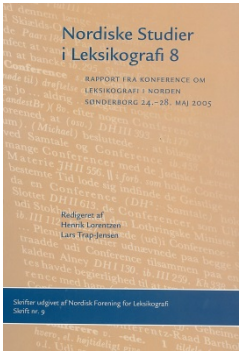


NORDISKE STUDIER I LEKSIKOGRAFI

Titel:	Multilingval elektronisk idiomordbog	
Forfatter:	Erla Hallsteinsdóttir	
Kilde:	Nordiska Studier i Leksikografi 8, 2006, s. 171-180 Rapport från Konferens om lexikografi i Norden, Göteborg 27.-29. maj 1999	
URL:	http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive	

© Nordisk forening for leksikografi

Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

Søgbarhed

Artiklerne i de ældre Nordiske studier i leksikografi (1-5) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

Multilingval elektronisk idiomordbog

A multilingual electronic dictionary of idioms. In this article, I present a research project on multilingual lexicography and phraseology. The aim of the project is to analyse linguistic and lexicographical theories and evaluate practical solutions in order to provide a basis for the conceptual work on and the compilation of an electronic German ↔ Icelandic ↔ Danish ↔ German dictionary of idioms.

1. Indledning: Projektets baggrund

Projektet multilingval elektronisk idiomordbog bygger på resultaterne fra et forskningsprojekt med en tysk-islandsk fraseologisk ordbog som blev finansieret af *The Icelandic Centre for Research*. Projektet har to overordnede formål: (1) At udarbejde et teoretisk koncept for multilingvale elektroniske ordbøger og (2) at omsætte konceptet til praksis i en tysk ↔ islandsk ↔ dansk ↔ tysk database med idiommer (med mulighed for udvidelse med fraseologismer (i bred forstand) og etordsleksemer).

2. Teoretisk grundlag

2.1. Fraseologi

Et af de vigtigste grundlag for fraseografisk arbejde dannes i den ensprogede fraseologiske grundforskning. Det drejer sig om udførlige ensprogede korpusbaserede analyser og beskrivelser af fraseologismer. Først når denne slags beskrivelser foreligger, kan man lave kontrastive analyser af fraseologiske relationer og udarbejde kontrastive fraseologikorpora, der så danner grundlaget for ordbøger. I dette projekt er der tale om tysk ↔ islandske, dansk ↔ islandske og dansk ↔ tyske fraseologismer. Det efterfølgende er en kort beskrivelse af det fraseografiske grundlag i de tre sprog.

2.1.1. Tysk fraseologi

Inden for den tyske fraseologiforskning findes der allerede et bredt teoretisk grundlag, der vil indgå i mit projekt (sml. udførlig litteraturliste i Hallsteinsdóttir 2001 og 2006).

Af aktuelle forskningsprojekter der fokuserer på en korpusbaseret eller kontrastiv beskrivelse af tyske idiommer, kan f.eks. nævnes projekterne “Kollokationen im Wörterbuch” ved Berlin-Brandenburgische Akademie der Wissenschaften (sml. Fellbaum et al. under udg. og kollokationen.bbaw.de/index.html) og “Syntax und Semantik idiomatisierter verbaler Phraseologismen in einem abgestuft kompositionellen Grammatikmodell” ved universitetet i Jena (sml. Soehn/Römer under udg. og

www.personal.uni-jena.de/~xcr/website/forschung/IVP-Web/main.htm). Yderligere er idiommer integreret som opslagsord i korpusordbogen “Projekt Deutscher Wortschatz” (www.wortschatz.uni-leipzig.de) ved universitetet i Leipzig. I EU-forskningsprojektet EPHRAS (sml. Jesenšek under udg. og www.ephras.org) ved universitetet i Maribor indgår tysk som et af sprogene ved udviklingen af en multilingual lærer-cd med fraseologismer.

Med hensyn til grundlaget for fraseografisk arbejde er der dog stadigvæk behov for forskning på bestemte områder, f.eks. vedrørende beskrivelsen af semantiske og pragmatiske egenskaber, afdækning af hhv. onomasiologiske og konceptuelle strukturer og beskrivelse af kontrastive relationer til islandsk og dansk.

2.1.2. *Islandsk fraseologi*

Inden for ensproget islandsk fraseologi og fraseografi har Jón Hilmar Jónssons forskning en særstilling. Sammen med Jón G. Friðjónssons (Friðjónsson 1993) og Halldór Halldórsson (Halldórsson 1991) værker danner hans arbejde (Jónsson 2002) et grundlag for beskrivelse af islandsk fraseologi i kontrastiv sammenhæng.

2.1.3. *Islandsk ↔ tysk fraseologi*

Bortset fra mine forskningsprojekter (Hallsteinsdóttir 2001, Hallsteinsdóttir under udg. a, b og c) og nogle korte artikler findes der kun en større teoretisk afhandling inden for kontrastiv tysk-islandsk fraseologi (Sverrisdóttir 1987). Her mangler altså stort set et grundlag for fraseografisk arbejde.

2.1.4. *Dansk fraseologi*

Inden for den danske fraseologi kan man konstatere øget forskningsinteresse i løbet af de sidste år, bl.a. med grundlæggelsen af Dansk Selskab for Fraseologi (DANFRAS). Der findes nyere oversigter over dansk fraseologiforskning og dansk fraseografi (f.eks. Farø 2004a+b og Farø under udg.), større idiomordbøger er udkommet og en elektronisk idiomordbog er blevet publiceret på internettet (www.idiomordbogen.dk) (sml. review article Farø 2004b og Almind et al. under forb.). Yderligere arbejdes der på et stort online-opslagsværk over det danske sprog, hvori beskrivelse af idiommer indgår (sml. Henrik Lorentzens og Lars Trap-Jensens artikel i denne publikation og www.ordnet.dk). Men alligevel er der stadig mange uløste problemer i forbindelse med systematiseringen og beskrivelsen af dansk fraseologi, eller, med Ken Farøs egne ord: “Der venter en opgave forude med at systematisere den fraseologiske forskning og beskrivelse i Danmark.” (Farø 2004b, 234).

2.1.5. *Tysk ↔ dansk fraseologi*

I den tysk-danske fraseologi findes nogle få nyere teoretiske og kontrastive undersøgelser (f.eks. Farø 2004a, Farø under udg. og Farø under forb.), der bl.a.

beskæftiger sig med ækvivalensproblemer og fraseografiske løsninger. Den danske og den dansk-tyske fraseologiforskning danner et forholdsvis solidt grundlag for videre forskning.

2.1.6. Dansk ↔ islandsk fraseologi

Selv om der i Island er en lang tradition for at undervise i dansk som fremmedsprog både i gymnasierne og i folkeskolerne, har der hidtil ikke været lagt nogen stor vægt på kontrastiv pædagogisk-leksikografisk forskning i disse to sprog. Bortset fra Auður Hauksdóttirs (Islands Universitet) nyere (p.t. ikke publicerede) undersøgelser i islandsk-dansk fraseologi findes der hverken teoretiske fraseologiske afhandlinger, større kontrastive arbejder eller fraseologiske ordbøger. Mit projekt er derfor et pilotprojekt i forbindelse med beskrivelsen af dansk-islandsk fraseologi.

2.2. Fraseografi

Mangler i fraseologiske ordbøger og forslag til udbedring af disse har været et hyppigt diskuteret tema i den tysksprogede fraseologiforskning inden for de sidste 20 år. Denne diskussion danner en bred teoretisk basis for fraseografisk arbejde. Diskussionen tog sin begyndelse med Werner Kollers forslag til en ideel tosproget idiomatisk lærerordbog (Koller 1987); dette forslag er delvis fulgt i trykte ordbøger, bl.a. findes onomasiologisk-tematisk opbygning (organisation efter betydning/koncepter) og øvelser i Regina Hesskys og Stefan Ettingers tyske onomasiologiske idiomordbog (Hessky/Ettinger 1997).

Først med muligheden for at bruge computere som leksikografisk medium er der kommet et arbejdsgrundlag for at realisere mange af de teoretiske forbedringsforslag. Det bliver muligt at bryde den todimensionale, lineære struktur trykte bøger har, op. Alligevel må man anse mange elektroniske ordbøger for at være biprodukter af trykte ordbøger. De opbygges efter koncepter for trykte ordbøger og har de samme begrænsninger, de udnytter altså ikke de muligheder computere byder på som medium (f.eks. ubegrænset plads, multimedia, hypertext, ikke-lineær opbygning).

Den teoretiske diskussion inden for elektronisk fraseografi bærer i stor grad præg af at forskningsområdet er under udvikling, og at der mangler teoretisk forarbejde i fraseologien og i leksikografien, f.eks. i forbindelse med ækvivalensproblematikken og spørgsmålet om grundlaget for lemmaselektion (sml. udførligt om fraseografiens problemstillinger i Hallsteinsdóttir 2006).

2.3. Metaleksikografisk arbejde

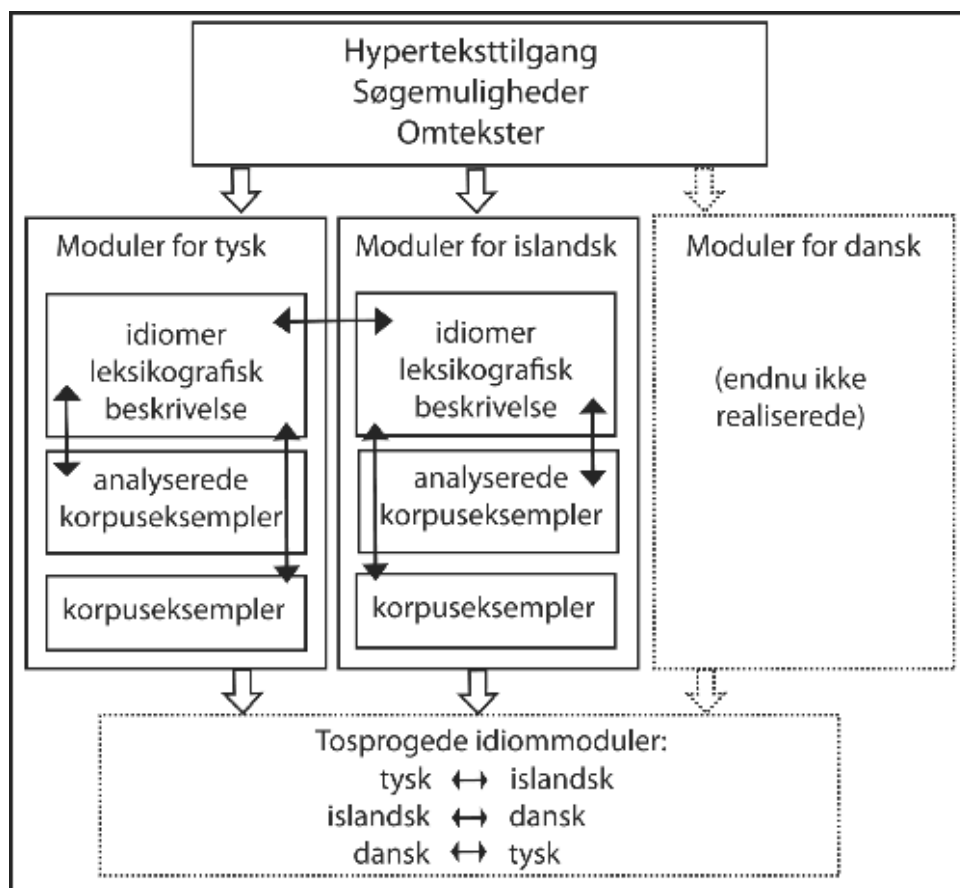
En stor del af arbejdet ved det aktuelle projekt består i metaleksikografiske overvejelser i forbindelse med udformningen af et leksikografisk koncept for en multilingval database, bl.a.:

1. design af multilingval databasestruktur,
2. design af ordbogens informationsenheder i forhold til ordbogens funktioner (i forbindelse med brugergrupper og ordbogsbrugernes behov) og
3. løsninger til ækvivalensproblematikken i multilingvale ordbøger.

2.3.1. Design af multilingval leksikografisk databasestruktur

Ved valget af databasesoftware skulle følgende kriterier kunne opfyldes:

1. Der skal være mulighed for databasemoduler for hvert sprog, dvs. den overordnede struktur skal være modulær (sml. figur 1), bl.a. med hensyn til senere udvidelse med flere sprog.



Figur 1. Overordnet leksikografisk struktur

2. Inden for modulerne skal strukturen være en ikke-lineær hypertextstruktur med en hierarkisk opbygning. Informationer inden for modulerne skal forbindes med hyperlinks.
3. Databaseprogrammet skal byde på forskellige muligheder for publicering, varierende tilgangsstruktur og layout, bl.a. forskellige søge- og rapportmuligheder.
4. De ikke-leksikografiske kriterier, brugervenlighed (der skal bruges minimal tid på tilegnelse af programmet) og pris (projektet har begrænset budget), spillede ligeledes en stor rolle da valget faldt på databasesoftwarens askSam (www.asksam.com).

2.3.2. Design af databasens informationsenheder

Design af ordbogens informationsenheder udarbejdes i forhold til ordbogens funktioner i forbindelse med brugergrupper og ordbogsbrugernes behov (se om ordbogsfunktioner i Bergenholtz 1997 og Tarp 1998). Selv om der findes en del teoretiske afhandlinger om problemet, er brug af ordbøger ud fra min erfaring tit meget individuel, for ikke at sige tilfældig, idet der hos de fleste brugere er tale om en selv lært kompetence. Ordbogsdidaktik er mere eller mindre ukendt i skolen, dvs. der finder nærmest ingen systematisk undervisning i ordbogsbrug sted. Alligevel prøver jeg at give et bud på hvordan man kunne strukturere de leksikografiske informationer i forhold til potentiel brug (sml. også Hallsteinsdóttir under udg. a og c).

Informationsstrukturen er opbygget efter ordbogsbrugernes potentielle behov for informationer for sprogreception (høre, læse), sprogproduktion (tale, skrive) og oversættelse. Der er tale om et informationshierarki som udfoldes ved hjælp af hyperlinks:

Niveau 1: Sprogreception.

Niveau 2: Sprogproduktion.

Niveau 3: Oversættelse af sprog.

I databasen bliver der yderligere taget højde for informationer til indlæring og korrektur, og den indeholder ligeledes en del metaleksikografiske informationer. Efterfølgende gennemgås databasens informationsenheder summarisk.

Databasen indeholder en del ensprogede informationer, såvel for sprogreception som sprogproduktion. Der er tale om:

- 1) En minimalform, f.eks. *grünes Licht*, som fungerer som metaleksikografisk identifikation af idiomet og har funktion af bookmark for hyperlinks.
- 2) De typiske syntaktiske strukturer, som på den ene side er vigtige med henblik

på identifikation af idiomer, især når formen eller dele af konstruktionen forekommer i andre kontekster (man kunne nok også sige: i ordret brug), og som på den anden side er uundværlige ved sprogproduktion:

- (a) *grünes Licht für jemanden/etwas*
 - (b) *jemand/etwas gibt (jemandem/etwas) grünes Licht (für/zu etwas)*
 - (c) *jemand/etwas erhält grünes Licht (für/zu etwas) (von jemandem) (mit etwas)*
 - (d) *jemand/etwas hat grünes Licht*
 - (e) *etwas bedeutet grünes Licht für jemanden/zu etwas*
- 3) Hyperlinks til analyserede korpuseksempler med en udførlig beskrivelse af grammatiske regler og restriktioner og eksempler på negation. Yderligere er der hyperlinks til andre idiomer med samme eller lignende form: *jemand rechnet mit grünem Licht*.
 - 4) Idiomets betydning(er) beskrives ud fra de analyserede korpuseksempler, og synonyme og antonyme idiomer angives. Desuden forklares idiomets etymologi og pragmatiske restriktioner.

De ensprogede informationer suppleres med tosprogede informationer som fungerer som hjælpemiddel fortrinsvis ved sprogreception og oversættelse. Disse er:

- 1) ordret oversættelse til islandsk/dansk,
- 2) betydningsangivelse på islandsk/dansk,
- 3) angivelse af ækvivalente islandske/danske idiomer og
- 4) en forklarende kommentar til ækvivalensrelationen.

Metaleksikografiske informationer i databasen bruges på den ene side ved struktureringen af de forskellige tilgangs- og søgemuligheder (alfabetisk liste med de enkelte komponenter, søgning efter de enkelte ord i idiomer, de enkelte komponenters semantiske felter, angivelse af koncepter til konstruktion af onomasiologisk tilgang til idiomernes betydning), på den anden side giver de mulighed for videregående forskning på grundlag af projektets resultater. Som eksempler på den slags metaleksikografiske informationer nævnes angivelse af frekvens i det tyske korpus *Deutscher Wortschatz* (www.wortschatz.uni-leipzig.de) og bekendthedsgrad hos tyske sprogbrugere (sml. Hallsteinsdóttir/Šajánková/Quasthoff, under forb.).

2.3.3. Ækvivalensproblematikken

Selv om der findes rigelig teoretisk litteratur om idiomer og ækvivalens, er en adækvat gengivelse af ækvivalens i en multilingval idiomatisk database stadigvæk en kompliceret problematik at redegøre for (sml. en udførlig teoretisk diskussion og empirisk undersøgelse af idiomer og ækvivalens i Farø under forb.). Ækvivalens beskrives i dette projekt altid kun mellem to sprog ad gangen. Der er ingen

begrundelse for bestemmelse af multilingval ækvivalens idet der ikke findes nogen potentiel brugssituation til denne.

Der er overordnet tale om to ækvivalenstyper:

- 1) *Idiom* ↔ *idiom(er)*. Denne type kræver beskrivelse af semantiske, pragmatisk og strukturelle forskelle mellem funktionelt ækvivalente idiommer. Hovedvægten i dette projekt ligger følgelig på situativ/funktionel ækvivalens (pragmatik og semantik, semantiske relationer). Idiomformen er kun relevant ved "tematiseret" brug (Farø under forb.). Sproglige (semantiske, strukturelle, etc.) relationer og/eller metaforik inddrages i begrænset omfang, først og fremmest med henblik på "falske venner".
- 2) *Idiom* ↔ *etordsleksem(er)*. Denne type betragtes i traditionel fraseologiforskning som nul-ækvivalens. Nul-ækvivalens er dog udelukkende en metodisk konsekvens af en teoretisk afgrænsning mellem idiommer og ikke-idiommer. Ægte funktionel nul-ækvivalens er med stor sandsynlighed meget sjælden. I første omgang er lemmataene i databasen dog af metodiske årsager begrænset til idiommer (med mulighed for udvidelse med andre sproglige enheder senere).

Idiommer indgår desuden i relationer til andre sproglige enheder (f.eks. synonymi, antonymi, orddannelse) som burde inddrages i den leksikografiske beskrivelse. Denne form for beskrivelse kræver dog at sproget betragtes i sin helhed. En konsekvent fortsættelse heraf ville være at sætte spørgsmålstegn ved "rene" idiomordbøger, ikke for at gøre mit eget arbejde overflødigt, men for at foreslå integration af idiommer i omfattende leksikografiske opslagsværker (leksika, tesaurusser). Integration af idiommer (og andre flerordsforbindelser) i leksikografiske opslagsværker og beskrivelse af intrasproglige relationer stiller den ensprogede leksikografi over for store udfordringer med hensyn til formulering af betydnings-, synonym- og antonymangivelser. I multilingval leksikografi med beskrivelse af intrasproglige og intersproglige relationer er udfordringerne endnu større idet de intrasproglige relationer ikke nødvendigvis er symmetrisk *intersprogligt* set, dvs. ens i to sprog.

3. Sammenfatning

Beskrivelse af fraseologismer er et højaktuelt forskningstema. En del af resultaterne i dette projekt forventes at blive en konkret beskrivelse af dansk, tysk og islandsk fraseologi og kontrastive analyser i en eksemplarisk database der f.eks. kan bruges som basis for lærebøger eller i andre undervisningsmaterialer. Databasen skal desuden, med forudgående kontrastive analyser af de pågældende sprog, kunne udvides med andre sprog. En brugbar teoretisk konception for multilingvale ordbøger danner forudsætningen for det leksikografiske arbejde. Denne konception åbner muligheder for fraseografiske arbejder i andre sprog.

Det forventes ligeledes at projektets resultater vil være et vigtigt bidrag til teoretisk viden om fraseologi som sprogligt fænomen. Endvidere forventes det at projektets resultater bliver til et grundlag for (og opmuntring til) yderligere fraseologiforskning, især i kontrastiv og leksikografisk sammenhæng.

Litteratur

Ordbøger:

- Friðjónsson, Jón G. 1993: *Mergur Málsins. Íslensk orðatiltæki. Uppruni, saga og notkun*. Reykjavík: Örn og Örlygur.
 Halldórsson, Halldór 1991: *Íslenskt Orðtakasafn*. Reykjavík: Almenna bókafélagið.
 Jónsson, Jón H. 2002: *Orðaheimur. Íslensk hugtakaorðabók með orða- og orðasambandskrá*. Reykjavík: JPV Útgáfa.
 Hessky, Regina/Ettinger, Stefan 1997: *Deutsche Redewendungen. Ein Wörter- und Übungsbuch für Fortgeschrittene*. Tübingen: Gunter Narr Verlag.

Anden litteratur:

- Almind, Richard/Bergenholtz, Henning/Vrang, Vibeke under forb.: Theoretical and Computational Solutions for Phraseological Lexicography. I: *Linguistik Online*.
 K. Farø og E. Hallsteinsdóttir (udg.): *Neue theoretische und methodische Ansätze der Phraseologieforschung*. <http://linguistik-online.com>.
 Bergenholtz, Henning 1997: Polyfunktionale ordbøger. I: *LexicoNordica* 4, 15-28.
 Farø, Ken 2004a: Am Anfang. Zur Erforschung der dänisch-deutschen Idiomatik. I: J. Wirrer og Cs. Földes (udg.): *EUROPHRAS 2002. Phraseologismen als Gegenstand sprach- und kulturwissenschaftlicher Forschung*. Baltmannsweiler: Schneider Verlag Hohengehren, 105-117.
 Farø, Ken 2004b: Idiomor på nettet: Den danske idiomordbog og fraseografien (<http://www.idiomordbogen.dk>). I: *Hermes, Journal of Linguistics* no. 32, 201-235.
 Farø, Ken under udg.: Danish Phraseography. I: H. Burger, D. Dobrovolskij, P. Kühn og N.R. Norrick (udg.): *Phraseology. Phraseologie. Ein internationales Handbuch der zeitgenössischen Forschung* (HSK). Berlin/New York: Walter de Gruyter.
 Farø, Ken under forb.: *Idiomatizität – Ikonizität – Arbitrarität: Beitrag zu einer Theorie phraseologischer Äquivalenz*. Ph.d.-afhandling ved Københavns Universitet.
 Fellbaum, Christiane/Kramer, Undine/Neumann, Gerald under udg.: Korpusbasierte lexikografische Erfassung und linguistische Analyse deutscher Idiome. I: A.H. Buhofer (udg.): *EUROPHRAS 04*.
 Hallsteinsdóttir, Erla 2001: *Das Verstehen idiomatischer Phraseologismen in der Fremdsprache Deutsch*. Hamburg: Verlag Dr. Kovač. http://www.verlagdrkovac.de/0435_volltext.htm.
 Hallsteinsdóttir, Erla 2006: Phraseographie. I: *Hermes* 36.
 Hallsteinsdóttir, Erla under udg. a: A bilingual electronic dictionary of idioms. I: H. Gottlieb og J.E. Mogensen (udg.): *Dictionaries, Lexicographical Options and User Needs. Selected*

- Papers from the 12th International Symposium on Lexicography*. Amsterdam: John Benjamins.
- Hallsteinsdóttir, Erla under udg. b: Vom Wörterbuch zum Text zum Lexikon. I: U. Fix, G. Lerchner, M. Schröder og H. Wellmann (udg.): *Zwischen Lexikon und Text – lexikalische, stilistische und textlinguistische Aspekte*. Leipzig: Sächsische Akademie der Wissenschaften zu Leipzig.
- Hallsteinsdóttir, Erla under udg. c: Konzeption und Erstellung einer computergestützten zweisprachigen Phraseologiesammlung Isländisch – Deutsch. I: A.H. Buhofer (udg.): *EUROPHRAS 04*.
- Hallsteinsdóttir, Erla/Šajánková, Monika/Quasthoff, Uwe under forb.: Phraseologisches Optimum für Deutsch als Fremdsprache. Ein Vorschlag auf der Basis von Frequenz- und Geläufigkeitsuntersuchungen. I: *Linguistik Online*. K. Farø og E. Hallsteinsdóttir (udg.): *Neue theoretische und methodische Ansätze der Phraseologieforschung*. <http://linguistik-online.com>
- Jesenšek, Vida under udg.: Zum Problem der Äquivalenz in mehrsprachigen phraseologischen Wörterbüchern. I: A.H. Buhofer (udg.): *EUROPHRAS 04*.
- Koller, Werner 1987: Überlegungen zu einem Phraseologie-Wörterbuch für Fremdsprachenunterricht und Übersetzungspraxis. I: H. Burger og R. Zett (udg.): *Aktuelle Probleme der Phraseologie*. Frankfurt a. M.: Peter Lang, 109-120.
- Soehn, Jan-Philipp/Römer, Christine under udg.: Wann ist ein Idiom ein Idiom? Eine Analyse von Phraseologismen ohne freie Lesart. I: A.H. Buhofer (udg.): *EUROPHRAS 04*.
- Sverrisdóttir, Oddný G. 1987: *Land in Sicht. Die kontrastive Untersuchung deutscher und isländischer Redensarten aus der Seemannssprache*. Frankfurt a. M.: Peter Lang.
- Tarp, Sven 1998: Leksikografien på egne ben. Fordelingsstrukturer og byggede i et brugerorienteret perspektiv. I: *Hermes, Journal of Linguistics* 21, 121-137.

Internetadresser:

- <http://corpora.informatik.uni-leipzig.de/>
<http://kollokationen.bbaw.de/index.html>
<http://www.asksam.com>
<http://www.dendanskenetordbog.dk>
<http://www.ephras.org>
<http://www.idiomordbogen.dk>
<http://www.ordanet.is>
<http://www.ordnet.dk>
<http://www.personal.uni-jena.de/~xcr/website/forschung/IVP-Web/main.htm>
<http://www.wortschatz.uni-leipzig.de>