

Résumés

Gregor Büchel, Bernhard Schröder

Le choix d'un système de codage pour les structures complexes de texte dépend de nombreux critères. Dans la plupart des cas le choix d'un système de codage basé sur SGML or XML répond à ces critères. Les données peuvent être contrôlées avec l'aide des systèmes de gestion de base de données. A cette fin les objets de texte et leurs relations doivent être représentés par les moyens expressifs du type de base de données choisi. Il dépend de la manière dans laquelle les documents sont structurés et qu'on utilisera les données quel modèle de base de données est le plus adéquat pour la recherche et la gestion.

Ingrid Schmidt, Carolin Müller

Les dictionnaires électroniques actuellement disponibles sur le marché n'utilisent qu'une partie très limitée de leur potentiel. Il est notamment frappant de constater que les versions électroniques de bons dictionnaires imprimés laissent souvent à désirer sur le plan de la qualité. Si l'on observe les exigences sans cesse changeantes du paysage médiatique, on remarque la nécessité de nouvelles bases dans les rapports avec les contenus lexicographiques. C'est avec cet objectif en tête que nous nous penchons d'abord sur certains aspects des thèmes suivants: SGML, Multiple-Media-Publishing, TEI, métalexicographie. Nous débattons ensuite sur cette base de l'élaboration d'un modèle lexicographique opposant la longévité à la rapidité des médias et avec lequel il doit être justement possible de réagir avec flexibilité aux exigences du marché.

Angelika Storrer

L'article rend compte des perspectives que la technologie hypertexte ouvre à l'élaboration de dictionnaires électroniques innovateurs.

Il explique d'abord les caractéristiques essentielles de l'idée de hypertexte: la liaison des unités d'information par des liens hypertextes, l'intégration de texte, images, son et vidéo dans le même logiciel, l'accès flexible aux données lexicaux. Il est démontré que la plupart des dictionnaires électroniques en vente aujourd'hui n'utilisent pas encore le véritable potentiel de la technologie hypertexte.

Pour cette raison, la partie principale de l'article présente sept thèses et discute les voies d'une meilleure utilisation de la plus-value de la technologie hypertexte dans les dictionnaires électroniques à l'avenir.

Ingrid Lemberg

Internet nous offre un moyen multifonctionnel de publication, de production et de communication. Les dictionnaires sur Internet possèdent les avantages suivants: liberté d'utilisation à partir de n'importe quel lieu et pendant une durée illimitée (chap. 4.1), possibilité de

correction et de mise à jour continues d'éléments du lexique déjà publiés (chap. 2.2), hypertextualisation et utilisation multimedia. Il est possible, grâce à ce nouveau moyen, de présenter les dictionnaires sous d'autres formes et de les élargir en créant par exemple des liens entre le dictionnaire et les textes de référence ou d'autres systèmes d'information (chap. 2.1.1). L'utilisation multimedia est considérée comme particulièrement appropriée pour présenter des connaissances lexicales ou lexicographiques avec divers moyens d'expression et contribuer par conséquent à une communication efficace du savoir (chap. 2.1.2). Internet fournit également à la lexicographie une toute nouvelle technique de production: l'accès à des banques de données lexicographiques sur Internet permet la création de dictionnaires en réseau, complets et homogènes, à partir de sites répartis dans le monde entier (chap. 3). Internet est enfin un moyen de communication qui permet le contact entre l'utilisateur et le lexicographe : les utilisateurs d'un dictionnaire sur Internet peuvent demander par E-Mail des compléments d'explication tout comme ils peuvent eux-mêmes apporter des renseignements supplémentaires et précieux au dictionnaire (chap. 4.2).

Annette Klosa

L'auteur de l'article entend d'abord donner une définition fondamentale de ce qu'est „la qualité“ d'un dictionnaire sur CD-ROM pour discuter ensuite les critères qualitatifs énoncés par les critiques de dictionnaires électroniques. Cette comparaison permet ensuite à l'auteur de développer ses propres critères de qualité pour la production d'un dictionnaire électronique avant de les confronter à la réalité des conditions de production éditoriale.

Ulrike Haß-Zumkehr

L'article contient des réflexions sur l'esquisse d'un programme concernant les micro-structures d'un système hypertextuel d'informations concernant le vocabulaire allemand contemporain. Ces réflexions sont ancrées sur un projet concret qui est en train d'être réalisé à L'Institut für Deutsche Sprache à Mannheim. Ce nouvel outil de travail met les lexicographes dans une situation et devant des conséquences nouvelles, comme par exemple devant la nécessité de séparer structuration et présentation des données, de créer plusieurs accès à l'article de dictionnaire à partir d'informations différentes mais de valeur égale (concernant la grammaire, la sémantique etc.), de formuler les informations hors contexte, de changer la structure souvent trop dense des textes explicatifs, de nommer explicitement les sortes d'informations afin de faciliter leur accès pour l'utilisateur. Ce nouvel outil nécessite également l'élaboration d'une typologie des renvois et des chaînes d'information ainsi qu'un choix entre une structure d'article basée sur la forme du mot ou basée sur ses emplois.

Thomas Gloning, Rüdiger Welter

La documentation d'un lexique sous forme de base de données offre plusieurs avantages par rapport à l'uni-dimensionalité des dictionnaires structurés de manière alphabétique: elle permet de saisir le vocabulaire de divers points de vue structurels et analytiques (p. ex.

réseaux onomasiologiques, différentes matières, types de textes, diachronie et évolution biographique, rendement fonctionnel, histoire des idées, etc.) tout en ouvrant à l'utilisateur des possibilités subtiles d'accès et de consultation suivant des critères combinés. En outre, les données électroniques peuvent, à tout moment, être corrigées et mises à jour.

Notre contribution explique d'abord les principes de l'architecture complexe et certains aspects fondamentaux de la structure lexicale (§§ 1 et 2) avant de fournir des précisions sur la structure interne et les particularités du vocabulaire de Goethe (§ 3). La partie principale (§ 4) traite des possibilités d'accès étendues qu'offre le dictionnaire de Goethe en tant que base de données SGML traitée à l'aide de marquages lexicologiques, p. ex.: usages variés des données de base, multiples possibilités d'analyse et de recherche en faveur des utilisateurs, mise en réseau avec d'autres documents. Importante du point de vue de la politique culturelle sera la publication gratuite des données sur la toile mondiale.

Thomas Burch, Johannes Fournier

Les dictionnaires du moyen haut allemand les plus significatifs sont actuellement le «Mittelhochdeutsches Wörterbuch» de Benecke/Müller/Zarncke, le «Mittelhochdeutsches Handwörterbuch» de Lexer, ses «Nachträge» et enfin le «Findebuch zum Mittelhochdeutschen Wortschatz». Ces dictionnaires sont étroitement reliés entre eux et doivent pour cela être utilisés simultanément. C'est pourquoi ils sont des candidats idéaux pour la composition d'un dictionnaire relié électroniquement. De plus, les informations lexicographiques contenues dans l'ensemble des quatre ouvrages de référence seront accessibles dans une banque de données pourvue d'options de recherche complexes. L'application du SGML, défini comme un standard ISO, fournit la base pour un encodage qui ne dépend pas d'un logiciel ou d'un matériel spécifique, que ce soit pour un CD-ROM ou pour Internet. En outre, cette codification garantira la longévité des données concernées. L'application des lignes directrices TEI qui ont déjà été utilisées avec succès dans divers projets s'est avérée essentielle pour la progression constante de la numérisation rétrospective.

La dissertation traite des problèmes qui surviennent lors de l'application des lignes directrices TEI des dictionnaires électroniques. Toutefois, il est évident que les problèmes ne proviennent pas de l'application des lignes directrices TEI en tant que telle, ils résultent plutôt principalement du fait que les articles du BMZ et du dictionnaire de Lexer ne suivent pas toujours une structure générale bien définie, ce qui fait de l'encodage automatique du SGML une tâche souvent difficile. Dans des nombreux cas, seulement l'encodage manuel permet d'obtenir des documents TEI maniables. Néanmoins, les résultats obtenus jusqu'à présent justifient pleinement la décision en faveur des lignes directrices TEI.

Ralf Plate, Ute Recker

L'article discute les chances et les perspectives de la rédaction d'un dictionnaire de référence (Belegwörterbuch) historique soutenue par l'ordinateur ainsi qu'elles se représentent pour le nouveau Dictionnaire du Moyen Haut Allemand au bout de cinq ans de travail préparatoire. L'élaboration de ce dictionnaire dont on a prévu une dimension de quatre volumes (chacun à 1000 pages environ) commencera en l'année 2001 et devrait être terminée dans l'espace de vingt années. Le nouveau dictionnaire remplacera ses prédécesseurs du XIX^e siècle qui se réfèrent aux sources de la période entre 1050 et 1350. Il sera basé dès

le début sur le traitement des données électroniques ce qui lui procure un caractère innovatif parmi les grands projets de la lexicographie historique en Allemagne. Le recueil des matériaux se fonde d'un part sur des archives électroniques qui contiennent les textes de toutes les sources du corpus de base et d'autre part d'un index d'environ 80.000 mots compilés à part des enregistrements dans les dictionnaires précédents. Sont produits ensuite des archives de référence lemmatisés à partir de deux composantes: d'un côté des textes préparés d'abord à l'aide d'un système informatique pour la lemmatisation semi-automatique, d'autre côté d'une liste qui comprend toutes les occurrences des textes déjà enregistrés. Le dictionnaire sera donc élaboré sur la base de ces deux archives et à l'aide d'autres systèmes informatiques qui soutiennent le travail à partir du premier arrangement de la référence jusqu'à la composition de l'article. Le dictionnaire ainsi achevé servira à une double fonction: À cause de son construction aux moyens électroniques, il sera utilisable non seulement dans sa forme traditionnelle comme livre imprimé mais encore comme instrument électronique qui offre des possibilités de recherche plus étendues.

Les chances de la lexicographie soutenue par l'ordinateur consistent en la possibilité – ici démontrée à l'exemple du nouveau Dictionnaire du Moyen Haut Allemand – de réorganiser et de rationaliser avec l'assistance des moyens électroniques les opérations lexicographiques traditionnellement utilisées. Il est ainsi possible d'économiser une grande partie du travail jadis nécessaire pour copier, corriger et lemmatiser chaque référence particulière. Une grande avantage est représentée par le fait que la choix des références anciennement entreprise au cours de l'excerptation rentre maintenant dans le domaine de l'auteur d'un article qui est instruit en lexicographie. Il peut se pencher sur l'interprétation du matériau tandis que le système informatique achève les principaux travaux nécessaires pour ranger, écrire et contrôler les articles. En plus, on trouve des informations sur la structure du dictionnaire directement dans l'arrangement du fichier de chaque article ce qui rend possible des recherches exactes dans la version électronique du dictionnaire. Au-delà de la préparation du dictionnaire même, il y existe une autre chance présentée par les résultats faits quasiment en passant qui pourront avoir du valeur pour d'autres projets linguistiques ou philologiques. Dans le cas du dictionnaire décrit ici, ce sont p.ex. les archives électroniques contenant les textes et les références ou le système électronique pour l'établissement des index et des concordances qui peut être utilisé pour n'importe quel texte en moyen haut allemand.

Les perspectives discutées dans l'article se réfèrent surtout aux avantages de la publication électronique du dictionnaire et aux questions de la conception d'un tel dictionnaire. D'un part, ces questions concernent le rapport entre le dictionnaire et les matériaux de base au moment où ils sont publiés ensemble ce qui offre la possibilité d'un accès direct de l'article aux archives. D'autre part, la question se pose dans quelle mesure le dictionnaire devrait, apart de sa fonction usuelle, être accessible pour des prétentions ultérieures comme p.ex. la production des registres.

Gerd Richter¹

Les formes électroniques de présentation, comme celle du hypertexte, offrent de nouvelles possibilités dans la présentation aussi bien que l'utilisation des répertoires toponymiques.

¹ Für die Übersetzungen danke ich Frau Sabine Prechter. – G.R.

Basé sur la version traditionnelle – imprimée – du *Südhessisches Flurnamenbuch* (SHFLNB), cet article va présenter les avantages et les potentiels supplémentaires de la version hypertexte face aux restreintes de la version traditionnelle. Une analyse modèle du SHFLNB qui inclura aussi une description détaillée de sa structure médiale servira comme point de départ pour une présentation du potentiel hypertextuel du genre. Finalement, une réalisation exemplaire du SHFLNB en tant que hypertexte montrera les possibles nouvelles formes de présentation offertes aux auteurs de répertoires toponymiques, surtout en comparaison avec les formes traditionnelles en tant que livre imprimé.

Krzysztof Petelenz

L'importance pratique et commerciale de la technologie hypertextuelle dans le domaine de la lexicographie est une conséquence de la vulgarisation d'ouvrages de référence sur CD-ROM. Le développement du World Wide Web, quant à lui, a contribué à améliorer la qualité des informations reliées par des liens ainsi que de promouvoir les systèmes ouverts, les contenus thématiques, une communication plus aisée entre les auteurs et leur public et, par conséquent, une meilleure prise en compte des besoins de l'utilisateur. Ainsi, il est utile de réfléchir sur les améliorations possibles dans les dictionnaires en ligne. Le présent article essaie de créer les bases pour la présentation d'un dictionnaire bilingue reposant sur un hypertexte de haute qualité. Les stratégies de la structuration et de l'encodage des données lexicographiques seront discutées en détail et illustrées à partir d'un échantillon d'un dictionnaire polonais-allemand. En préliminaire, nous situerons les dictionnaires électroniques bilingues dans le contexte des systèmes hypermédia et nous évoquerons les particularités de la lexicographie bilingue qui déterminent la conception de dictionnaires bilingues en ligne.

L'intégration d'images, auxquelles ce nouveau média permet enfin d'accorder l'espace qui leur est dû, ainsi que leur catégorisation feront l'objet d'une étude approfondie. Finalement, nous traiterons quelques caractéristiques du dictionnaire en ligne qui apportent une «survaleur», comme par exemple l'interconnectabilité avec d'autres projets Web et l'expansion suivie et de la macrostructure et de la microstructure.

Claudia Kunze, Andreas Wagner

Nous présentons dans cet article GermaNet, réseau lexico-sémantique et thésaurus «en ligne», couvrant le vocabulaire allemand de base. GermaNet intègre les relations sémantiques élémentaires comme la synonymie, l'hyponymie, l'antonymie, ainsi que les relations casuelles des éléments lexicaux traités. Ce réseau sémantique offre une ressource importante pour le traitement de la désambiguïsation des mots, pré-requis pour de nombreuses applications en traitement automatique des langues, comme la recherche d'information ou l'annotation sémantique de corpus.

Lothar Lemnitzer

Les connaissances des besoins de l'utilisateur sont vitales pour la réalisation de nouveaux dictionnaires et leurs mises à jour. Cependant, ces données sont très difficiles à acquérir.

Heureusement, le Oueb nous permet maintenant de mettre en place un nouveau partenariat avec l'utilisateur. Celui-ci peut utiliser des dictionnaires sans frais, la contre-partie étant un enregistrement et une analyse de sa recherche. Cet article présente une telle utilisation avec un ensemble de dictionnaires bilingues consultables sur Internet. Les résultats présentés illustrent les avantages d'une telle recherche.