

method based on mnemonic principles. The underlying idea of this method (Gilbert calls it the "Mebiwegal Method") is, that Chinese characters are easy to learn if you tag them with mnemonic labels and form verbal picture systems ("Bildsysteme") for their graphic components. Based on his method Gilbert published a Chinese Learners Dictionary which is in turn described and analysed. Finally the advantages and disadvantages of Gilberts method are discussed.

Der Computer als chinesisch-deutsches Lexikon

Anton Lachner, Joachim Heinzl

1. Einführung

Einen Großteil seiner Arbeit verbringt der Übersetzer (auch der Dolmetscher bei der Vorbereitung) mit dem Lexikon. Entweder schlägt er ein ihm nicht geläufiges zu übersetzendes Wort nach, oder er sucht nach einer kontextuell passenderen Übersetzung. Der gewissenhafte Übersetzer konsultiert oft auch ein Wörterbuch, das von der Zielsprache (einsprachig oder wiederum mit Übersetzung) ausgeht, um Konnotation und Idiomatik des Translats zu überprüfen. So erst ist gewährleistet, daß das im Quell-Zielsprachen-Wörterbuch gefundene Lemma tatsächlich ein gebräuchliches Äquivalent ist, und nicht nur eine mehr oder weniger umschreibende „Übersetzung“.

Auf Ökonomie bedachte Übersetzer haben den Computer längst als unentbehrlichen Begleiter ihrer Tätigkeit entdeckt und nutzen ihn zumindest als Schreibgerät. Seine Vorteile bei der Aktualisierung von Texten, durch den Zugriff auf häufig wiederkehrende Textbausteine und die Flexibilität beim Ausdruck – selbst mit „anspruchlosen“ Textprogrammen – brauchen nicht extra hervorgehoben zu werden. Ein elektronisches Wörterbuch, das jederzeit auf dem selben Computer, auf dem der übersetzte Text entsteht, verfügbar ist und nachgeschlagene Einheiten in der Zielsprache in den am Monitor entstehenden Text übernimmt, könnte die Arbeit des Übersetzers erheblich beschleunigen. Braucht er dann nämlich die zutreffenden Translate nicht mehr direkt einzutippen, sondern kann sie über eine Menüziffer an jede beliebige Stelle seines Textes transferieren.

Ein solches System soll hier vorgestellt werden. Es handelt sich um einen deutsch-chinesischen elektronischen Thesaurus auf der Basis des *Neuen Chinesisch-Deutschen Wörterbuchs*¹ (fortan *CDW*), der in ein chinesisches Texteingabesystem, das so genannte *Münchner System*² integriert ist. Das *Münchner System* und das elektronische Lexikon wurden am Lehrstuhl für Feingerätebau der Technischen Universität München entwickelt und sind unter MS-DOS auf XT und AT-kompatiblen Rechnern lauffähig, wobei nicht auf die Zeichengeneratoren der Betriebssysteme zugegriffen wird. Die chinesische Texteingabe erfolgt zweckmäßigerweise in *Pinyin* in Einheiten von (mehrsilbigen) Wörtern; andere standardisierte Eingabeverfahren wie GB-Code (bzw. *Qūwèi*-Ziffern)³ sind ebenfalls vorhanden.

Wir werden nicht nur die für den Übersetzer wichtigen Charakteristika des elektronischen Lexikons anführen, sondern auch solche, die für Lernende, Lehrende und

¹Das neue Chinesisch-Deutsche Wörterbuch 1985, ed. Běijīng wàiguó yǔ xuéyuàn Déyǔdì. [Nebent.:] 新汉语词典.

²Cf. Zhang Shuying, 1981, „Neues Konzept einer Schreibmaschine für chinesische Schrift“ (Diss. München) und Ding T., Heinzl J.L., Lachner A. 1988, „Pinyin-Worteingabe – ein effizientes Eingabeverfahren für die chinesische Schrift“, in *Chinesisch und Computer* 1:12-29 (Bremen). Der zweite Aufsatz ist die deutsche Übersetzung eines Vortrages auf der „1987 International Conference on Chinese Information Processing“ in Běijīng.

³Zhōnghuá Rénmín Gònghéguó guójiā biāozhǔn xìnxī jiāohuàn yòng hànzi biānmǎ zìfú jí. Jīběn jí. 中华人民共和国国家标准信息交换用汉字编码字符集. 基本集. Běijīng 1981. Dies ist der Standard-Code für chinesische Schriftzeichen, vergleichbar dem ASCII-Code (American Standard Code for Information Interchange) für lateinische Buchstaben.

- [mit *ling jiàn* 另见 („siehe auch“) eingeleitet: Querverweis nur in Schriftzeichen auf ein homographes aber heterophones Lemma;]
- [mit *jiàn* 见 („siehe“) eingeleitet: Querverweis nur in Schriftzeichen auf einen gleich- oder anderslautenden Eintrag; oft Schriftzeichenvarianten;]

Kapitelzeichen

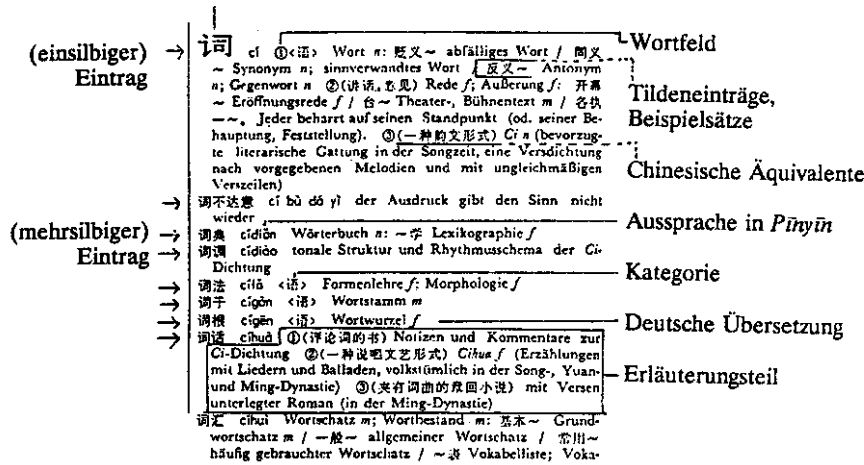


Abb. 3: Beispiele für Einträge des CDW mit Erklärungsteil

2.3. Kategorien

Die Angabe von Kategorien bei den chinesischen Stichwörtern umfaßt drei Bereiche, die rein äußerlich nicht unterschieden werden: (1) Wortarten, (2) Stilmittel, (3) Fachwortschätze. Auf den Seite 6-7 des CDW findet sich eine Liste der im Wörterbuch abgekürzten 75 Kategorien. Die nicht abgekürzten Kategorien sind dort nicht aufgenommen.⁴ Insgesamt werden 116 unterschiedliche Kategorien⁵ benutzt, deren Verteilung die Abbildungen 4a und 4b zeigen.

⁴Beim Abdruck dieser alphabetisch angeordneten Liste ist ein Fehler unterlaufen. Die auf zwei Seiten mit je zwei Spalten abgedruckte Liste hat auf der ersten Seite in der linken Spalte die Ordnung *biān ... jūn* [1], *wān ... yàn* [3], in der rechten Spalte *kǒu ... wǎi* [2], *yízhōu ... zhù* [5], auf der zweiten Seite in der linken Spalte *yào ... yú* [4], in der rechten Spalte *zì ... zūn* [6]. Aus der Numerierung in Klammern ergibt sich die richtige Ordnung.

⁵Die Statistik basiert auf 62,4 % der Einträge (i.e. Seiten 1 mit 692, Buchstaben A mit 5). In Wirklichkeit sind es 125 Kategorien, darunter befinden sich aber Mehrfacheinträge wie z.B. *měishù* 美术 und *měi* 美 oder *shè* 摄 und *shèyǐng* 摄影, die beide jeweils das gleiche bedeuten.

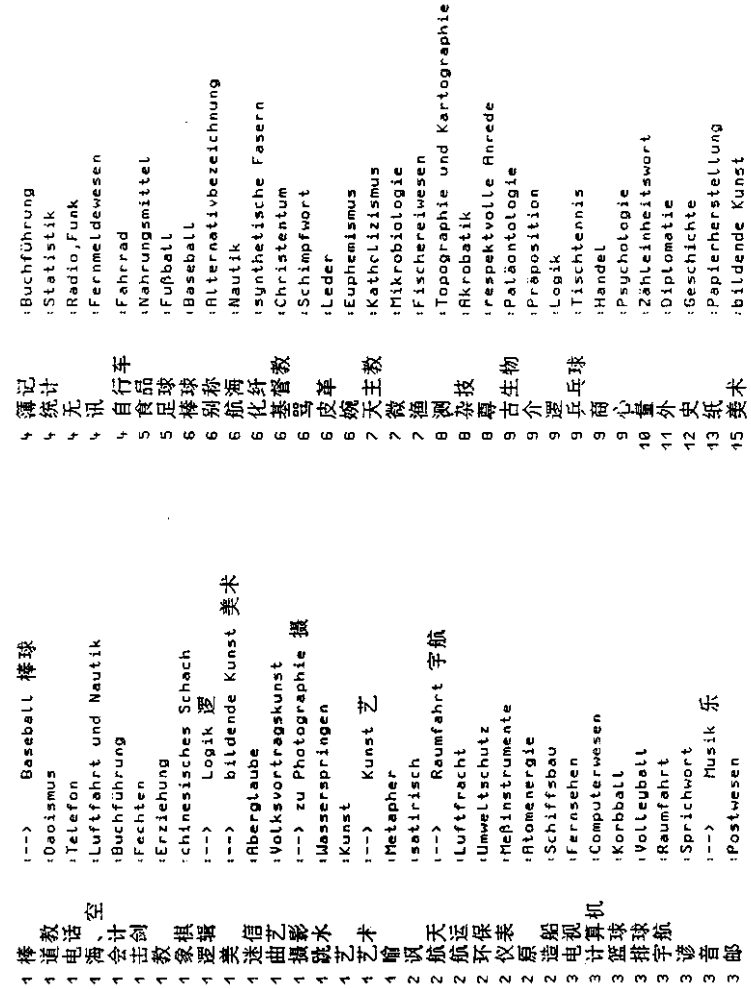


Abb. 4a: Kategorien des CDW und ihre Verteilung (1. Teil)

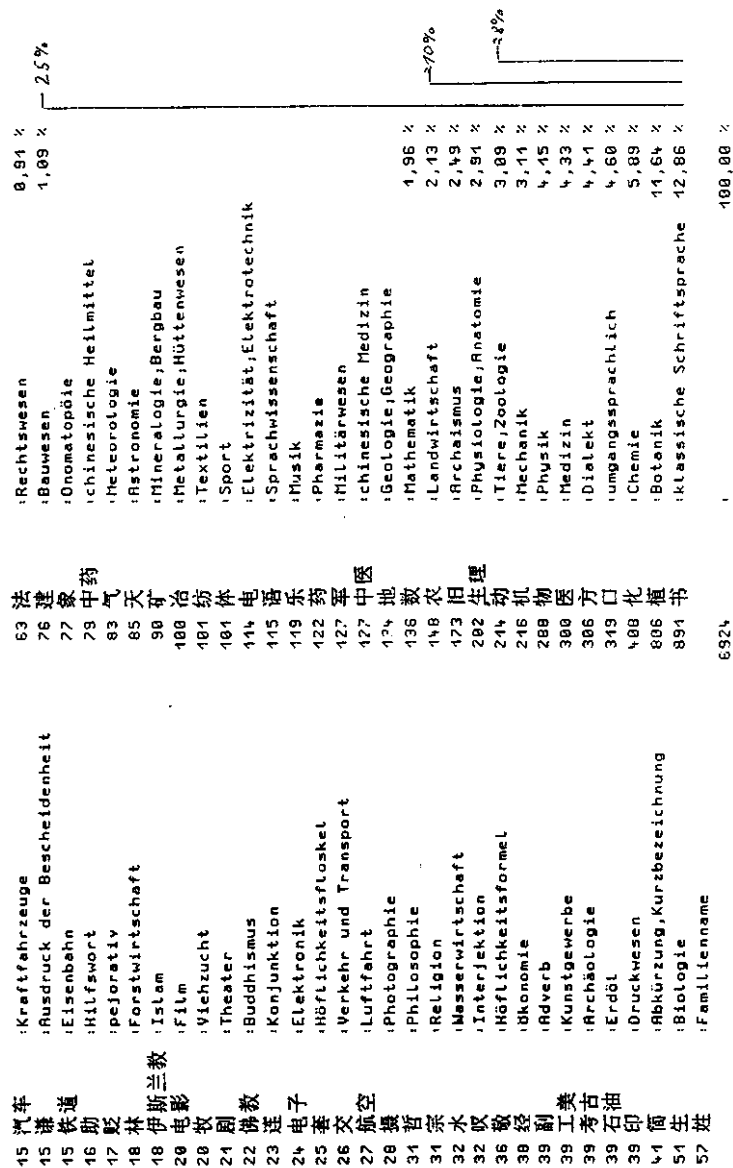


Abb. 4b: Kategorien des CDW und ihre Verteilung (2. Teil)

Von den in Hinblick auf Zuweisung von Kategorien für das elektronische Lexikon bereits bearbeiteten 31.516 Einträgen sind 6.924 mit Kategorien versehen, das sind 21,96 %⁶. Die Gegenüberstellung zweier unterschiedlich langer Teile der bisher bearbeiteten Lemmata, für die eine proportional gleiche Verteilung ermittelt wurde, läßt für das ganze Wörterbuch die gleiche Verteilung erwarten.

Für die Erstellung von Fachlexika ist die vorliegende Kategorisierung zu fein; sie weist auch einige Redundanzen und Kuriosa auf: Es gibt die Kategorie „Sport“ als Oberbegriff, aber auch „Korbball“, „Volleyball“, „Tischtennis“ usw. Um das gleiche Problem handelt es sich bei der Kategorie „Religion“, zu der wir auch „Christentum“, „Katholizismus“, „Islam“, „Buddhismus“ usw. finden.

Die Bedeutung, die in China dem Fahrrad zugemessen wird, kann man daran erkennen, daß es eine eigene Kategorie „Fahrrad“ mit nur 4 Einträgen gibt. Es sind dies:

- dǎngníbǎn 挡泥板 (Schutzblech)
- jiǎozhǎ 脚闸 (Rücktrittbremse)
- liàntào 链套 (Kettenschutz)
- liànzhào 链罩 (Kettenschutz)

Die fehlende Signifikanz der Kategorien in der vorliegenden Aufteilung für die Erstellung von Fachlexika läßt sich an folgendem sehen: Über die Hälfte der Kategorien – begonnen bei denen mit nur 1 Eintrag – stellen zusammengenommen nur 4 % aller mit Kategorien versehenen Einträge. Nur 29 Kategorien (25 %) aller mit Kategorien versehenen Einträge stellen mehr als 1 % der Vorkommen innerhalb der Kategorien, nur mehr 12 (9,44 %) stellen mehr als 2 % der Vorkommen. Bei über 3 % sind es die Kategorien Zoologie (3,09 %), Mechanik (3,11 %), Physik (4,15 %), Medizin (4,33 %), Dialekt (4,41 %), Umgangssprache (4,60 %), Chemie (5,89 %), Botanik (11,64 %), klassische Schriftsprache (12,86 %).

Um die Kategorien sinnvoll als Fachwortschätze zu nutzen, müssen noch die innerhalb von Bedeutungsfeldern vermerkten Kategoriebezeichnungen ausgewertet und alle Kategorien zusammengefaßt und auf etwa ein Dutzend reduziert werden. Die Berücksichtigung der Kategorien ist jedoch bei der Auswertung von Homophonen bei der Texteingabe mit *Pīnyīn* wichtig. Wenn man bedenkt, daß man bei der Erarbeitung von allgemeinsprachlichen modernen Texten durch das gezielte Abschalten der Kategorien „klassische Schriftsprache“, „Botanik“, „Chemie“, „Dialekt“, „Medizin“, „Physik“, „Mechanik“ und „Zoologie“ 5.600 Einträge weniger berücksichtigt werden und die Homophonie somit geringer ausfällt.

3. Das elektronische Lexikon

3.1. Umfang und Leistung

Das elektronische Wörterbuch hat mit 50.531 chinesischen Lemmata und ihren deutschen Übersetzungen genau den Umfang des CDW ohne Tildeneinträge. Die Tildeneinträge werden derzeit aufgenommen. Alle Angaben aus dem Erläuterungsteil, wurden – zum Teil zunächst mit Platzhaltern – übernommen; so sind z.B. 22 % aller chinesischen Einträge mit Markierungen der Zugehörigkeit zu den obengenannten Kategorien versehen.

⁶Hier sind nur die berücksichtigt, die in ihrer Ganzheit eine Kategoriebezeichnung tragen, also nicht solche bei denen nur ein (oder mehrere) Bedeutungsfeld gekennzeichnet sind.

Mit einem Programm, auf das wir weiter unten noch eingehen werden, wurden die Positionen der chinesischen Lemmata und der deutschen Wörter des Erläuterungsteils ausgetauscht. Unter Ausschluß einiger Füllwörter, der Artikel und satzverbindender Wörter⁷ ergab sich für den Erläuterungsteil die absolute Zahl von 171.774 Datensätzen. Nach Eliminierung von Mehrfachvorkommen kann das chinesische Korpus des elektronischen Lexikons nunmehr über 47.772 deutsche Stichwörter erschlossen werden. Da diese Stichwörter nicht von Hand nachbearbeitet wurden, sind in dieser Zahl auch alle morphologischen Ableitungen enthalten; neben „kommen“ etwa gibt es auch die Stichwörter „gekommen“, „komme“, „kommst“ usw. Tatsächlich sind es etwa 30.000 bedeutungsunterschiedliche deutsche Stichwörter. Ein Beispiel chinesischer Äquivalente für das deutsche Wort „Computer“ gibt Abb. 5.

1	elektronisches Gehirn; Computer m 电脑 diànnǎo
2	Rechenmaschine f; Rechner m; Computer m: 计算机 jìsuànjī

Stichwort: Computer Einträge: 2

Abb. 5: Chinesische Äquivalente für „Computer“

Wie man sich vorstellen kann, werden in der Regel jeweils mehrere chinesische Übersetzungen angeboten, unter denen das gewünschte Äquivalent auszuwählen ist. Es sind selten so wenige wie in diesem Beispiel. Das deutsche Stichwort „machen“ erschließt z.B. 742 chinesische Wörter und Redewendungen. Das Programm zeigt jeweils 10 Übersetzungen auf einmal; die nächsten Fenster können durch Betätigung der „+“-Taste angezeigt werden (Abb. 6).

⁷Darin sind enthalten: *jn, jm in, im so, ob du, er es an zu js am um, und ein, der, die, das, auf, von, für, mit, usw, als, sie, aus, dem, den, auf, zur, vor, man, bzw, des, ist, bei, wie, vom, ins, ich, wir, bis, hat, bei, zum, auch, etw, Der, Die, Das, Ich, Sie, Ein, sich, sein, wäre, wird, denn, oder, kann, eine, über, nach, sehr, alle, ohne, auch, einem, einer, eines, nicht, seine, unter, haben, einen, jedem, durch, würde, deren, selbst, werden, seinen, seiner, seines, seinem.* Da diese Stichwörter zu viele Belegstellen hätten, ist es sinnvoll, ihnen treffende chinesische Äquivalente von Hand zu zuzuordnen.

1	herstellen vt; machen vt; fabrizieren vt; produzieren vt; erzeugen vt: 制作 zhìzuò
2	(1) (*) [#] tun vt; machen vt: (2) (*) Werkstatt f: 作 zuò
3	(1) (*) tun vt; machen vt; verrichten vt; sich mit etw. beschäftigen: (2) (*) fertigen vt; erzeugen vt; schaffen vt: (3) (*) gründen vt; errichten vt; erbauen vt: (4) (*) verschaffen vt; auftreiben vt: (5) [#] herbeiführen vt: 搞 gǎo
4	(1) (*) tun vt; handeln vi; machen vt: (2) (*) dienen als; gelten als: (3) (*) zu etw. werden: (4) (*) sein vi; bedeuten vt: (5) <#> [#] : (6) <#> <#> [#] : 为 wéi
5	(1) (*) Hauptteil m; Stamm m; Rumpf m; Rückgrat n: (2) <#> (*) Kader m; Funktionär m: (3) (*) tun vt; machen vt: (4) (*) bekämpfen vt; kämpfen vi; sich zanken; streiten vi: (5) (*) fähig; tüchtig: 干 gàn
6	(1) (*) machen vt; schaffen vt; bauen vt; erzeugen vt: (2) (*) erfinden vt; aushecken vt; zusammenlügen vt; vorschwindeln vt: (3) (*) ausbilden vt; erziehen vt: (4) <#> (*) hingehen vi; sich (zu jm) begeben: (5) (*) eine von zwei Vertrags- oder Prozeßpart 造 zào
7	(1) (*) machen vt; herstellen vt: (2) (*) kochen vt; zubereiten vt: (3) (*) sich mit etw. beschäftigen; an etw.

Stichwort: machen Einträge: 742

731	honigsüße Worte machen , aber im Herzen einen Dolch tragen; Honig im Mund, aber einen Degen in der Hand; Honig im Mund und Galle im Herzen 口蜜腹剑 kǒumì-fùjiàn
732	eine stürmische See aufrühren - ein Übel noch ärger machen ; eine schlechte Lage verschlimmern 推波助澜 tuībō-zhùlán
733	alte oder weggeworfene Sachen reparieren und wieder gebrauchsfähig machen 修旧利废 xiū jiù lì fèi
734	jn hineinlegen; gegen jn intrigieren; jn fälschlich beschuldigen und ihn zum Opfer behördlicher Verfolgungsmaßnahmen machen 诬害 wūhài
735	(1) (*) jn anderswohin versetzen; etw. anderswohin schaffen: (2) (*) eine Dienstreise machen , um bei anderen Institutionen Auskünfte in personellen Fragen einzuholen 外调 wàidiào
736	Leere-Stadt-Kriegslist f (die Feinde irreführen, indem man die Tore einer schwach bewachten Stadt weit aufsperrt); nach außen hin kühn gegen jn Front machen , um eine schwache Verteidigung zu tarnen 空城计 kōngchéngjì
737	auch die tüchtigste Hausfrau kann kein Mahl ohne Reis bereiten - ohne die notwendigen Bedingungen ist nichts zu machen 巧妇难为无米之炊 qiǎofù nán wéi wú mǐ zhī chuī
738	auf dem Boden einen Kreis zeichnen und ihn zum Gefängnis

Stichwort: machen Einträge: 742

Abb. 6: Erstes und letztes Fenster von chinesischen Äquivalenten für das Stichwort „machen“

Die Suchstrategie ist natürlich wichtig: man würde nicht unter „schlimmer machen“, sondern unter „verschlimmern“ suchen, wenn man ein Äquivalent in dieser Bedeutung benötigt (Abb. 7 und 8).

- 1 sich **verschlimmern**; sich verschlechtern; sich zum Schlechten wenden: 恶化 èhuà
- 2 (1) (*) schwerer werden; erschweren vt; das Gewicht vermehren: (2) (*) etw. ärger machen; verschärfen vt; schlimmer werden; **verschlimmern** vt: 加重 jiāzhòng
- 3 eine stürmische See aufrühren - ein Übel noch ärger machen; eine schlechte Lage **verschlimmern** 推波助澜 tuībō-zhùlán

Stichwort: verschlimmern

Einträge: 3

Abb. 7: Chinesische Äquivalente für das Stichwort „verschlimmern“

- 1 (1) (*) schwerer werden; erschweren vt; das Gewicht vermehren: (2) (*) etw. ärger machen; verschärfen vt; **schlimmer** werden; verschlimmern vt: 加重 jiāzhòng
- 2 auf einer niedrigeren Stufe stehen; noch **schlimmer**; noch minderwertiger: 等而下之 dēng ér xià zhì
- 3 (*>) sogar; mehr noch; **schlimmer** noch: 甚至 shènzhì
- 4 die Lage wird immer **schlimmer**; es geht bergab; auf dem absteigenden Ast sein 每况愈下 měi kuàng yù xià

Stichwort: schlimmer

Einträge: 4

Abb. 8: Chinesische Äquivalente für das Stichwort „schlimmer“

Auch ein Benutzer mit nur geringen Chinesischkenntnissen kann aus dem deutschen Erläuterungsteil schließen, welches chinesische Äquivalent für ihn in Frage kommt. Dies soll an den verschiedenen Bedeutungen des deutschen Stichworts „aufheben“, gezeigt werden (Abb. 9)

- 1 **aufheben** vt; einer Sache ein Ende setzen; abschaffen vt; annullieren vt: 废止 fèizhǐ
- 2 **aufheben** vt; etw. außer Kraft setzen; etw. ungültig machen; annullieren vt: 废除 fèichú
- 3 **aufheben** vt; aufnehmen vt: 撮 duō
- 4 verbieten vt; ein Verbot gegen etw. erlassen; **aufheben** vt 取缔 qǔdì
- 5 abschaffen vt; widerrufen vt; **aufheben** vt; tilgen vt; etw. für ungültig erklären; etw. rückgängig machen; annullieren vt 取消 qǔxiāo
- 6 (1) (*) abschaffen vt; **aufheben** vt; beseitigen vt: (2) (*) jn entlassen; jn seines Postens (od. Amtes) entheben; jn aus einer Stellung entfernen; absetzen vt 革除 géchú
- 7 (1) (*) heben vt; **aufheben** vt; hochheben vt: (2) (*) (zu zweit oder mehreren) tragen vt: 抬 tái
- 8 (1) (*) aufbewahren vt; aufspeichern vt; **aufheben** vt: (2) (*) Reserve f; Vorrat m: 储备 chúbèi
- 9 (1) (*) sammeln vt; ansammeln vt; anhäufen vt: (2) (*) aufkrepeln vt; **aufheben** vt; hochstreifen vt: (3) (*) (Beld) zusammenraffen vt; etw. an sich reißen: 搂 lōu
- 10 (1) (*) etw. mit einer Hand tragen: (2) (*) hochheben vt; **aufheben** vt: (3) (*) einen Termin vorverlegen: (4) (*)

Stichwort: aufheben

Einträge: 32

Abb. 9: Chinesische Äquivalente für das Stichwort „aufheben“ (nur erstes Fenster)

Sobald auch die Beispielsätze des CDW voll in das elektronische Lexikon eingebunden sind, fällt dem Benutzer die Entscheidung noch leichter; oft werden dann ganze Beispielsätze kopierbar sein.

Ein weiterer Unterschied zu herkömmlichen fremdsprachlich-chinesischen Lexika ist der, daß den chinesischen Schriftzeichen die Aussprache in der richtigen Orthographie mit Tönen beigegeben ist. Dies ist auch für den Fortgeschrittenen in Anbetracht der Mehrfachlesungen vieler Schriftzeichen sehr hilfreich.

Es kann auch die deutsche Übersetzung chinesischer Wörter nachgeschlagen werden. Dazu hat der Benutzer zwei Optionen zur Auswahl:

- (1) Die chinesischen Stichwörter werden unmittelbar über Tastatur eingegeben; dabei stehen verschiedene Eingabeverfahren zur Verfügung (Pinyin, GB-Code, Randzeichen, Telegraphencode).
- (2) Die chinesischen Stichwörter können durch gezielte Positionierung des Cursors aus einem vorhandenen Text, der Schriftzeichen enthält, abgerufen werden.

In beiden Fällen wird die deutsche Kurz-Übersetzung samt Umschrift mit Tönen des ausgesuchten Zeichens und aller Lexikoneinträge, die dieses Zeichen an beliebiger Stelle enthalten, angezeigt. „Blättern“ kann man vorwärts und rückwärts. Unter „Kurz-Übersetzung“ ist eine automatisch gewonnene prägnante Übersetzung eines chinesischen Lemmas zu verstehen; Bedeutungsnuancen werden hierbei aus Platzgründen nicht

berücksichtigt, um bis zu zehn Einheiten gleichzeitig auf dem Monitor anzeigen zu können. Diese Kurz-Übersetzungen reichen aber in der Regel zur vorläufigen Orientierung aus. Der volle „Erläuterungsteil“ wird angezeigt, wenn man aus den Komposita eines Schriftzeichens ein bestimmtes per Platzziffer auswählt.

Neben der Übersetzung einzelner Begriffe kann der Benutzer auch Wort-für-Wort-Übersetzungen ganzer Sätze anfertigen lassen (Abb. 10 und 11).

他	tā	er
打算	dǎsuàn	planen
今天	jīntiān	heute
晚上	wǎnshàng	abends
请	qǐng	bitten
女朋友	nǚpéngyǒu	Freundin
去	qù	(hin)gehen
吃饭	chīfàn	essen
或	huò	oder
去	qù	(hin)gehen
跳舞	tiàowǔ	tanzen

Zeilenende Seite 802 Stamm-Nr.: 36765

Abb. 10: Erstes Beispiel einer Wort-für Wort-Übersetzung

Diese Wort-für-Wort-Übersetzungen helfen dem fortgeschrittenen Benutzer beim Verständnis unbekannter Wörter, und selbst jemand mit äußerst beschränkten Chinesischkenntnissen kann erkennen, um was es in dem Text geht. In der vorliegenden Form können die Daten als rudimentäre Hilfe bei der Übersetzung gebraucht werden; sie bilden sicher die Grundlage für die Entwicklung mächtiger Übersetzungssysteme.

她	tā	sie [f]
说	shuō	sprechen
汉语	Hànyǔ	Chinesisch
说	shuō	sprechen
得	dé	und zwar, [dei:] müssen
不错	bùcuò	nicht schlecht
,		
尤其	yóuqí	besonders
发音	fāyīn	Russprache
很	hěn	sehr
好	hǎo	gut
.		

Zeilenende Seite 322 Stamm-Nr.: 14596

Abb. 11: Zweites Beispiel einer Wort-für Wort-Übersetzung

Zusätzlich zu den rein übersetzungsbezogenen Funktionen bietet das elektronische Lexikon auch einen Einblick in die Kombinatorik der chinesischen Schriftzeichen. Der Zugriff auf alle Einträge, die ein bestimmtes Schriftzeichen an beliebiger Stelle enthalten, läßt Aussagen über die Produktivität der einzelnen Schriftzeichen, auch im Hinblick auf deren Position im Wort, bei der Wortbildung zu. Diese Funktion ist für Lehrende und Lernende, Lehrbuchautoren und Sprachforscher gleichermaßen interessant. Es steht hier sozusagen ein traditionell angeordnetes und ein rückläufiges Wörterbuch in einem zur Verfügung (Abb.12).

Gegenüber rückläufigen Wörterbüchern hat man den Vorteil, über alle Positionen eines Zeichens Zugriff auf entsprechende Einträge zu haben; im Unterschied zum *Chángyòng gòucí zìdiǎn*⁸, das auch diese Möglichkeit bietet, wird beim elektronischen Lexikon des *Münchner Systems* außerdem eine deutsche Übersetzung gegeben. Einträge, deren deutsche Übersetzung für die Bildschirmbreite zu lang ist, können über die Platzziffer und eine Funktionstaste in extenso angezeigt werden.

⁸ *Chángyòng gòucí zìdiǎn* 常用构词字典, 1982, von Fù Xīnglíng und Chén Zhānghuàn, Běijīng, Zhōngguó Rénmín Dàxué chūbǎnshè.

1	子	zǐ	Kind
2	子	zǐ	
3	子畜	zǐchù	junges Tier
4	子代	zǐdài	Spaltungsgeneration
5	子弹	zǐdàn	Patrone
6	子弟	zǐdì	Söhne und jüngere Brüder
7	子房	zǐfáng	Fruchtknoten
8	子妇	zǐfù	Sohn und Schwiegertochter
9	子宫	zǐgōng	Gebärmutter
10	子规	zǐguī	Kuckuck

Zeichen: GB-Code: 5554 Anzahl: 938

931	被管制分子	bèiguǎnzhìfēnzǐ	überwachtes Element
932	虎毒不食子	hǔ dú bù shí zǐ	vor den eigenen Kindern macht di
933	虎父无犬子	hǔfù wú quǎnzǐ	die Größe des Vaters geht auf de
934	打肿脸充胖子	dǎzhǒng liǎn chōng pàngzi	sich mit durchgewalk
935	以子之矛，攻子之盾	yǐ zǐ zhī máo, gōng zǐ zhī dùn	mit sein
936	有其父，必有其子	yǒu qí fù, bì yǒu qí zǐ	wie der Vater, so
937	山中无老虎，猴子称大王	shānzhōng wú lǎohǔ, hóuzi chēng dàiwáng	ist kein Tiger me
938	不入虎穴，焉得虎子	bù rù hǔxué, yān dé hǔzǐ	man kann kein

Zeichen: GB-Code: 5554 Anzahl: 938

Abb. 12:

Ein Beispiel zur Kombinatorik des Schriftzeichens 子
Erstes und letztes Fenster der 938 Einträge des elektronischen Lexikons

Abgesehen von den genannten Möglichkeiten, kann man einen chinesischen Lexikoneintrag auch über seine Adresse (Seitenzahl des CDW plus n-tes Lemma) im Lexikon abrufen; dort wird ebenfalls die deutsche Übersetzung in voller Länge angezeigt. Dieses „Blättern“ im Lexikon kann durch den Parameter der Wortlänge (Silbenanzahl eines Wortes – möglich sind 1..13; 15) gesteuert werden. Man kann z.B. durch das Lexikon gehen und gezielt alle Einträge – etwa ab einer bestimmten Seite – die aus vier Schriftzeichen bestehen, abrufen (Abb. 13).

一步登天	yī bù dēng tiān	mit einem Schritt in den
一差二错	yīchā-èrcuò	im Fall von Fehlern; bei
一场春梦	yī chǎng chūnmèng	ein Frühlingstraum; ein
一唱百和	yī chàng bǎi hè	beginnt einer zu singen,
一唱一和	yīchàng-yīhè	singen die einen, stimmen die
一尘不染	yī chén bù rǎn (1) (*)	von keinem einzigen
一成不变	yī chéng bù biàn	unveränderlich bleiben;
一筹莫展	yī chóu mò zhǎn	(sich) keinen Rat wissen;
一触即发	yī chù jí fā	beim ersten Anstoß sofort
一触即溃	yī chù jí kuì	schon beim ersten Schlag
一锤定音	yī chuí dìng yīn	der Letzte Hammerschlag
一蹴而就	yī cù ér jiù	etw. über Nacht vollenden; im
一刀两断	yī dāo liǎng duàn	mit einem Schwerthieb
		spalten; entschieden mit jm/etw. brechen;

Chinesisch-Deutsches Wörterbuch Seite 949 Stamm-Nr.: 43604

Abb. 13:

Beispiel für viersilbige Einträge des CDW, die mit yī — beginnen

Abgesehen von den translatorischen und kombinatorischen Funktionen können mit dem Lexikon-Programm auch chinesische Schriftzeichen-Texte analysiert werden, wobei die Verknüpfung der chinesischen Schriftzeichen zu Wörtern oder Kontextgruppen in den Formen der chinesischen Einträge des Wortspeichers erfolgt. Dabei werden identifizierte Wörter oder Wortgruppen ins Deutsche übersetzt (Kurzform) und mit *Pinyin* (mit Tonzeichen) versehen. Hier wurden die Voraussetzungen für eine Auswertung von über Scanner gewonnene Schriftzeichentexte geschaffen, in denen die Wortgrenzen ja nicht markiert sind. Dabei sind allerdings noch einige Probleme zu lösen: Gegenwärtig werden die Einträge nach dem „longestmatch“-Verfahren im Wortspeicher gesucht, und da kann es vorkommen, daß Schriftzeichen fälschlicherweise zu Wörtern oder Kontextgruppen verbunden werden. Die Schriftzeichenkette *fāzhǎn / Zhōngguó jiātíng / zhèngcè* 发展/中国/家庭/政策 (die chinesische Familienpolitik voranbringen) würde fälschlicherweise als *fāzhǎn zhōng guójiā / tíng / zhèngcè* 发展中国国家/庭/政策 (Entwicklungsland/Hof/Politik) analysiert. Durch Markierung der Ambiguität vom Programm aus kann der Benutzer auf die Mehrdeutigkeit und somit auf etwaige Fehler aufmerksam gemacht werden.

Da das Wörterbuch in ein Programm für chinesische Textein- und -ausgabe integriert ist, können die nachgeschlagenen Einheiten unmittelbar in den Text übernommen und mit abgespeichert werden. Diese Funktion macht das Programm für Übersetzer in beiden Sprachrichtungen interessant: Das Nachschlagen erfolgt bei der Texterstellung direkt am Bildschirm. Von den Bildschirmanzeigen können Papieraudrucke angefertigt werden.

3.2. Datenaufnahme

Die Aufnahme der Kapitelzeichen, der Polynome und der Erläuterungsteile erfolgte getrennt: Zunächst wurden in der Reihenfolge des *CDW* die Kapitelzeichen in eine separate Datei aufgenommen. Die Aufnahme der 6.647 Zeichen erfolgte im Dialogmodus mit sofortiger Richtigstellung; die Schriftzeichen wurden mit den Informationen des GB-Codes und der Silbenkennzahlen abgespeichert. In eine andere Datei wurde nur die Pīnyīnform (ohne Töne) der Kapitelzeichen und der zu jedem Kapitelzeichen gehörenden Polynome aufgenommen.

Die Umwandlung der einzelnen Pīnyīn-Artikel in Schriftzeichen erfolgte halbautomatisch. Dazu wurden beide Dateien synchronisiert und der im *Münchener System* bis dato vorhandene Wortspeicher aktiviert. Zur Identifikation dienten die isolierten Kapitelzeichen. Immer wenn in der Pīnyīn-Datei eine einzelne Silbe angetroffen wurde, wurde jeweils auf ein neues Kapitelzeichen zurückgegriffen und die jeweils erste Position der Polynome aus der Pīnyīn-Datei mit diesem Schriftzeichen besetzt bis wieder eine isolierte Silbe auftrat. Waren im alten Wortspeicher Einträge vorhanden, deren Aussprache mit den neuen Pīnyīn-Einträgen identisch war und deren erste Position mit dem zutreffenden Kapitelzeichen übereinstimmte, so wurden auch die weiteren Schriftzeichen dieses Eintrags automatisch vergeben.

Von Hand mußten dann nur noch die zweiten und weiteren Positionen derjenigen Einträge richtiggestellt werden, die nicht im alten Speicher enthalten waren. Dieser Korrekturlauf lief automatisch ab. Der Cursor sprang von einer als mehrdeutig markierten Stelle zur nächsten, und das Auswahlménü für die zu prüfenden Silben wurde selbsttätig geöffnet. 60 % der angebotenen Schriftzeichen waren von vorneherein korrekt. Das heißt, es brauchten nur weniger als die Hälfte der Schriftzeichen ab der 2. Position korrigiert werden. Zur Kontrolle im Falle möglicher Abweichungen diente die in der Pīnyīn-Datei enthaltene Angabe der Seitenzahl aus dem *CDW*.

Die Eingabe des Erläuterungsteils (im wesentlichen die deutschen Übersetzungen) erfolgte wiederum unabhängig von den beschriebenen Arbeitsgängen in der Reihenfolge des *CDW*. Dazu wurde in eine mit den Stammmummern der *CDW*-Einträge (das sind die Ordnungsnummern in aufsteigender Reihenfolge des Lexikons) versehenen Datei zu jedem Eintrag die deutsche Übersetzung aufgenommen, wobei zunächst nur die im *CDW* mit Pīnyīn versehenen Haupteinträge aufgenommen wurden.

Für die Aufnahme der „Tildeneinträge“, die derzeit ergänzt werden, wird das gleiche Verfahren angewendet.

3.3. Datengenerierung, Organisation

Der Speicherbedarf des *Münchener Systems* inklusive Lexikon beträgt 7 MByte auf der Festplatte; davon werden für das Wörterbuch benötigt:

chinesische Einträge	655,2 KB
Zeigerdateien	721,14 KB
deutsche Stichwörter	515,55 KB
deutsche Kurzeinträge	911,89 KB
deutsche Einträge	2.889,63 KB.

Zur Erweiterung um eine weitere Fremdsprache sind jeweils 4,3 MB nötig.

Die 50.531 chinesischen Einträge des Lexikons sind in der Form einer relationalen Datenbank organisiert; sie haben eine der Anordnung im *Chinesisch-Deutschen Wörterbuch* entsprechende Stammmummer. Der chinesischsprachige Teil der einzelnen Lexikoneinträge ist nach der Wortlänge (= Silbenzahl) in 15 Dateien eingeteilt, die folgende Informationen enthalten: Stammmummer; Wortart; GB-Codes, Silbenkennzahlen, Silbenton, Silbenstatus⁹ der einzelnen Silben. Bei der Übersetzung einer chinesischen Zeichenfolge ins Deutsche beginnt die Suche beim größtmöglichen String (15 Silben), bis in einer der 14 Dateien (es gibt keine 14-silbigen Polynome) ein Eintrag gefunden wird. Der passende Eintrag wird mittels Zeiger über die Stammmummer gefunden. Die deutschen Einträge (je unterschiedlicher, beliebiger Länge, ohne 255-Byte-Schranke) liegen in einer ASCII-Datei vor, wobei die einzelnen Einträge durch die Stammmummer und einen Marker voneinander getrennt sind.

Der Speicherzugriff hängt vom Prozessor des Computers, seiner Taktfrequenz und auch davon ab, ob ein virtueller Speicher (RAM-Disk) eingerichtet ist. Auf einem AT oder AT-kompatiblen Computer mit einem Intel 80386 Prozessor einer Taktfrequenz von 16 MHz und einer Festplatte mit 40 ms mittlerer Zugriffsgeschwindigkeit dauert die Suche nach einem beliebigen Eintrag weniger als 0,2 Sekunden.

3.3. Das elektronische Lexikon im Vergleich mit anderen deutsch-chinesischen Lexika

Da die chinesischen Einträge aus einem chinesisch-fremdsprachigen Lexikon gewonnen wurden, weisen die chinesischen Übersetzungen der deutschen Stichwörter größere Authentizität auf als in herkömmlichen deutsch-chinesischen Wörterbüchern, wo man sich oft nur mit Umschreibungen, die noch nicht einmal als solche gekennzeichnet sind, zufrieden geben muß. Da das elektronische Lexikon den Kontext (cf. Abb. 6) für die verschiedenen Bedeutungen eines Wortes mit anzeigt, kann das kontrollierende Nachschlagen in einem Chinesisch-Deutschen Lexikon unterbleiben. Insbesondere Sprichwörter und Redewendungen, hier vor allem die *Chéngyǔ*, haben in dem elektronischen Wörterbuch breiten Raum und helfen das Translat stilistisch zu glätten. Verb-Komplement-Verbindungen, etwa *xiǎngqǐ lái* 想起来 sind gegenwärtig noch – wie in herkömmlichen deutsch-chinesischen Wörterbüchern – relativ selten. Sie werden aber durch die bevorstehende Einbindung der Tildeneinträge und Beispielsätze ergänzt werden.

4. Ausblick

Das elektronische Lexikon gewinnt an Funktionalität, wenn es in Verbindung mit auf Diskette gespeicherten Textdaten genutzt wird, die nach Textsorten unterteilt sein sollten. Wir denken z.B. an die Textsorten „Wissenschaftssprache“, „Literatur“, „Umgangssprache“ und „Zeitung“.

Der Benutzer kann sich dann über Zuschalten von als GB-Codes gespeicherten Dateien – nach Textsorten oder sogar nach einem bestimmten Artikel, Autor oder gar bestimmtes Werk näher spezifiziert – alle Belege zeigen lassen, in denen ein bestimmtes Schriftzeichen oder Wort vorkommt. Für die verschiedensten Anwendungen wird dies fruchtbar sein. Denkbar ist auch die satzorientierte Suche, also z.B. nach Sätzen, in denen etwa sowohl *gāoxìng* 高兴 (sich freuen) als auch *lǐ wù* 礼物 (Geschenk) steht. Die Satzstellung für eine (oder mehrere verschiedene) chinesische Übersetzung(en) des deutschen Satzes „Ich habe

⁹Der Silbenstatus ist eine Ziffer, die die Schreibung der Silbe in *pīnyīn* repräsentiert – also großer bzw. kleiner Anfangsbuchstabe, Bindestrich am Silbenende, Status der Zusammenschreibung mit der nächsten Silbe.

mich über dein Geschenk sehr gefreut“ wird dann automatisch mitgeliefert. Ganz interessant wird es, wenn man nach einer Kombination von chinesischen Äquivalenten sucht, von denen man nur deren deutsche Stichwörter eingibt: das Angebot wird stark eingengt und ist dann auch schon zutreffend – Man sucht beispielsweise nach Äußerungen, in denen die chinesischen Äquivalente von „Bett“ und „machen“ enthalten sind und erhält vollständige chinesische Sätze, ohne auch nur ein chinesisches Wort selbst eingetippt zu haben.

Durch die Kombination mit authentischen Texten werden Lücken im elektronischen Lexikon spielend entdeckt, und das Lexikon kann spielend erweitert werden – und zwar mit authentischem Material. Hier liegt eine große Herausforderung. Wir können es nicht leisten, so große Mengen chinesischer Texte einzugeben. Dies sollte überhaupt zentral geschehen, um Mehrfachaufnahmen zu vermeiden. Wir hoffen auf die Zusammenarbeit mit chinesischen Stellen. Die Texte brauchen nur in Disketten-Form, codiert mit dem Standard-Code der VR China, am besten noch mit Markern für die Wortgrenzen, vorliegen. Sie können von unserem Programm sofort ausgewertet werden.

Eine halbautomatische Wort-für-Wort-Übersetzung von in Diskettenform vorliegenden Texten ist jetzt bereits realisiert. Durch geeignete Schritte wird sich auch hier eine Verbesserung in Richtung auf ein Übersetzungssystem erzielen lassen. Das Korpus der chinesisch-fremdsprachlichen Wörterbücher des Fremdspracheninstituts Beijing, die für die Sprachen Englisch, Deutsch, Französisch, Italienisch, Spanisch und Russisch vorliegen, ist identisch. Die Erweiterung des elektronischen Lexikons um die englischen Stichwörter steht vor ihrer Vollendung. Erweiterungen um die genannten Sprachen erfordern im wesentlichen Tipparbeit. In Experimenten wird sich zeigen, ob das Chinesische als Metasprache bei der Übersetzung zwischen anderen Sprachen geeignet ist.

Literaturverzeichnis

- Chángyòng gòucí zìdiǎn* 常用构词字典, 1982, Fù Xīnglǐng, Chén Zhānghuàn. Běijīng: Zhōngguó Rénmín Dàxué chūbǎnshè.
- Ding T., Heinzl J.L., Lachner A. 1988, „Pinyin-Worteingabe – ein effizientes Eingabeverfahren für die chinesische Schrift“, in *Chinesisch und Computer* 1:12-29 (Bremen).
- Hànyǔ cíhuì de tǒngjì yǔ fēnxī* 汉语词汇的统计与分析 1985, ed. Běijīng yǔyán xuéyuàn yǔyán jiàoxué yánjiūsù.
- Hànyǔ pīnyīn cíhuì* 汉语拼音词汇 ²1964(1963), ed. Zhōngguó wénzì gǎigé wēiyuánhuì cíhuì xiǎozǔ.
- Das neue Chinesisch-Deutsche Wörterbuch* 1985, ed. Běijīng wàiguóyǔ xuéyuàn Déyǔxì [Nebent.: 新汉德词典].
- Vietze, Hans-Peter 1987, „Ein Beispiel des Nutzens rechnergestützter Lexikographie“, in *Z. Phon. Sprachwiss. Kommunik.forsch. (ZPSK)* 40,6:780-785.
- Xiàndài Hànyǔ cídiǎn* 现代汉语词典 1972, ed. Zhōngguó kēxuéyuàn yǔyán yánjiūsù cídiǎn biānjǐshí.
- Xiàndài Hànyǔ pínǜ cídiǎn* 现代汉语频率词典 1986, ed. Běijīng yǔyán xuéyuàn yǔyán jiàoxué yánjiūsù.
- Zhang Shuying, 1981, „Neues Konzept einer Schreibmaschine für chinesische Schrift“ (Diss. München).
- Zhōnghuá Rénmín Gònghéguó guójiā biāozhǔn xìnxī jiāohuàn yòng hànzi biānmǎ zifú jí. Jīběn jí. GB 2312-80.* 中华人民共和国国家标准信息交换用汉字编码字符集. 基本集. Běijīng 1981.

Abstract

On the basis of the *Das neue chinesisch-deutsche Wörterbuch* (新汉德词典) compiled by the Beijing Foreign Languages Institute in 1985, the Department of Precision Instruments of the Technical University of Munich has developed an electronic thesaurus. Its 50,531 entries are made accessible by means of character-codes, their *pīnyīn*-forms and 47,772 German keywords. The expansion of the thesaurus on the basis of the *Chinese-English dictionary* (1978) will soon be accomplished.

As the electronic thesaurus is incorporated into the *Münchner System*, a Chinese character input-output system based on *pīnyīn* word input, it can be accessed at any phase of text-editing. Entries of the thesaurus which have been looked up in the thesaurus can directly be entered into the text.

In addition to the simple function as a dictionary, the system can analyze Chinese character texts stored in the form of GB-Code on floppy-disks, identifying the ends of words and finding context groups; the identified dictionary entries are also translated into German. Moreover, all dictionary entries in which a chosen character appears in any position (positions 1 up to 16 are possible) can be shown with their German equivalents.

The *Münchner System* and the electronic Chinese-German thesaurus can be run on IBM/AT compatible computers under MS-DOS; the size of the data bank is 7 MBytes on the hard-disk. The storage access time is less than 0.2 seconds.

摘要

多功能汉德电子词典简介
蓝安东, 汉字乐

在北京外国语学院于一九八五年所出版的《新汉德词典》的基础上, 慕尼黑科技大学精密仪器系发展研制了一个多功能电子词典。它五万零五百三十一条词条可通过汉字编码、汉语拼音及四万七千七百七十二个德文词条查着。目前正在把《汉英词典》的范围扩充进来。

由于电子词典是「慕尼黑汉语拼音输入输出系统」的组成部分, 在输入汉语文章时, 可随时查词典以及把查获的词条直接引用用于文章。

除了纯粹的词典功能以外, 本系统亦可分析磁盘上以国标字符储存的文体, 从而识别词与词之间的界限; 识别出来的词条也提供德语翻译。除此之外, 还可通过某一组成汉字把词库的任何词条及其德语翻译寻获, 无论本汉字在词条中的位置。词条长短达到15个音节之长。

「慕尼黑汉语拼音输入输出系统」及其电子词典可安装于IBM/AT计算机或兼容机上, 整个系统存储量在硬盘上为7兆。存储器存取时间小于零点二秒。