

Transkriptionsverfahren zur maschinellen Verarbeitung deutscher und japanischer Diskurse

Michel, Wolfgang
College of General Education, Kyushu University

<https://hdl.handle.net/2324/9574>

出版情報：独仏文學研究. 31, pp.111-125, 1981-11. 九州大学独仏文学研究会
バージョン：
権利関係：

DOKU-FUTSU BUNGAJU KENKYU

Studien zur deutschen und französischen Literatur
Bulletin annuel des essais sur la littérature française et
allemande

Nr. 31 (1981)

Transkriptionsverfahren zur maschinellen
Verarbeitung deutscher und japanischer Diskurse

WOLFGANG MICHEL

KYUSHU-UNIVERSITÄT
FUKUOKA JAPAN

Wolfgang MICHEL

Transkriptionsverfahren zur maschinellen Verarbeitung deutscher und japanischer Diskurse

1.0 Zweck

Die in der Literatur üblichen Notationen gesprochener Sprache sind - mit einer Ausnahme - infolge ihrer Kodierungsform der elektronischen Speicherung und Verarbeitung fast unzugänglich. Andererseits setzt jede Verallgemeinerung im Bereich des von Individuum zu Individuum beträchtlich variierenden Kommunikationsverhaltens eine umfangreiche empirische Arbeit an möglichst vielen Beobachtungsfällen voraus. Dies und auch die unbestreitbaren Vorteile eines jederzeit zugänglichen Daten-Korpus zu Referenzzwecken legen es eigentlich nahe, eine Lösung zu suchen, die einerseits die technischen Bedingungen zur elektronischen Verarbeitungen erfüllt, ausbaufähig und möglichst reichhaltig ist und andererseits zugleich der Forderung nach Lesbarkeit und Benutzbarkeit gerecht wird.

Unter diesem Aspekt möchte ich den Kern eines bereits ein Jahr getesteten Verfahrens zur Notierung deutschsprachiger Kommunikation sowie zwei im Hinblick auf vergleichende Konversationsanalysen entwickelte Systeme für das Japanische vorstellen.

2.1 Bisherige Verfahren

EHLICH und SWITALLA hatten 1976 die bis dahin verbreiteten phonetisch-phonologischen sowie die diskursanalytischen Transkriptionssysteme einer genauen Diskussion unterzogen. Zu diesen Systemen gesellte sich zwei Jahre später eine Art synthetischer Zusammenfassung von HENNE und REHBOCK, der mit der

Abb. 1

<p>5</p>	<p>hA < 100 + und da hat er Aufgedreht den hAhn = o-hC KOPFNICKEN hD 108 was = (99) -gehau-n =</p>	<p>hC < BEUGT SICH LANGSAM VOR, GREIFT NACH BIERGLAS</p>	<p>f weiter bedächtig erzählend, stark rhythmisch glie- dernd, sehr kehlig in den Haupt- betonungen</p>	<p>fB dreht kurz den Kopf nach rechts, ordnet mit der l. Hand die Haare, dann bis 5,42 o-hA. hB schaut bei 108 kurz zu hD, dann langsam zurück zu hA, bleibt in dieser Haltung (immer noch vorgebeugt).</p>	<p>über, Druck oder Überdruck ?</p>
<p>2:30 .</p>	<p>hA 109 nich wAhr' denn</p>	<p>mf hD: 108: kurzer erstaunter Ausruf</p>	<p>mf überlegend, fra- gend</p>	<p>fC und fD: o-hA (fC vorgebeugt).</p>	
<p>6</p>	<p>< 1 geh-n ja = wieviel atÜ = sEchs R. HAND FRAGENDE GESTE c-(o</p>	<p>mf überlegend, fra- gend</p>	<p>mf 10f.: wiederauf- nehmend</p>	<p>hC faßt bei 5,1 sein Bierglas, trinkt bis 15 und setzt bei 22 ab.</p>	
<p>< 6</p>	<p>oder Acht oder was, FRAG. WEGWER- BEUGT LEHNT HEBT GESTEN. FENDE SICH SICH R. HINTEN BEW. VOR ZURUCK ARM</p>	<p>mf 10f.: wiederauf- nehmend</p>	<p>hC faßt bei 5,1 sein Bierglas, trinkt bis 15 und setzt bei 22 ab.</p>		

Notation 'segmentaler', 'suprasegmentaler Zeichen', 'sichtbarer personengebundener Ereignisse', 'nicht personengebundener Ereignisse', mit 'Bedeutungsangaben', 'Zeitzählung' und 'Angaben zum Gesprächsbereich und Gesprächstyp' bislang wohl sicher detaillierteste Entwurf (s. den kleinen Textausschnitt in Abb.1).

Die Vor- und Nachteile der einzelnen Verfahren stellen sich in praktischen Anwendungsversuchen schnell heraus. Während z. B. die in amerikanischen Publikationen angebotenen Notierungen - die meist nur auf das jeweils gestellte Problem ausgerichtet sind - einerseits aufgrund ihrer Nähe zu der Dialogdarstellung in fremdsprachlichen Lehrwerken leicht lesbar sind, eignen sie sich andererseits infolge ihrer Informationsarmut kaum als Arbeitsmaterial, wenn man die Fragestellung variiert. Zum Aufbau eines Datenkorpus zu Referenzzwecken taugen sie daher kaum. Demgegenüber setzt die Detailfülle in den Transskriptionen bei HENNE/REBOCK (s. Abb. 1) sowohl dem Umfang der auf diese Weise herstellbaren Transskripte als auch der Lesbarkeit und Nachvollziehbarkeit sehr enge Grenzen, von der Unmöglichkeit einer elektronischen Speicherung ganz zu schweigen. Da in jede Transskription überdies stets die Interpretationen des Transskribienten einfließen, wächst die Gefahr von Fehldeutungen sowohl während der Herstellung als auch beim Verarbeiten. Schnelle Orientierung ist unmöglich, zum Erfassen des Gesamtablaufes die Originalaufnahme unerlässlich. Einen sinnvollen mittleren Kurs zwischen diesen Polen schlagen die Freiburger, Bielefelder bzw. Düsseldorfer Notationen ein (s. EHLICH/SWITALLA, S. 86ff. , 95f. , 91f.). Da die beiden letztgenannten 'handschriftlichen' Notationsweisen leicht in das zu maschinellen Zwecken entwickelte Freiburger System aufgenommen werden können, beziehe ich mich auf dieses unter der hypothetischen Annahme einer bereits vollzogenen entsprechenden Verbesserung und konzentriere mich auf die Struktur sowie die Nutzungs- und Ausbaumöglichkeiten solcher als Datei ('File') gespeicherten Transkriptionen. Daß zumindest durch editorische Bearbeitung der zu 'Lesezwecken' erschienen Publikationen auch

Abb. 2

- (a) her die Jungrinder stationiert werden +, , + die über ein Jahr sind +, , + und wieviel tragende Färsen wir dann nachher verkaufen +, . das macht wirklich k + Spaß + k .

uuuuuuxxxxxxnn

k + ... + k k + ... + k

uuuuuuxxxxxxad

/k + (ja) ich k + möchte + k

uuuuuuxxxxxxac

k + (ja) + k

uuuuuuxxxxxxad

das k + unterstreichen + k .

uuuuuuxxxxxxaa

k + (bitte) das sollen sie ja + k .

uuuuuuxxxxxxad

wir haben + k ja als Wissenschaftler ständig einen sehr engen Kontakt mit den Praktikern . + g + für uns ist es

- (b) A A ich hab mir das überlegt .es isf doch vielleicht
 10 viel besser ,+ wenn ich zu dir rein komme +,
AB niest
 ich sitze da in meinemAuto . und mir pfeift der Wind
 nicht um die Ohren

A B (ach so) (nein) (nein) ./ ich ich fahr gern

15 ein bißchen...

A A (ach so) (ja) ,+ wenn du gerne fährst +,

A B (na ja) | (eben) (ja) (doch) (doch). |

ich möcht gern ein bißchen fahren |

A A ... (ah so) (ja) (ja) . |

20 ich wollte dir nur sagen s+ sonst wäre ich zu dir
 gekommen | (f + ne + f?) +s. |

A B | (f + och + f) | das is aber auch sehr

im Freiburger Umkreis die Vorteile bestimmter Notationsweisen anerkannt werden, demonstrieren zwei Ausschnitte aus den 'Texten gesprochener deutscher Standardsprache', Band 1 (1971, S. 201) und Band 3 (1975, S. 68).

2.2 Zur Transkription deutscher Diskurse

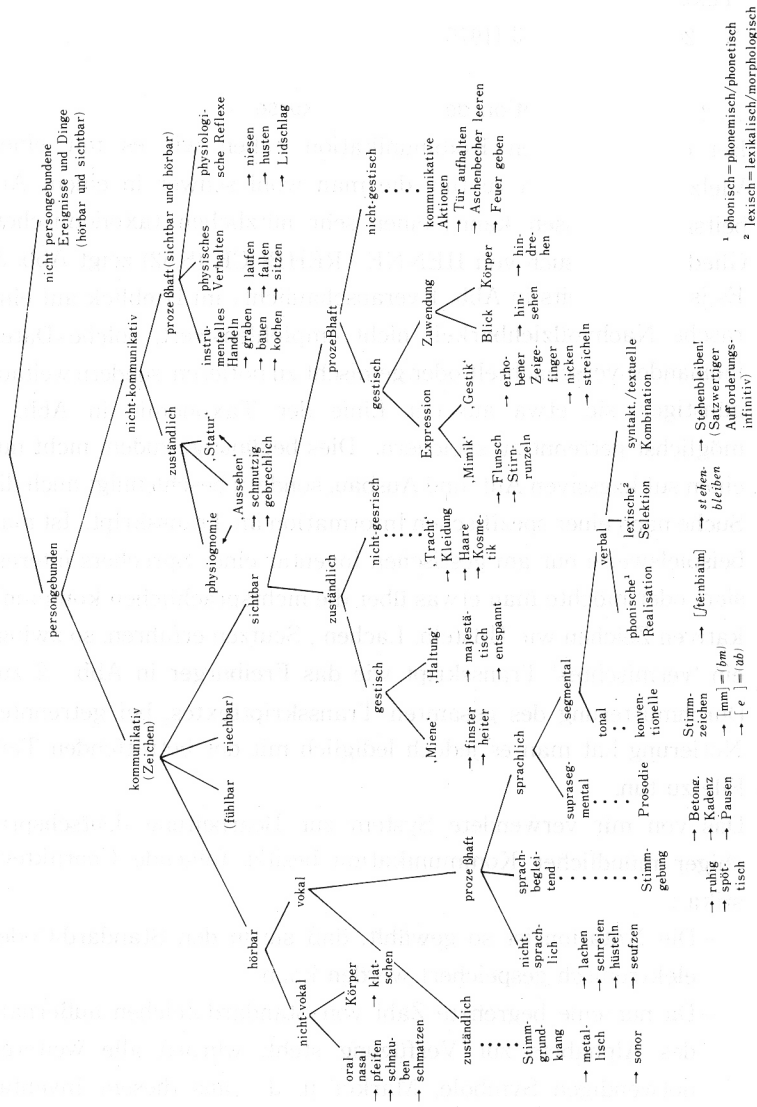
Bei der mündlichen Kommunikation haben wir es mit einer Vielzahl von Daten zu tun, die man wohl schwer in einem Arbeitsgang erfassen kann; einen sehr nützlichen taxonomischen Gliederungsversuch von HENNE/REHBOCK (S.62) zeigt Abb. 3. Es ist, wie bereits in Abb. 1 veranschaulicht, im Hinblick auf eine rasche Nachvollziehbarkeit nicht empfehlenswert, solche Daten ineinander verschachtelt oder gemischt zu notieren, sondern weitaus günstiger, sie etwa auf der Linie der Taxonomie in Abb. 3 möglichst getrennt zu speichern. Dies begünstigt zudem nicht nur einen sukzessiven Auf- und Ausbau, sondern beschleunigt auch die Suche nach einer spezifischen Information im Transskript. Ist man beispielsweise nur am gestischen Inventar eines Sprechers interessiert oder möchte man etwas über die nichtsprachlichen kommunikativen Zeichen wie Hüsteln, Lachen, Seufzen erfahren, so zwingt ein 'vermisches' Transskript wie das Freiburger in Abb. 2 zur Durchmusterung des gesamten Transskripttextes, bei getrennter Notierung hat man es jedoch lediglich mit der betreffenden Teil-File zu tun.

Das von mir verwendete System zur Bearbeitung deutschsprachiger mündlicher Kommunikation besitzt folgende Charakteristika :

- Die Notation ist so gewählt, daß sie in den Standard-Codes elektronisch gespeichert werden kann.
- Da nur eine begrenzte Zahl von Standard-Zeichen außerhalb des Alphabets zur Verfügung steht, wurden alle weiteren notwendigen Symbole, Marker u. ä. aus diesem Inventar gewählt bzw. kombiniert, und zwar so, daß sie technisch leicht über Kartenlocher, Display-Terminals bzw. optische Lesege-

Abb.3

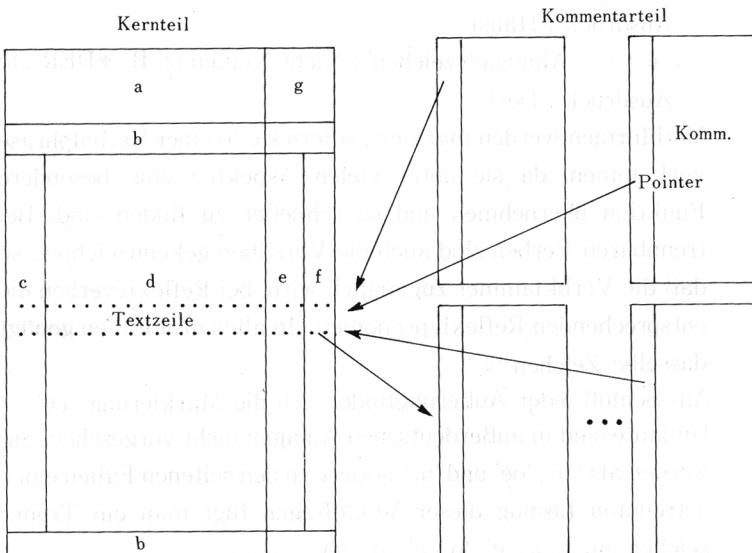
Taxonomie der Daten in der mündlichen Kommunikation



1 phonisch=phonemisch/phonetisch
2 lexisch=lexikalisch/morphologisch

- räte (OCR = Optical Character Reader) eingegebbar sind (Häufigkeit, Position auf den Tastaturen etc.).
- Die Darstellung der segmentalen lautlichen Seite, die den Kern des Transskriptes bildet, wurde weitgehendst von der Beschreibung der suprasegmentalen lautlichen Seite und der außersprachlichen Seite getrennt, um Lesbarkeit und schnelles Information-Retrieval zu gewährleisten.
- Die sich um diesen ‘Kernteil’ gruppierenden weiteren möglichen Kommentare, Beschreibungen etc. sind nach Kategorien getrennt in einem ‘Kommentarteil’ untergebracht. ‘Pointer’, d. h. Zeigefelder, in denen die Positionsangaben der Bezugsstelle enthalten sind, stellen dann die jeweilige Zuordnung sicher (s. Abb. 4).
- Da Diskurstransskripte auch im Hinblick auf Wortschatz- oder Syntaxanalysen u. ä. interessant sind, werden einige Kategorien bzw. Phänomene besonders markiert, um eine entsprechende Weiterverarbeitung zu erleichtern.

Abb. 4



Ein Transkriptkernteil hat die in Abbildung 4 veranschaulichte Form:

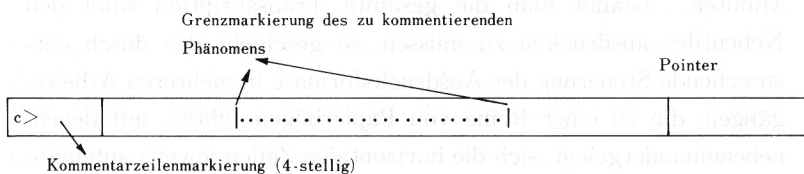
- (a) Angaben zum Gesprächsbereich, Gesprächstyp (Teilnehmer, Zeit, Ort, Anlaß, Art des Gespräches, Thematik, Beziehung der Teilnehmer, Aufnahmebedingungen, Angaben zur Transkribierung etc.)
- (b) Markierung des Diskursbeginns bzw. -endes
- (c) Bereich für die Sprecherkürzel (4 Zeichen pro Zeile)
- (d) Transskriptionstext (72 Zeichen pro Zeile)
- (e) Positionschiffre
- (f) Pointerstellen zu den Kommentarteilen
- (g) Reservestellen

Im Transkriptionstext selbst verfährt man wie folgt:

- Zum Ausdruck von 'Leseversionen' ist eine Differenzierung von Groß- und Kleinbuchstaben nötig, die die Rechenanlagen nicht direkt leisten können. Majuskel müssen besonders gekennzeichnet werden, wobei man noch weiter unterscheiden kann zwischen Nomina und Nicht-Nomina (z. B. am Satzanfang).
 - '#' + 'Alphabetszeichen' : Nomen (z. B. #HAUS ; im Ausdruck : Haus)
 - '*' + 'Alphabetszeichen' : Nicht-Nomen (z. B. *DER ; im Ausdruck : Der)
- Verbformen werden markiert, sofern sie in einer Verbalphrase vorkommen, da sie unter vielen Aspekten eine besondere Funktion übernehmen und so schneller zu finden sind. Bei trennbaren Verben sind auch die Vorsilben gekennzeichnet, so daß die Verbklammer zugänglich wird, bei Reflexivverben die entsprechenden Reflexivpronomen. In allen drei Fällen genügt dasselbe Zeichen '='.
- Am Schluß jeder Äußerung findet sich die Markierung '@'.
- Umlaute sind in außerdeutschen Anlagen nicht vorgesehen. Sie werden als 'ae', 'oe' und 'ue' notiert, in den seltenen Fällen einer getrennten Lesung dieser Vokalfolgen fügt man ein Trennzeichen ein ('a-e', 'o-e', 'u-e').

- Ähnliches gilt für das deutsche 'ß', für das man die Währungssymbole '\$' oder '¥' einsetzen kann.
- Vokaldehnungen werden durch Vermehrung des betreffenden Zeichens repräsentiert: 'ach sooo'.
- Nicht lexikalisierte Laute bzw. Lautfolgen werden grundsätzlich so notiert, daß die standardsprachliche Lesung der tatsächlichen Lautung am nächsten kommt.
- Besondere Akzentuierungen zeigt ein ':' vor der jeweiligen Stelle an.
- Nicht identifizierbare Laute, die potentiell eine kommunikative Funktion haben, werden durch '++', solche, die vermutlich zufällig geäußert wurden, mit '+' indiziert.
- Nicht lexikalisierte, doch identifizierbare Lautfolgen werden zusammen mit Kontakt- und Kommentarparenthesen in Klammern gefaßt : ' (also eee ja) '.
- Dialektformen sind durch 'D>' und '<D' eingerahmt, Eigennamen durch 'P>' und '<P', Nebensätze oder Nebensatzfolgen durch 'S>' und '<S'.
- Sprechpausen notiert man in Einheiten von 1/3 Sek. mit '-': 'alsoo ---'.
- In all den Fällen, wo man den Beginn und die Dauer eines suprasegmentalen oder nonverbalen Phänomens direkt im Kernteil anzeigen zu müssen glaubt, wird eine Kommentarzeile folgender Struktur eingeschoben:

Abb. 5



- Offenkundige Versprecher erkennt man an einem nachgestellten '*'.

- Zitate beginnen und enden mit Anführungszeichen .
- Um das Auffinden der Stellen zu erleichtern, an denen ein Sprecherwechsel stattfindet oder mehrer Sprecher parallel sprechen, ist die Grenzmarkierung ' | ' vorgesehen.
- Alle Wörter, alle Markierungs- und Satzzeichen sind von je einer Leerstelle vorne und hinten umgeben.

Nachfolgend ein kleines Textbeispiel in der Form, wie es im Speicher erscheint (Abb.6). Die Satzzeichen, sowie die Majuskel-Markierungen am Satzanfang dienen der Herstellung

Abb.6

```
MI *EINE #GRUNDLAGE - DAFUER =IST , S> DAS$ DIE #FRAGEN NEU =AUFKOMMEN ,
MI DIE MAN =GLAUBTE VOR ZEHN ODER FUENFZEHN #JAHREN F> AD ACTA <F
MI =LEGEN ZU =KDENNEN . <S |
DL:          | + P> #HERR #STAMMLER <P BITTE . |
ST:          | - *ALSO
ST ICH =GLAUBE , - + - S> ES =SIND MEHRERE #ARGUMENTE . <S *SIE =HABEN
ST ,SCHON -- (EE) - E!* EIN GANZ WICHTIGES =GENANNT . --- *ES =IST DIE
ST GERADEZU - !IRRATIONALE #ANGST , - S> DIEE - MIT DER ATOMAREN #KRAFT
ST =VERBUNDEN =IST . <S --- (EE) *DAS =ZEIGT =SICH BEI DN* - IM #FALLE
ST DER #KERNKRAFTWERKE GENAUSO - WIE BEI DER #ATOMBOMBE , -- UND (EHM)
```

von 'Leseausgaben'. In diesem Falle wird der Kernteil zunächst elektronisch kopiert und dann entweder über eine Serie von Display-Terminal-Kommandos oder aber ein kleines Rechenprogramm von allen nicht benötigten Markierungen (mit Ausnahme der Majuskel-Marker), Symbolen etc. 'gereinigt'. Danach folgt ein Majuskel-Programm, daß den Ausdruck über den entsprechenden Drucker steuert; das gesamte Verfahren dauert nur wenige Minuten. Glaubt man die gesamte Transskription samt den Nebenfiles ausdrucken zu müssen, so geschieht das durch entsprechende Steuerung der Ausdrucksformate in mehreren Arbeitsgängen, die zu einer Reihe von Papierbögen führen, auf denen, nebeneinandergelegt, sich die horizontalen Zeilen jeweils aufeinander beziehen; auch diese Arbeit ist ähnlich rasch getan.

Das System wurde etwa ein Jahr erprobt. Seine Vorzüge zeigen sich schnell, wenn man Texte aus anderen Systemen umzunotieren

versucht, wobei z. B. beim Freiburger Korpus immer wieder Unklarheiten hinsichtlich des genauen zeitlichen Ablaufes, des genauen Verlaufes von Sprecherwechseln etc. auftauchen, die man angesichts der auf der anderen Seite so detaillierten syntaktischen Markierungen nicht erwartet hätte. Auch scheint mir das Abbilden von Vokaldehnungen, die differenzierte Notation nicht-lexikalischer Lautäußerungen und der 'Mikropausen' unerlässlich. Beim Information-Retrieval bleibt der Verbal- und Nominalbereich des Freiburger Korpus mangels Markierung unzugänglich, die Suche nach außersegmentalen Phänomenen, Kommentaren o. ä., soweit sie überhaupt notiert sind, wird, wie schon weiter oben erwähnt, durch die Vermischung der Daten außerordentlich zeitraubend. Das spätere Eingliedern weiterer Befunde (Up-Dating) ist bei dieser Struktur ebenfalls mühselig, da jedes Mal Platz geschaffen werden muß, um solche Nachträge einzuschieben. Eine Parallelstruktur wie die vorgeschlagene hingegen ist offen für beliebige Ergänzungen, neue Teil-Files lassen sich jederzeit schaffen und über Pointer am Kernteil fixieren.

3.1 Zur Transskription japanischer Diskurse

Ein weiterer Vorteil der vorgeschlagenen Notation liegt darin, daß sie sich mit geringen Variierungen auch auf das Japanische anwenden läßt und den Einsatz derselben Computer-Programme möglich macht.

Wenn man japanische Transskripte nicht als potentielles Ausgangsmaterial für Lesetexte zu verwenden beabsichtigt, genügt durchaus eine latinisierte Notation, da die Sprecher in der mündlichen Kommunikation alles tun, um Mißverständnisse auszuschließen, wodurch die Angst vor der vielzitierten Homophonität des Japanischen fast belanglos wird. Kurze Auszüge in dieser Form verwandte beispielsweise HINDS in seinen Untersuchungen zur japanischen Konversationsstruktur (S. 88):

Abb.7

A:	soo ne, nan to naku, koo, tyotto, nitumatte kita mitai zya nai?	sonoo, sannin.
B:		soo.
C:		
A:	hitori ...	sorya soo, n n anoo, mitei to iu
B:	sono hito doomo ne, amari yoku nasasoo, watakusi ni totte wa ... tada sa, tosi ga	
C:		n n
A:	koto dake zya nakute, maa, tosi mo wakai kara, hoka ni mo torai dekiru ka mo si-	
B:	wakai to iu koto de,	
C:		zentaiteki ni mite mo ne,

Daß die HINDsche Darstellungsweise zwar illustrativ nützlich ist aber gehobenen Ansprüchen nicht genügt, dürfte aus dem weiter oben Gesagten hinreichend deutlich geworden sein. Zu Versuchszwecken wurden aus einem Korpus japanischer Diskussionen, den ich z. Zt. aufbaue (Korpusname: JADIS), Teile in romanisierter Form gespeichert. Die Umschreibung folgt im wesentlichen dem traditionellen Hepburnschen System. Die langen Laute sind durch Verdoppelung repräsentiert, nicht lexikalisierte Dehnungen durch entsprechende Vervielfachung des betreffenden Vokalzeichens. Auf das Markieren von Nomen und Verben wurde zunächst verzichtet. Während die Postpositionen (no, ni etc.) meist als selbstständige Wörter auftreten, sind die Verbalsuffixe der besseren Lesbarkeit halber über ein '=' mit dem jeweiligen Verb verbunden. Ansonsten gelten die weiter oben aufgeführten Konventionen, Umlaut und 'ß' sind nicht relevant. Die maschinelle Verarbeitung solcher Texte - einen Ausschnitt in der gespeicherten Form zeigt Abb. 8 - unterscheidet sich kaum von der deutscher Transkripte. Allerdings sind die Resultate von mehr grammatisch den diskursanalytisch orientierten Programmen, wie z. B. die automatische Wortformenregistrierung, zuweilen nicht ganz zufriedenstellend, da mit der Isolierung einzelner Worte das Homophonie-Problem wieder aktuell wird. Auch zur Herstellung langer 'Leseausgaben'

Abb.8

M1 NATTE=ORIMASU GA: - ZEIKIN NO OMOMI WO - HONTO NI - HI:SHIHISHI
M1 TO KANJITE=ORIMASU . |
IN: | + |
M2: | (EE) WATAKUSHI WAA - SUIDEN SAN
M2 HEKUTAARU WOO KOUSAKU SHITE=IMASU GA , - SAN NIN NO ROU*
M2 ROUDOURYOKU DE - ISSHOUKENMEI HATARAITEMOO - NENKAN SANBYAKU
M2 - GO ,ROKUJUU MAN TEIDO NO - SHUUNYU SHIKA AGERU KOTO GA
M2 DEKIMASEN . -- SONOO NANBAI NI MOO SOUTOU SURUU - ZEIKIN GA --
M2 GYOUSEI NO DESAKIKIKAN DE - KARASHUCHOU YAA - ZUSAN NA KEIRI NI
M2 YOTTEE --- TSUKAWARETE=IRU TTE IU KOTO NI TASHI + FUNGAI NI
M2 TAEMASEN . |
IN: | --- GOKATEI NO - SHU:FU WA - IKAGA DESU | KA .
F1: | - HAI -
F1 (ANOO) - KODOMO WO MOTSUU OYA TOSHITE . - MAINICHI - HOUDOU
F1 SARETE=IMASU -- GENZAI NO SONOO TSUKAIKATA NO FUSEISA TTE IU NO
F1 NI TAISHITE , - KODOMO NI - DOU SETSUMEI SHITE I I NO KA - HONTOU
F1 NI HAZUKASHII NA: TO + OMO* OMOTTE=ORIMASU .
IN: | ++ --- HAI - USHIRO NO
F2: | - WATASHI NO TOKORO MO SARARIIMAN
F2 DESU KARA , - MAIGETSU NO ZEIKIN WA KAKUJITSU NI -

und zum schnellen Durchlesen eignen sich romanisierte Transkripte nicht allzu sehr; selbst der westliche Leser erfaßt die mit ihren Flexionsendungen und Suffixen recht langen Gebilde nur langsam.

Technisch ist inzwischen die elektronische Speicherung und Verarbeitung von Kana-Texten mit chinesischen Schriftzeichen ('Kanji') kein Problem mehr. Eine Reihe von Computer-Herstellern bieten Zusatzeinrichtungen für japanische Texte an, z. B. JEF (Japanese Processing Extended Feature) von Fujitsu . Der Textausschnitt von Abb. 8 nimmt dann folgende Form an:

Abb. 9

男₁ : て来る—十年代に—なっておりますが：—税金の重みを—
男₁ : 本当に—ひ：しひしと感じております。 |
石野 : | + |
男₂ : | (エエ) 私
男₂ : はア水田さんヘクターをオ耕作していますが、—さん人
男₂ : の労*—労働力で—一生懸命働いてもオ—年間さんびやく
男₂ : —ご、ろくじゅう万程度の—収入しか上げることが出来ま
男₂ : せん。—そのオ何倍にもオ相当するウ—税金が—行政

男₂ の出先機関で—空出張やア—杜撰な經理によってエ—
 男₂ 使われているっていうことに対し+憤慨に耐えません。|
 石野： |
 石野 ——ご家庭の—主：婦は—いかがです|か。
 女₁： |—はい— (ア
 女₁ ノオ)—子供を持つウ親として、—毎日—報道されていま
 女₁ す—現在のそのオ使い方の不正さっていうのに対して、
 女₁ —子供に—どう説明していいのか—本当に恥ずかしいな：
 女₁ と+思*思っております|。
 石野： |++——はい—後の方にも伺い
 石野 ましょう。|
 女₂： |—わたしのところもサラリーマンですから、
 女₂ —毎月の税金は確実に—支払っております。+その—

Die Notationsprinzipien sind dieselben wie die für latinisierte Versionen, nur werden nicht-lexikalisierte Laute bzw. Lautfolgen in Katakana gehalten, wodurch die Lesbarkeit weiter steigt und die Ausgabe einer nahezu standardsprachlichen 'Leseversion' möglich ist. Morphologische Analysen verlaufen mit dem entsprechend intensiven Einsatz chinesischer Zeichen weitaus befriedigender, auch verlangen die Resultate von Wortschatzanalysen u. ä. weniger Nachbereitung als im vorherigen Fall. Der einzige bisher erkennbare schwache Punkt im Umgang mit solchen Transskripten ist die umständliche Eingabe über die zu diesem Zweck notwendigen speziellen Display-Terminals. Falls die Verdatung künftig noch beschleunigt werden könnte, scheint bei japanischen Diskursen diese Notation langfristig gesehen die vielversprechendste zu sein.

Literatur:

Texte gesprochenener deutscher Sprache I, II, III. Erschienen in der Reihe 'Heutiges Deutsch' als Band II/1, II/2 und II/3, München 1971, 1974 bzw. 1975.

Konrad Ehlich/Bernd Switalla: Transkriptionssysteme - Eine

exemplarische Übersicht. In : Studium Linguistik. Nr. 2,
Kronberg 1976.

Helmut Henne/Helmut Rehbock: Einführung in die Gesprächs-
analyse. Berlin 1978.

John Hinds: Conversational Structure: An investigation based on
Japanese interview discourse. In: Hinds/Howard:
Problems in Japanese Syntax and Semantics. Kaitaku-
sha, Tokio 1978.