

Ordbøger, konkordanser og lemmatisering.

1. Indledning.

Ordbøger og de fleste konkordanser har mange træk til fælles: bortset fra det faktum, at de giver oplysninger om betydningen og (eller) brugen af ord, repræsenterer de hver resultatet af en operation, som vi kan kalde med en fælles betegnelse lemmatiseringen. Lemmatiseringen er en manipulation af teoretisk-praktisk karakter, som gør det muligt at forvandle en løbende tekst (parole) til en ordnet liste af kanoniske former (langue). Disse kanoniske former kalder man stikord, når vi har at gøre med ordbøger, mens vi bruger lemma (lemmaer, lemmata), når vi omtaler samme enheder i konkordanser (de såkaldte 'lemmatiserede' konkordanser). Ordbøger og konkordanser kan således betragtes som ordnede lister af kanoniske former, men det materiale til indsamling af disse kanoniske former er principielt meget forskelligt: lemmata stammer traditionelt fra de løbende ord i en begrænset tekstenhed (f.eks. et værk eller et forfatterskab), mens det materiale, der ligger til grund for udvælgelse af stikord, stammer fra større tekstenheder, eller mindre vel afgrænsede tekstenheder (f.eks. en større periode som nudansk eller oldfransk). Lemmatiseringen består i at klassificere og reducere visse ordformer til principielt "mindre komplicerede" ordformer. Denne operation forudsætter en konvention af morfologisk og semantisk art. Lemmatiseringen kan i datamatisk sprogbrug for så vidt også godt betragtes som en generel beskrivelse af de adresser, under hvilke man kan finde de enkelte ordformer.

Nærværende artikel har som formål at studere en række ordbøger og en række lemmatiserede konkordanser og udlede deraf de kriterier, som har været herskende for at lemmatisere eller udpege en for til stikord. Det materiale, som er brugt til dette studium, omfatter især franske værker, i særdeleshed ordbøger og konkordanser om oldfransk eller om oldfranske forfattere. Enkelte eksempler er dog taget fra danske ordbøger og fra nufransk materiale. Der bliver også sat spørgsmålstegn ved, om lemmatiseringen i det hele taget er en nødvendig operation. Artiklen forsøger sluttelig at opstille en liste over

de betingelser, der skal være til stede, for at lemmatiseringen (om nogen) kan være tilfredsstillende.

2. Kanoniske former og konvention.

De fleste grupperinger af ordformer under en fælles nævner støtter sig til en almindelig accepteret konvention: f.eks. serier af ordformer som book og books hører sammen med fællesnævneren BOOK, men også work, works og worked med fællesnævneren WORK. Dette kaldes af Ladislav Zgusta i Manual of Lexicography "word-and-paradigm" modellen, og denne antagelse af en "word-and-paradigm model" er grundlæggende for ordbogforfatteres virke og alt andet leksikografisk arbejde. "When using this model (and for the lexicographer, it is really the best one), the lexicographer assumes that the lexical meaning of every single series of forms such as those quoted above remains the same, while only the different grammatical categories are expressed by different forms...", og længere hen ang. de 1350 mulige former af verbet på sanskrit: "It is only natural that the lexicographer cannot indicate all these forms in his dictionary. For every paradigm, i.e. for every inflected word and for the totality of its forms, he indicates only one form that is considered basis: this is the so-called canonical form of a word. The canonical form represents the whole paradigm...". (Zgusta, op.cit., p. 119)

For at forene ligeledes chevaux og cheval under CHEVAL, meilleur, bonnes, bons, bon til BON, vais, irai, allant til ALLER må vi bruge samme antagelse. Man kan hermed konstatere, at traditionen uvægerligt rangerer flertal under ental, femininum under maskulinum, bøjede verbalformer under infinitte former (helst infinitiv), komparativ (og superlativ) under den ikke komparerede form osv. At konventionen afviger somme tider fra sprog til sprog behøver ikke at undre os, konventionen udspringer bare fra en tradition, og det er derfor ikke mærkeligt i sig selv at skulle søge en verbalform under infinitiv på fransk eller på dansk, men under første person ental indikativ præsens for latin eller græsk. "The guiding principle for the choice of the canonical form is that it should be as good a starting point for the construction of the other forms of the paradigm as possible. In the majority of languages, the canonical forms are by now fixed by tradition and the lexicographer will do well not to abandon the tradition unless the reasons for a change are really imperative...". (Zgusta, ibid. p. 120) Problemerne opstår virkeligt, når man ikke kan finde

nogen mulig "kandidat" i ental til en flertalsform for eksempel, eller en maskulinum form til en femininum form osv. Hvad skal man stille op med følgende ord, hvis ikke de selv skal udpeges til kanoniske former:

Fransk: vacances VACANCE? VACANCES?
 fastes (subst)
 fiançailles

Engelsk: trousers
 Dansk: bukser
 Fransk: enceinte (adj.)

Har man desuden lov til at sætte som kanonisk form en lidet brugt form som GÉSIR (for mere almindeligt forekommende former som gisant, gît, og ligeledes TISTRE for tissu? Og i fald man går med til at samle under ét en række forskellige rødder som vais, allai, irai under ALLER hvad er der som forhindrer én i at samle je me rappelle og je me suis souvenu under SE SOUVENIR? Zgusta forestiller sig i disse tilfælde, at man udnævner en form til at være en kanonisk ad hoc form ("a canonical ad hoc"). Disse betragtninger sammen med en række andre gør, at det at finde kanoniske former eller lemmata byder på ikke så få problemer, for ikke at sige faremomenter (især i skabelse af ens egne "ad hoc" former

3. Den vanskelige norm.

Hvis en konkordansforfatter strengt taget kan lade være med at lemmatisere sin konkordans, og dermed vælger at give sine brugere de rå resultater af konkordansen, kan man ikke gøre det samme, når det drejer sig om ordbøgerne. Selv om forfatterne til ordbøgerne er bundet til at finde kanoniske former, er de på en måde bedre stillet, eftersom de mange gange selv kan vælge deres egne kriterier for normen: For eksempel viser forfattere til den franske ordbog Le Petit Robert en del frihed, når de ca. 10 år efter udgivelsen af deres ordbog tillader sig at omfatte slang, tabuord m.m. i 2. udgave. Således siger Alain Rey i forordet af 1977-udgaven (side xviii) "Quant aux mots tabous, argotiq ou malsonnants, il nous a semblé que l'évolution des moeurs devait être reflétée, ne serait-ce que pour répondre aux besoins de compréhension les plus évidents".. og senere: "Le Petit Robert s'est donné comme objectif de combler les lacunes les plus criantes de l'éducation offi-

cielle". Her er det privilegium, som moderne forfattere har, når de behandler et sprogtrin, de selv kender, fordi de selv oplever sproget og dets mulige skiften men også tidens skiftende moralske begreber. Det forholder sig anderledes, når en ordbog eller en konkordans afspejler en tilstand, som hører til historien, og hvor man kan have svært ved at genetablere ord-paradigme-modellen, for eksempel fordi man mangler belæg på visse ordformer. Da lemmatisering eller kanonisering af visse former i forhold til andre altid i sidste analyse hviler på et skøn, kan man selvfølgelig vælge andre kriterier end de, som er tilgængelige fra teksterne selv. Man vil umiddelbart have en tendens til at bearbejde et materiale, som afspejler et ældre sprogtrin, med de kriterier, som gælder for det nyere sprogtrin (denne tendens er meget markant, hvis man oven i købet tilhører gruppen, som taler det sprog på et nyere trin end det, man studerer). En omvendt procedure, som man ofte iagttager hos franske forskere, består i at "gøre teksten ældre", fordi man f.eks. sammenligner med latin eller andre mulige etymologiske former.

For at etablere normen for en tekst af ældre sprogtrin er der dog følgende andre teoretiske muligheder:

- (1) Teksten indeholder teoretisk alle mulige former, det vil sige også de former, man vil udnævne til kanon. Teksten må derfor være af et vist omfang, men sprogforskeren vil stadig skulle sortere og etablere værdinormer for kanoniske former og lemmata.
- (2) Teksten er ikke selvtilstrækkelig, men der eksisterer gode beskrivelser af samme sprogtrin: pålidelige morfologiske og syntaktiske beskrivelser kan bruges til at udlede former, som ikke er fundet i teksten.
- (3) Teksten er ikke selvtilstrækkelig, og der findes ikke gode morfologiske eller syntaktiske beskrivelser af samme sprogtrin. Der findes til gengæld ordbøger, som afspejler samme sprogtrin. Her vil det være oplagt at følge disse ordbøger, vel at mærke hvis disse redskaber er pålidelige. Selv om de skulle være pålidelige, skal man altid gardere sig for mulige forvrængninger, der kan stamme fra forskellige opfattelser af den ydre virkelighed, en anden kultur, forskellige leksikografiske normer osv.

Der forbliver dog en række problemer i etaperne (2) og (3), fordi man principielt introducerer elementer, som ikke er til stede i den oprindelige tekst. Og hvad værre er, man skal vogte sig for,

at den udledte form ikke bliver en re-konstruktion, en tænkt form.

4. En komparativ undersøgelse af 5 konkordanser over oldfransk.

Materiale: 3 konkordanser fra det leksikologiske laboratorium ved Liège Universitet, som er etableret af C. Dubois, M. Dubois-Stasse, A. Fontaine-Leauve, G. Lavis og M. Graitson.

(a) Chretien de Troyes: Philomena (udg. 1969-70)

(b) Les Chansons de Blondel de Nesle (udg. 1971)

(c) Chretien de Troyes: Guillaume d'Angleterre I-II (udg. 1974) samt 2 andre konkordanser, den ene over oldfransk, den anden over en nyere tekst (15. årh.), som er etableret af et andet belgisk team, 'Service de linguistique française de l'université de Gand' med følgende medarbejdere: A. Boone, H. Cuyvers-Lippens, G. de Poerck, R. van Deyck-Bauwens, D. Willems, R. Zwaenepoel-Dhanis.

(d) Le Charroi de Nîmes (udg. 1970)

(e) François Villon (udg. 1974)

(Jeg henviser til bibliografien for de fuldstændige titler.)

Konkordanserne a, b og c fra Liège er blevet etableret med samme grundprincipper: ordformerne er blevet klassificeret alfabetisk under deres kanoniske former eller lemmata, og disse er også ordnet alfabetisk. Disse lemmata (herunder også grafierne af ord (et vigtigt element i ældre sprogtrin)) er hentet i Tobler-Lommatzschs Altfranzösisches Wörterbuch, en stor ordbog på 10 bind over oldfransk, men hvis udgivelse endnu i 1979 ikke er færdiggjort.

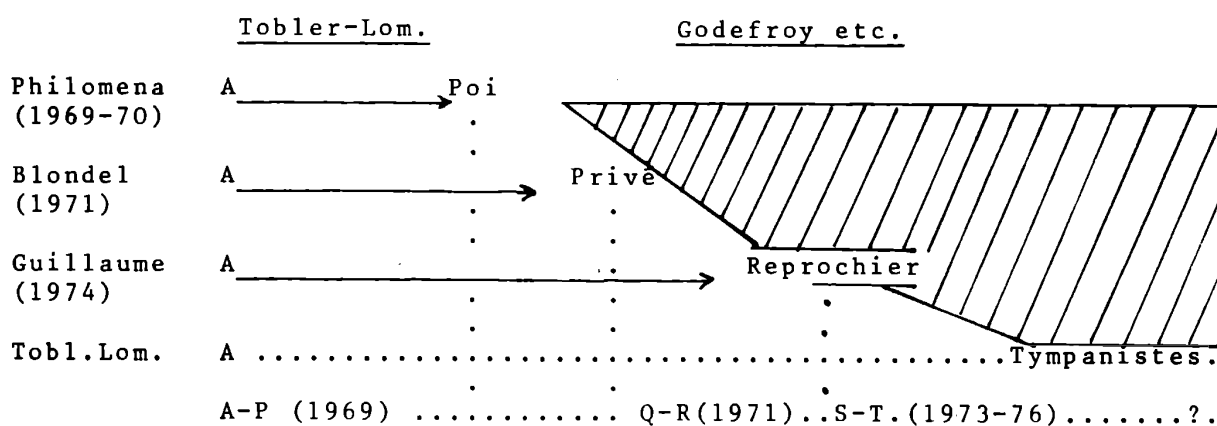


Fig. 1. Udviklingen af det lemmatiserede ordforråd iflg. Tobler-Lommatzsch i 3 konkordanser over oldfransk.

Fig. 1 viser fordelingen af lemmata ordnet efter Tobler-Lommatzsch for de 3 konkordansers vedkommende, efterhånden som tiden går. Forfatterne fra Liège har selvfølgelig måttet supplere deres kilde-materiale og har derfor måttet bruge andre forhåndenværende ordbøger, f.eks. Fr. Godefroys Dictionnaire de l'ancienne langue française et de tous ses dialectes du IX^e siècle au XV^e siècle (8 bind + 4 tillægsbind), Flûtres Table des noms propres avec toutes leurs variantes figurant dans les romans du moyen âge (CESCM, Poitiers, 1962), Walther von Wartburgs Französisches Etymologisches Wörterbuch, m.m.

En sammenligning af ordet QUE i de tre konkordanser viser, at man for Philomenas vedkommende har måttet bruge Godefroys ordbog, og får derefter 2 lemmata; for Blondel får man tre lemmata for QUE, stadig væk med Godefroy som kilde, mens man for Guillaume, som er lemmatiseret ifølge Tobler-Lommatzschs ordbog, får 5 forskellige lemmata.

Konkordanserne (d) og (e) er også blevet lemmatiseret på direkte eller indirekte vis med Tobler-Lommatzsch som forbillede. Konkordansen (e) om Villons forfatterskab introducerer en såkaldt "dobbel lemmatisering" ("entrée" eller "lemme" siden "forme de base" og endelig "occurrences"). Det mest bemærkelsesværdige er, at man for at lemmatisere en tekst fra det 15. århundrede bruger kriterier fra et ældre sprogtrin. I Villons tid er f.eks. kasus bortfaldet, mens det trin, Tobler-Lommatzsch beskriver, er det såkaldte klassisk oldfransk, som bl.a. endnu byder på et to-kasus-system.

5. En undersøgelse af 4 oldfranske ordbøger og deres valg af lemmata eller stikord.

Bortset fra Tobler-Lommatzsch og Godefroy vil jeg også omtale Greimas' Dictionnaire de l'ancien français jusqu'au milieu du XIV siècle samt Grandsaignes d'Hauterives Dictionnaire d'ancien français. Sidstnævnte ordbøger er snarere "skole"-udgaver end reelt videnskabelige ordbøger.

De ord, der er afprøvet, er følgende:

(1) verbet finirs ækvivalenter i oldfransk. Det drejer sig om at vide, om man vælger fenir, finir eller finer, som er 3 parallelle former i oldfransk.

(2) om den etymologiske dublet faillir/falloir svarer til 2 eller kun til et verbum i oldfransk.

(3) hvilken form er valgt som kanonisk form for craindres repræsentanter i oldfransk (6 eller flere forskellige varianter ifølge Kr. Nyrops Grammaire historique).

	<u>fenir</u>		<u>finir</u>		<u>finer</u>
Tobl. Lom.	FENIR	←	(FINIR) (cf. FENIR)		FINER
Godefroy	(FENIR) (cf. FINIR)	→	FINIR fenir		FINER finyer
Greimas	(FENIR) (cf. FINIR)	→	FINIR, FENIR		FINER
Grandsaignes	(fenir)		0		FINER

Fig. 2. Fordelingen af de former, der betyder finir i 4 ordbøger over oldfransk.

Fig. 2 viser fordelingen af former for finir.

	<u>falir, fallir, faillir</u>		<u>faloir, falloir</u>
Tobl. Lomm.	FALIR		0
Godefroy	FAILLIR fallir, falir, falyr		0
Godefr. Compl.			FALLOIR
Greimas	FAILLIR		FALOIR
Grandsaignes	0		0

Fig. 3. Fordelingen af formerne af dubletten: faillir/falloir i 4 ordbøger over oldfransk.

Fig. 3 viser, hvilke valg af kanoniske former, der er foretaget for faillir/falloir.

	<u>criembre</u>	<u>criendre</u>	<u>craindre</u>	<u>cremer</u>	<u>cremir</u>	<u>cremoir</u>
Tobl. Lom.	CRIEMBRE	0	0	cremer v. Godefroy	v.criembre	v.criembre
Godefroy	0	0	0	CREMER cromer	CREMIR	CREMOIR cremmoir
Godefr. Compl.	CRIEMBRE	0	CREINDRE	0	0	0
Greimas	CRIEMBRE, CREIMBRE V. CREMER	0	0	CREMER, IR, OIR CRIEMBE (?)		
Grandsaignes	CRIEMBRE	0	0	CREMER	CREMIR, OIR	

Fig. 4. Fordelingen af mulige infinitivformer af "craindre" i 4 ordbøger over oldfransk.

Fig. 4 viser den interessante fordeling af de mange former afledt af latin trēmere. Iøjnefaldende ændringer som det at skifte meningen mellem selve ordbogen og dens supplement (Godefroy), "Parti pris" mod det moderne hos Tobler-Lommatzsch og samtlige andre, bortset fra Godefroy Compl. Det allerinteressanteste består i valget af "primus inter pares"-formen, når der findes mange sideformer. Tobler-Lommatzsch er den eneste, som vælger CRIEMBRE fremfor de andre former. En større undersøgelse har siden vist mig, at denne form ikke har noget belæg hos Tobler-Lommatzsch selv, men heller ikke i hans kilder. (Jeg henviser til en anden artikel, jeg har skrevet om emnet i en lidt modificeret form, i "Proceedings from the Conference on General and Jewish Lexicography, Delaware, 1979)(i trykken).

Man kan derfor sætte et spørgsmålstegn ved denne form for kilder, når samtlige under punkt 4 omtalte konkordanser har gjort Tobler-Lommatzsch's principper til deres, hvad angår valg af lemmata.

6. Undersøgelse af typisk "kvindelige" begreber i nogle moderne ordbøger.

(a) ordet enceinte (gravid).

Dette ord er brugt som kanonisk ad hoc form i både 1968- og 1977-udgaver af Petit Robert. Desuden også i TLF's frekvensordbog. I dette tilfælde kan man sige, at man ved at udvælge formen enceinte til kanonisk ad hoc følger den metalingvistiske virkelighed.

(b) ordet infirmière i en række franske ordbøger. Gougenheims Dictionnaire fondamental giver som stikord infirmier, og derunder skal man finde infirmière (jvfr. épiciers-épicière, employeur-employeuse osv.). Dette faktum er interessant i en ordbog, som skulle afspejle de 3500 hyppigste ord! Petit Robert udgaver fra 1968 og 1977 giver samme præsentation af ordet infirmière. Man skal søge det under infirmier. Man skal dog bemærke, at samtlige eksempler henviser til en kvindelig sygeplejerske, mens 2 modeksempler er af hankøn.

(c) ordet sage-femme (jordmor). Dette ord forekommer som stikord i le Petit Robert. En interessant henvisning leder læseren til accoucheuse, som til gengæld ikke findes som stikord. Ordet findes, men under stikordet accoucheur.

(d) ordet sygeplejerske og dets behandling i Nudansk ordbog. Dette ord er blevet studeret udelukkende for at se, om der var forskel i opfattelse mellem 2 kulturer. Der er desuden det spændende, at dette ord på dansk kan siges at tilhøre et af de få paradigmer, som udtrykker hunkøn på dansk, jvfr.:

lærer-lærerinde/arbejder-arbejderske.

Man kan på dansk også udtrykke køn med følgende midler: dvs. ved at danne komposita: han-hun vs hun-hund, mandlig kollega vs kvindelig kollega.

Ang. sygeplejerske er det af interesse at se, under hvilket stikord formen bliver givet.

Nudansk ordbog, som er en tæt pakket ordbog på 2 bind, har åbenbart pladsmangel, og dens præsentation ser ud til at være præget deraf. Stikordet for sygeplejerske er sygemelde, som også fungerer som stikord for sygemelding, sygepasser, sygepleje, sygeplejer, -plejerske, -seng. Man skal også her bemærke, at et vigtigt begreb bliver givet under dets afledning og ikke som første valg.

Valget af eksempler i denne artikel, men også andre steder er meget lærerig: "forstanderinde for hospitalets sygepleje.." Sygeplejerske indgår som definerende element i definitionen af ordet sygehjælper, et ord som er medtaget noget før under stikordet syge er endnu mere afslørende: sygehjælper, "om person, der har mindre uddannelse end en sygeplejerske, og som yder lettere sygepleje". Alt dette viser, at den semantiske sfære åbenlyst er knyttet til det kvindelige, mens det rent lingvistiske valg opereres på gængse mandlige kriterier.

7. Konklusion.

En række iagttagelser om henholdsvis konkordanser, ordbøger over ældre sprogtrin, og ordbøger over nyere sprogtrin viser, at den leksikografiske tradition byder på en række dårligt definerede områder, hvad angår valget og kriterierne for valget af den såkaldte makrostruktur. Man skal derfor her som konklusion forsøgsvis opstille en række betingelser, som burde være til stede for at sikre, at en eventuel lemmatisering er forsvarlig:

- (1.) Den kanoniske form X repræsenterer ordformerne y, z, osv., hvis og kun hvis x, y, x er en del af samme paradigme.
- (2.) For at være i samme paradigme må x, y og z nødvendigvis være samtidige i tid (samme periode).
- (3.) Og samtidige i rummet (én og samme tekst, ét og samme sprog-niveau, samme stil, samme dialekt osv.).
- (4.) X må være mere repræsentativt end y, x osv. Det vil sige, at y, z, osv. må være lette at udtrække fra X.
- (5.) X's frekvens må være mindst lige så høj som y's, z's (X må altså ikke være en teoretisk, rekonstrueret, ikke belyst eller mindre brugt form af y, z osv.).
- (6.) Den lingvistiske enhed X må samtidig være relateret til den ekstralingvistiske virkelighed, og den må repræsentere denne virkelighed bedst muligt.

BIBLIOGRAPHI

1. Leksikografiske problemer.

- Baldinger, Kurt: Introduction aux dictionnaires les plus importants pour l'histoire du français, Klincksieck, Paris, 1974.
- Brackenier, R.: "Index et Concordances d'auteurs français modernes. Etudes critiques" Travaux de linguistique 3 (1972) 1-43, et 4 (1975) 1-61.
- Dubois, Jean et Claude: Introduction à la lexicographie: le dictionnaire, Larousse, Paris, 1971.
- Hanon, S.: Compte rendu de J.J. Duggan: A Concordance of the Chanson de Roland et G. de Poerck et al. Le Charroi de Nîmes I-II, Revue Romane, VIII, (1973) 421-423.
- Hanon, S.: "Chronique. Colloque sur l'analyse des corpus linguistiques. Problèmes et méthodes de l'indexation maximale" Cahiers de Lexicologie, 23, (1973) 117-124.
- Hanon, S.: "Mots dans le texte, mots hors du texte: réflexions méthodologiques sur quelques index et concordances appliquées à des oeuvres françaises, italiennes ou espagnoles." Revue Romane, XII, (1977), 272-296.
- Hanon, S.: Compte rendu de Rika van Deyck et Romana Zwaenepoel: François Villon Oeuvres d'après le manuscrit Coislin, textes, variantes et concordances. Revue Romane, XIII, (1978), 340-344.
- Hausmann, Franz Josef: Einführung in die Benutzung der neufranzösischen Wörterbücher, Niemeyer, Tübingen, 1977.
- Matoré, Georges: Histoire des dictionnaires français, Larousse, Paris, 1968.
- Nyrop, Kr.: Grammaire historique de la langue française, tome II, Gyldendal, Copenhague, 1960.
- Quemada, Bernard: Les dictionnaires du français moderne 1539-1863, Didier, Paris, 1968.
- Rey, Alain: "Les Dictionnaires: forme et contenu". Cahiers de Lexicologie, VII, (1965), 65-102.
- Rey, Alain: Le lexique: images et modèles du dictionnaire à la lexicologie. Colin, Paris, 1977.
- Rey-Debove, Josette: Etude linguistique et sémiotique des dictionnaires français contemporains, Mouton, The Hague, Paris, 1971.
- Wagner, R.L.: Les vocabulaires français I-II, Didier, Paris, 1967-1970.
- Zgusta, Ladislav: Manual of Lexicography, Mouton, Academia, The Hague, Paris, Prague, 1971.

2. Konkordanser.

Blondel de Nesle: Les Chansons de Blondel de Nesle. Concordances et index établis d'après l'édition. L. Wiese par G. Lavis. Traitement automatique: C. Dubois. Publications de l'Institut de Lexicologie française de l'Université de Liège. Faculté de Philosophie et Lettres de l'Université de Liège, (1971).

Chanson de Roland: A Concordance of the Chanson de Roland, compiled by Joseph L. Duggan. Ohio State University Press, 1969.

Charroi de Nîmes: Le Charroi de Nîmes, chanson de geste, par G. de Poerck, concordances, R. van Deyck, texte et variantes, R. Zwaenepoel, traitement automatique. Tome I-II. Textes et traitement automatique. Librairie- Editions Mallier, Saint-Aquilin-de-Pacy (Eure) 1970.

Chrétien de Troyes: Philomena. Concordances et index établis d'après l'édition C. de Boer, par C. Dubois, M. Dubois-Stasse et G. Lavis. Faculté de Philosophie et Lettres de l'Université de Liège. Publications de l'Institut de Lexicologie française de l'Université de Liège, (1970).

Chrétien de Troyes: Guillaume d'Angleterre. Concordances et index établis d'après l'édition M. Wilmotte par M. Dubois-Stasse, A. Fontaine-Lauve. Traitement automatique: C. Dubois, M. Graitson, I-II. Faculté de Philosophie et Lettres de l'Université de Liège. Publications de l'Institut de Lexicologie française de l'Université de Liège. (1970).

Joachim du Bellay: Joachim du Bellay, La Deffence et Illustration de la Langue Francoyse. Concordance établie par Suzanne Hanon. Traitement automatique: Poul Bonne Jørgensen et Ulf Hagen Kølgaard. Etudes romanes de l'Université d'Odense, vol. 6, Odense University Press, 1974.

Villon: François Villon: Oeuvres d'après le manuscrit Coislin. Rika van Deyck: Texte, variantes et concordances. Romana Zwaenepoel: Traitement automatique. I-II, Mallier, Saint-Aquilin-de-Pacy 1974.

3. Ordbøger.

de Gorog, Ralph: Lexique Français moderne - Ancien français. University of Georgia Press, Athens, 1973.

Dictionnaire alphabétique et analogique de la Langue française par Paul Robert, Société du Nouveau Littré, Paris, 1968.

idem, 2^e ed. Paris, 1977.

Dictionnaire des Fréquences. Vocabulaire littéraire des XIX^e et XX^e siècles. Centre de Recherche pour un Trésor de la Langue française, I-IV. C.N.R.S.- T.L.F. Didier, Nancy, 1971.

- Dictionnaire fondamental de la Langue française par Georges Gougenheim, Didier, Paris, 1958, nelle édition revue et augmentée.
- Diez, Friedrich: Etymologisches Wörterbuch der Romanischen Sprachen Bonn, Heilbronn, Leipzig, 1887-1889 et Georg Olms Verlag, Hildesheim, New York, 1969.
- Gamillscheg, Ernst: Etymologisches Wörterbuch der Französischen Sprache, Heidelberg, Carl Winther, 1969.
- Godefroy, Frédéric: Dictionnaire de l'ancienne langue française et de tous ses dialectes du IX^e au XV^e siècle (I-VIII + Compl. I-IV), Vieweg, Paris, 1881-1902.
- Grandsaignes d'Hauterive, R.: Dictionnaire d'ancien français. Moyen âge et Renaissance, Larousse, Paris, 1947.
- Greimas, A.J.: Dictionnaire de l'ancien français jusqu'au milieu du XIV^e siècle, Larousse, Paris, 1969.
- Mayer-Lübke, W.: Romanisches Etymologisches Wörterbuch, (REW) 4. auflage, Heidelberg, C. Winther, 1968.
- Nudansk ordbog udgivet med støtte af Undervisningsministeriet I-II. Politíkens Forlag, København, 1964 red. Lis Jakobsen et. al.
- Tobler-Lommatzsch: Altfranzösisches Wörterbuch (I-X), Franz Steiner Verlag, Wiesbaden, 1915-1976.
- Wartburg, Walther v.: Französisches Etymologisches Wörterbuch, (FEW), Eine Darstellung des galloromanischen Sprachschatzes, Band XIII, 2. Teil, Basel, Zbinden Druch und Verlag AG, 1965.