktur* menes strukturen af de eksplicite og implicite indikatorer, som leder ordbogsbrugeren rundt i ordbogen for at supplere eller uddybe de allerede fundne oplysninger (Bergenholtz et al., 1997: 134).

³² Med *diasystematisk markering* menes markering af, at en leksikalsk enhed på en eller anden måde tydeligt og systematisk afviger fra hovedmængden af de øvrige leksikalske enheder i ordbogen (Bergenholtz et al., 1997: 99).

ophavsmand (i tilfælde af oversættelse angives i reglen såvel oversætter som forfatter) og udgivelsestidspunkt); 12) bibliografisk apparat (eksplicit og direkte henvisning til kildefortegnelsen); 13) henvisning til et andet regulært opslagsord eller en nicheartikel; 14) henvisning til en anden artikel, hvor det pågældende opslagsord indgår i en talemåde; 15) leksikografisk afhjæmning (dvs. henvisning til andre, tidligere udgivne ordbøger, som også har optegnet det pågældende ord).

Ifølge Bergenholtz (1995a: 9) kan der skelnes mellem omkring tredive forskellige ordbogsstrukturer. Af pladsmangel kan vi her kun kommentere nogle af de vigtigste, og omtalen vil af samme grund være meget kort.

Bortset fra at der gives en større mængde nicheartikler, konstitueres makrostrukturen³³ af den glatalfabetiske opslagsordsliste med initialalfabetisk præsentationsform ordnet strengt alfabetisk efter bogstav for bogstav-princippet (også kaldet tegn for tegn-princippet).

Henvisningsstrukturen er alt for kompleks til, at pladsen i denne artikel tillader en blot tilnærmelsesvis tilstrækkelig udredning. Vi må nøjes med en kort opremsning af de strukturelle elementer: Der findes syv former for eksplicite henvisninger: 1) henvisningsartikler bestående af et henvisningssted og en henvisningsadresse (mellem disse er der en "L.", forkortelse for "lásd" [se]); 2) henvisningsartikler med angivelse af ordklasse og med henvisning til nicheartikler (førsteledsartikler); 3) henvisning til bøjningsparadigme i ordbogsgrammatikken; 4) artikelfinale henvisninger indledt med forkortelsen *Vő.*, som henviser til andre leksikografiske opslagsværker (det drejer sig især om de i denne artikel tidligere omtalte værker); 5) artikelfinale henvisninger indledt med forkortelsen *Ö.*, som henviser til selvstændige nicheartikler (førsteledsartikler); 6) *Fr.*, fraseologisk henvisning, dvs. henvisning til et andet opslagsord, som udgør kerneordet i en talemåde, som det opslagsord, der bliver henvist fra, indgår i (sml. m. angivelse nr. 14 for oven); 7) ordbogseksterne henvisninger i form af kildefortegnelsen i bind I samt i selve artiklerne, idet mange af de i litteraturlisten anførte titler også anføres her.

Angående den ligeledes komplekse mikrostruktur³⁴ må vi her nøjes med at henvise til afsnittet om artikelopbygningen længere oppe i nærværende kapitel.

Afsluttende bemærkninger

Som vi skrev i begyndelsen af denne artikel, har vi valgt at afholde os fra at give en kritik af akademiordbogen. Men eftersom værket ifølge planen først vil være afsluttet om ca. 25 år og ifølge al erfaring sikkert

³³ Med *makrostruktur* menes rækkefølgen af og relationerne imellem opslagsord, som indgår i en bestemt opslagsordsmængde (Bergenholtz et al., 1997: 178).

³⁴ Med *mikrostruktur* menes rækkefølgen af og relationen mellem de angivelser, der er adresseret til et opslagsord hhv. en artikel.



først om meget længere tid, vil der være rig lejlighed for leksikografer verden over til at diskutere og kritisere det i anmeldelser, hvorfor en sådan mangel, som vi her efterlader, sikkert snart vil blive afhjulpet andetsteds. Der vil man formentlig også kunne finde en grundigere udredning af ordbogens byggede, strukturer, funktioner og målgruppe(r), – emner, som kalder på grundig behandling i netop en anmeldelse – end hvad det af pladmæssige men også konceptuelle grunde har været muligt at give her.

Ud fra ønsket om at give et adækvat billede af ungarsk modersmålsleksikografi fra begyndelsen til nyeste tid, har vi også afholdt os fra at inddrage den moderne leksikografiske funktionslære, eftersom denne teori stort set er ukendt eller i hvert fald ikke etableret i de leksikografiske miljøer i Ungarn.

Den, som interesserer sig for ungarsk leksikografi – og i øvrigt behersker sproget – kan få et fuldstændigt overblik over ungarsk ordbogsskrivning i en nyligt udkommen bibliografi (Magay, 2004), som regelmæssigt opdateres i de nye numre af *Lexikográfiai füzetek*³⁵, som udgives af Akadémiai Kiadó. En bibliografi over ungarsk metaleksikografisk litteratur er også på vej, den befinder sig på nuværende tidspunkt i planlægningsfasen og forventes at udkomme om nogle år.

Loránd-Levente Pálfi er forskningsmedarbejder ved Center for Leksikografi, Handelshøjskolen, Aarhus Universitet.

Júlia Pajzs, ph.d., er seniorforsker ved Afdeling for Leksikologi og Leksikografi, Institut for Sprogvidenskab, Det Ungarske Videnskabernes Selskab.

Litteratur

Bakos, F.; Ruzsicky, É.; Szávai, J. (1994): *Magyar Larousse. Enciklopédikus szótár I-III* [Ungarsk Larousse. Encyklopædisk ordbog]. Budapest: Akadémiai Kiadó.

Bárczi, G. et al. (1959-1962): *A magyar nyelv értelmező szótára I-VII* [Betydningsforklarende ordbog over det ungarske sprog]. Udgivet af Magyar Tudományos Akadémia, Nyelvtudományi Intézet [Det Ungarske Videnskabernes Selskab, Institut for Sprogvidenskab]. Budapest: Akadémiai Kiadó.

Bárczi, G. et al. (2004): *A magyar nyelv értelmező szótára I-VII* [Betydningsforklarende ordbog over det ungarske sprog]. Udgivet af Magyar Tudományos Akadémia, Nyelvtudományi Intézet [Det Ungarske Videnskabernes Selskab, Institut for Sprogvidenskab]. Budapest: Arcanum CD.

Bárdosi, V. (2003): *Magyar szólástár. Szólások, helyzetmondatok, közmondások értelmező és fogalomköri szótára* [Ungarsk talemådeordbog. Betydningsforklarende ord-

bog over talemåder, fraser og ordsprog]. Budapest: Tinta Könyvkiadó.

Baróti, D. Sz. (1784): *Kisded szótár, mely a ritkább magyar szókat az abc rendi szerént emlékeztető versekben előadja* [Lille ordbog, som optegner de mere sjældent brugte ord og ordner dem alfabetisk i vers for at gøre dem lettere at huske]. Kassa.

Benkő, L. (1967-1984): *A magyar nyelv történeti-etimológiai szótára I-IV* [Historisk-etymologisk ordbog over det ungarske sprog]. Budapest: Akadémiai Kiadó.

Benkő, L. (1993-1997): *Etymologisches Wörterbuch des Ungarischen I-III*. Budapest: Akadémiai Kiadó.

Bergenholtz, H. (1988): "Empiriske metoder i sprogvidenskabelig forskning". *Hermes* 1: 7-23.

Bergenholtz, H. (1992): *Frekvensordbog. Baseret på tekster fra danske romaner, ugeblade og aviser 1987-1990*. København: Gad.

Bergenholtz, H. (1995a): "Grammatik i bilingvale ord-bøger". *LexicoNordica* 2: 5-18.

Bergenholtz, H. (1995b): "Leksikografi. Hvad er det?" In: Svavarsdóttir, A.; Kvaran, G.; Jónsson, J.H. (red.): *Nordiske studier i leksikografi 3. Rapport fra Konferanse om leksikografi i Norden. Reykjavík 7.-10. juni 1995*. Reykjavik: Nordisk forening for leksikografi: 37-49.

Bergenholtz, H. (1996): "Grundfragen der Fachlexikographie". In: Gellerstam, M.; Järborg, J.; Malmgren, S.-G.; Norén, K.; Rogström, L.; Papmehl, C.R. (eds.): *Euralex '96. Proceedings I-II. Papers submitted to The Seventh EURALEX International Congress on Lexicography in Göteborg, Sweden*. Göteborg: Göteborg University: 731-758.

Bergenholtz, H.; Cantell, I.; Fjeld, R.V.; Gundersen, D.; Jónsson, J.H.; Svensén, B.; Haraldsson, H.; Mikkelsen, H.K.; Sivula, J. (1997): *Nordisk leksikografisk ordbok*. Oslo: Universitetsforlaget.

Bergenholtz, H. (1999): "Das Schlaue Buch. Vermittlung von Informationen für textbezogene und textunabhängige Fragestellungen". In: Zettersten, A.; Mogensen, J.E.; Hjørnager Pedersen, V. (eds.): *Proceedings of the Eighth International Symposium on Lexicography May 2.-5., 1998 at the University of Copenhagen*. Tübingen: Max Niemeyer: 93-110.

Bergenholtz, H.; Gouws, R. (2006): "How to do policy with dictionaries". *Lexikos* 16: 13-45.

Bergenholtz, H.; Pálfi, L.-L. (2008): "Danmarks første nationalordbog: Om Videnskabernes Selskabs Ordbog I-VIII (1793-1905)". *Fund og Forskning* (submitted).

³⁵ Da.: Leksikografiske hæfter.



- Bergenholtz, H.; Vrang, V. (2006): "Den Danske Ordbog: en ordbog for lingvister!" *LexicoNordica* 13: 185-196.
- Csengery, K. (2006): "Az elektronikus korpusz". In: Ittész et al.: *A magyar nyelv nagyszótára I. Segédletek*: 18-19.
- Csepregi, M. (red. 1998): *Finnugor kalauz*. Budapest: Panorama.
- Duden. *Das große Wörterbuch der deutschen Sprache in zehn Bänden*. 3., völlig neu bearbeitete und erweiterte Auflage. Herausgegeben vom Wissenschaftlichen Rat der Dudenredaktion. Mannheim/Leipzig/Wien/Zürich: Dudenverlag 1999.
- Dyrberg, G.; Faber, D.; Hansen, S.L.; Tournay, J. (1988): "Etablering af et juridisk tekstkorpus". *Hermes* 1: 209-227.
- Élesztős, L.; Rostás, S.; Glatz, F. (1993-2004): *Magyar nagylexikon I-XVIII* [Den store ungarske encyklopædi]. + Kiegészítő kötet [supplementsbind]: *Magyar nagylexikon XIX*. Budapest: Akadémiai Kiadó.
- Eöry, V. (red.) (2007): *Értelmező szótár+ I-II* [Betydningsforklarende ordbog+]. Budapest: Tinta Könyvkiadó.
- Forgács, T. (2003): *Magyar szólások és közmondások tára* [Ordbog over ungarske talemåder og ordsprog]. Budapest: Tinta Könyvkiadó.
- Forgács, T. (2007): *Ungarische Grammatik*. Wien: Praesens.
- Gáldi, L. (1960): "Mutatvány A Magyar Irodalmi Nyelv Nagyszótárából". *Magyar Nyelvőr*: 182-196.
- Czuczor, G.; Fogarasi, J. (red.) (1862-1874): *A magyar nyelv szótára I-VI* [Ordbog over det ungarske sprog]. [1862: Bind 1: A-D; 1864: Bind 2: E-H; 1865: Bind 3: I-L; 1867: Bind 4: M-Ö; 1870: Bind 5: P-S; 1874: Bind 6: T-Z]. Udgivet af Magyar Tudományos Akadémia [Det Ungarske Videnskabernes Selskab]. Pest: Emich Gusztáv.
- Gerstner, K. (2006): "A magyar nyelv nagyszótárának áttekintő története". In: Ittész et al.: *A magyar nyelv nagyszótára I. Segédletek*: 10-17.
- Grimm, J.; Grimm, W. et al. (1854-1971): *Deutsches Wörterbuch*. 16 Bde. in 32 Bdn. + Bd. 33 [Quellenverzeichnis]. Leipzig: Hirzel.
- Hjorth, E.; Kristensen, K.; Norling-Christensen, O. (red. 2003-2005): *Den Danske Ordbog 1-6*. Udgivet af Det Danske Sprog- og Litteraturselskab. København: Gyldendal.
- Imbs, P. et al. (1971-1994): *Trésor de la langue française. Dictionnaire de la langue de XIX^e et du XX^e siècle*. 1-16. Ed.
- Du Centre National de la Recherche Scientifique, Centre de Recherche pour un Trésor de la Langue Française : Nancy, Paris.
- Ittész, N. (2006): "Tájékoztató a szótár szerkesztési elveiről, szerkezetéről és használatának módjáról". In: Ittész et al.: *A magyar nyelv nagyszótára I. Segédletek*: 23-50.
- Ittész, N. et al. (2006ff): *A magyar nyelv nagyszótára* [Storordbog over det ungarske sprog]. [Bind] I: *Segédletek* [(nøglebind med kildefortegnelse mm.), 1119 sider]. [Bind] II: *A-Azsiúroz* [1550 sider]. Budapest: Udgivet af Magyar Tudományos Akadémia, Nyelvtudományi Intézet [Det Ungarske Videnskabernes Selskab, Institut for Sprogvidenskab]. Budapest.
- Juhász, J.; Szőke, I.; Nagy, G. Ó. (1972): *Magyar értelmező kéziszótár* [Betydningsforklarende håndordbog over ungarsk]. Budapest: Akadémiai Kiadó.
- Kis, J.; Pusztai, F. (red.) (2005): *Magyar Nyelvtörténet*. Budapest: Osiris Könyvkiadó.
- Kiss, G. (1998): *Magyar szókinctár. Rokon értelmű szavak, szólások és ellentétek szótára* [Ungarsk begrebsordbog. Ordbog over betydningsbeslægtede ord, talemåder og betydningsmodsatte ord]. Budapest: Tinta Könyvkiadó.
- Kiss, L. (1994): "Nyelvtörténeti szótáraink típusai". In: *A lexikológia és lexikográfia elmélete és módszertana*. Budapest: Magyar Nyelvtudományi Társaság: 13-32.
- Kiss, L.; Pajzs, J. (1989): "A magyar irodalmi és köznyelv nagyszótára. 1533-1990". *Magyar Nyelv*: 129-136.
- Ludwig, K.-D. (1998): "Wörterbücher als Spiegel gesellschaftlicher Veränderungen". In: Zettersten, A.; Hjørnager Pedersen, V.; Mogensen, J.E. (eds.): *Symposium on Lexicography VIII. Proceedings of the Eighth International Symposium on Lexicography May 2.-4., 1996, at the University of Copenhagen*: 49-64.
- Lundblad, C.-E. (1999): "Ordbok eller encyklopedi – en fråga om hänsyn till användaren?" In: Slotte, P.; Westerberg, P.; Orava, E. (red.): *Nordiska studier i lexicografi IV. Rapport från Konferensen om lexicografi i Norden. Esbo 21.-24. maj 1997*. Helsingfors: Nordisk forening for lexicografi: 265-273.
- Lőrinczy, É.B. (red. 1979-2002[ff?]): *Új magyar tájszótár I-IV* [Ny ungarsk dialektordbog]. Udgivet af Magyar Tudományos Akadémia, Nyelvtudományi Intézet [Det Ungarske Videnskabernes Selskab, Institut for Sprogvidenskab]. Budapest: Akadémiai Kiadó.
- Maegaard, B.; Ruus, H. (1981a): *Hyppige ord i danske børnebøger. 4000 hyppigste ord i danske børnebøger ordnet alfabetisk og efter hyppighed*. København: Gyldendal.



- Maegaard, B.; Ruus, H. (1981b): *Hyppige ord i danske romaner. 5000 hyppigste ord i de mest læste danske romaner ordnet alfabetisk og efter hyppighed*. København: Gyldendal.
- Magay, T. (2004): *A magyar szótáriróladalom bibliográfiája* [Bibliografi over ungarsk ordbogsskrivning]. Budapest: Akadémiai Kiadó.
- Malige-Klappenbach, H. (1989): "Sprache und Ideologie, insbesondere bei der Wörterbucharbeit". *Muttersprache* 99/2: 153-159.
- Ország, L. (1960): "Problems and principles of the new dictionary of the Hungarian language". *Acta Linguistica* X/3-4: 211-273.
- Pajzs, J. (1991): "The Use of a Lemmatized Corpus for Compiling the Dictionary of Hungarian". In: *Using Corpora. Proceedings of the 7th Annual Conference of the OUP & Centre for the New OED and Text Research*. Waterloo: University of Waterloo: 129-136.
- Pálfi, L.-L. (2005): "Den Danske Ordbog – en kritisk opsamling af den offentlige ordstrid". *LexicoNordica* 12: 123-148.
- Pálfi, L.-L. (2007): "Hvad er leksikografi?". *Referencen* 37/1: 23-29.
- Pálfi, L.-L.; Bergenholtz, H. (2007): "Dansk Glossarium – til at hjælpe de af Lægfolk tilrette, som gjerne gad læst en dansk Bog". *Nomos* 5/1: 55-91.
- Pedersen, K.M. (1994): "Ordbog og encyklopædi". In: Garde, A.; Jarvad, P. (red.): *Nordiske Studier i Leksikografi II. Rapport fra Konference om Leksikografi i Norden*: 248-256.
- Prószycki, G. (1996): "HUMOR – A Morphological System for Corpus Analysis". In: *Proceedings of the first TELRI Seminar in Tihany*. Budapest: Research Institute for Linguistics: 149-158.
- Prószycki, G.; Kis, B. (1999): "Agglutinative and Other (Highly) Inflectional Languages". In: *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*. Maryland: College Park: 261-268.
- Pusztai, F. (red. 2003): *Magyar értelmező kéziszótár* [Betydningsforklarende håndordbog over ungarsk]. [2. udgave]. Budapest: Akadémiai Kiadó.
- Révai nagy lexikona. *Az ismeretek enciklopédiája I-XXI* [Révais konversationsleksikon] (1911-1935). Budapest: Révai Testvérek.
- Imre, S. (1977): "A mai magyar nyelvtudomány néhány szintézise". *Magyar Nyelv*: 279-287.
- Sinor, D. (ed.) (1988): *The Uralic Languages. Description, History and Foreign Influences*. Leiden/New York/København/Köln: E.J. Brill.
- Szabó, T.P. (2006): "Tájékoztató a forrás- és a névjezőgyékek használatához". In: Ittész et al.: *A magyar nyelv nagyszótára I. Segédletek*: 53-64.
- Szarvas, G.; Simonyi, Zs. (1890-1893): *Magyar nyelv-történeti szótár. A legrégibb nyelvémlékektől a nyelvújításig. Lexicon linguae Hungaricae aevi antiquioris. I-III*. [Sproghistorisk ordbog over ungarsk. Fra de ældste sprogmindesmærker til sprogfornelsen [≈ Ordbog over old- og middelungarsk]]. [1890: Bind 1: A-I; 1891: Bind 2: J-S; 1893: Bind 3: Sz-Zs + Register]. Udgivet af Magyar Tudományos Akadémia [Det Ungarske Videnskabernes Selskab]. Budapest: Hornyánszky Viktor Akadémiai Könyvkereskedése.
- Szily, K. (1902-1908): *A magyar nyelvújítás szótára I-II* [Den ungarske sprogfornelses ordbog]. Budapest: Hornyánszky Viktor.
- T. Litovkina, A. (2005): *Magyar közmondástár. Közmondások értelmező szótára példákkal szemléltetve* [Ungarsk ordsprogsordbog. Betydningsforklarende og med eksempler belyst ordbog over ordsprog]. Budapest: Tinta Könyvkiadó.
- Wiegand, H.E. (Hrsg. 2003): *Untersuchungen zur kommerziellen Lexikographie der deutschen Gegenwartssprache*. Bind I. Tübingen: Max Niemeyer.
- Wierzbicka, A. (1995): "Dictionaries and Ideologies: Three Examples from Eastern Europe". In: Kachru, B.B.; Kahane, H. (eds.): *Cultures, Ideologies and the Dictionary. Studies in Honor of Ladislav Zgusta* (= Lexicographica Series Maior 64): 181-195.
- Zaicz, G. (2006): *Etimológiai szótár – magyar szavak és toldalékok eredete* [Etymologisk ordbog – ungarske ords og endelsers oprindelse]. Budapest: Tinta Könyvkiadó.