

国立国語研究所学術情報リポジトリ

漢字プリンタによるターンアラウンド・システム

メタデータ	言語: Japanese 出版者: 公開日: 2020-06-29 キーワード (Ja): キーワード (En): 作成者: 斎藤, 秀紀 メールアドレス: 所属:
URL	https://doi.org/10.15084/00002870

漢字プリンタによるターンアラウンド・システム

齋藤 秀紀

0. 序

国立国語研究所（以下 国研）では各種の用語用字調査を進めてきたが、調査規模の拡大と期間の短縮を図るため、昭和41年にHITAC-3010（20KC）型コンピュータを導入し、「新聞の語彙調査」を実施した。さらに昭和49年「高校教科書の用語調査」が新たに発足するとともに、新型コンピュータHITAC-8250（98KB）にリプレースした。ついで50年には高速の漢字ラインプリンタNEAC-C5210を導入した。漢字プリンタは従来の漢字テレタイプ低速印字装置に変わるものとして導入したものである。本稿では、これら漢字プリンタ導入によって可能となった、新システムについて概観し、システムの基本的な考え方、従来のシステムとの相違点、適用可能な応用面について述べる。

1. 漢字プリンタ導入の背景

昭和41年に漢字入出力装置として導入された装置は、印字速度が毎分120字と非常に低速であった。漢字プリンタ導入は、これら印字装置の低速を補うものとして計画されたが、機器構成の検討を進めるうちに、汎用漢字処理システムへとより大きな機能を持ったものに要求が拡大していった。

一つには、従来の調査形態が大量処理向きのバッチ処理方式を基本としていたために、必ずしも、研究者の研究上の要請と一致していなかったことによる。

ターンアラウンド処理方式は、このような背景のもとで計画され、各装置は可能な限り、この線によって構成されている。特に、このターンアラウンド処理を行う上で留意した点は次の二点である。

- (1) プリント用紙は入力装置媒体として使用可能なこと。
- (2) 入出力媒体は通常の手作業で使用される形態に近いこと。

また、漢字プリンタに対して導入時点で要求した基本的機能は、次の五項目である。

- (1) 高速印字機能
- (2) 任意のコード体系への変換機能
- (3) 漢字ディスプレイによる情報の検索と校正機能
- (4) 外字挿入の容易性
- (5) 用紙はOMRまたはカードリーダ装置規格のものが使用可能なこと

その他、制御用ミニコンピュータについては、ホスト・コンピュータ側で新たに開発しなければならないソフトウェアを極力おさえるため、記憶装置及びコンピュータ能力が十分大きいことが要求される。

これらは総合的な漢字処理を指向する場合、コンピュータ側の周辺装置の接続能力に余裕があることが必要であり、将来のシステム拡張の面においても、コンピュータの余裕度を確認することは重要であ

る。

2. ハードウェア構成

漢字プリンタはNEAC-3200型ミニコンピュータによって制御されるスタンド・アローン形式である。各機器の機能は、磁気テープ装置について、HITAC側とのデータ交換、検索、校正処理、コード交換等プリンタ側で独立に処理される入出力用に使用する。印字処理のためのNEAC標準コードへの変換テーブルは磁気ドラムに格納するが、内容は、漢字テレタイプ盤内字2400字、盤外字約6000字分である。

他の一台のドラムについては、使用頻度の低い字種約3500字分の漢字パターンの格納に使用する。使用頻度の高い字種約5000字は高速文字発生装置(ROM)に登録するが、ROM内の5000字分の容量については、現有漢字テレタイプコードが12ビットであるため、盤内字2400字について、直接ハードによる印字処理が可能である。この場合、変換テーブルによるコード変換は不要となる。

漢字ディスプレイについては、付属の漢字キーボード、ライトペンによる、用例検索、原データの校正、に使用される。また、OMRシートとの併用によって、漢字パターン・デザイン処理も可能である。漢字パターン及びコードについては、漢字辞書によって統一的に扱われており、内容は次の項目から構成されている。

図書	シーケンス番号
漢字コード	NECコード
ユーザー・コード	現在使用せず
部首	部首コード
画数	漢字の総画数
分類	当用、補正、人名漢字等の判別記号
級	漢字セットの領域番号
漢字の読み	音訓別の代表的な読み
パターン	漢字のドットパターン

OMR装置、プリンタ装置による、ターンアラウンド処理に使用する場合、プリンタ印字方式が多針電極による湿式静電記録であるため、用紙はシートまたは80欄カード、いずれの場合も特殊加工されたものを使用する必要がある。その他プリンタの1行当りの印字数は多針電極方式の場合、ポイントに関係なく(9P、12P共に最大80字)一定であるため、ポイントを上げてのデータ印字が可能である。

3. ソフトウェアシステムの概要

従来の用語用字調査システムの処理形態が磁気テープを中心としたバッチ処理であったため、入力データ作成に多くの経費、期間などが必要であった。それは、調査対象となるデータの大部分がコンピュータ入力以前に人手によって処理され、集計、編集、データ管理等の比較的単純な作業にのみコンピュータが使用されていたためである。これらは、磁気テープによる処理形態であったため、システム構成的には無理な面が多かったことは否めない。しかし、最近の周辺装置の開発状況と高速の漢字プリンタ

の出現は、従来調査の前作業として処理されてきた大部分の人間作業をコンピュータとの、より緊密な共同作業として進めることが可能となり、本稿で提案するシステムの運用についても、従来から研究者がとってきた分析方法に近い、コンピュータ処理方式が適用できるようになった。

ここで提案するシステムは、入力装置のOMR機能を利用したターンアラウンド処理であるが、単位切り部分と付加情報部分の二つに分けられる。図3はシステム全体を示したものであり、図4は入出力装置とプリンタ及びホスト・コンピュータの相互関係を示したものである。

また、これらのシステムを従来のシステム構造と比較した場合、次のような表で表すことができる。

表1 処理構造

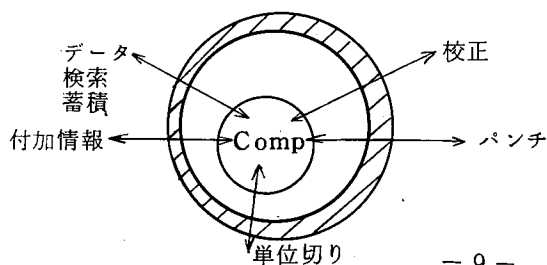
Post-edit Type		Pre-edit Type	
人	パンチ 校正	人	単位切り 校正
Comp	Computer	人	清書
人	単位切り 校正	人	付加情報 校正
Comp	Computer	人	清書
人	付加情報 校正	人	パンチ 校正
Comp	Computer	Comp	Computer

すなわち、処理のパターンにもとづいて、従来の方式と、ここに言うターンアラウンド方式とを比較すると、前者では、各々の機械処理の前に、人手作業が挿入されるが、後者では、人手作業を一括して進めた上で機械処理に入るといふ違いがある。この点から、従来の方式は、“Pre-edit type”と呼ぶことができ、ターンアラウンド方式は“Post-edit type”と名づけることができよう。

ここでは後者の型を中心に解説する。この型の特長は、システム階層が

人間処理←人間処理+コンピュータ処理

反復処理であるため、コンピュータの外側の環を人間—機械系のインターフェースとして見る事が可能となり、これらの環を入出力に対する各装置と媒体間の操作に置き換えることができる。これらの関係は、システムに追加される新規の作業に対しても、インターフェースとなる媒体形式が、その目的に合致するかどうかの確認問題のみとなる。例えばカードによる付加情報処理は、そのまま同語異語処理用カードとして併用可能であり、従来のシステム方式に比べ、いわゆる、一般付加情報処理と同一時点で処理が可能となる。その他、装置の増設の場合も Open-ended な形となる。



これによって、システム全体は、従来の方式に比較し、少なくとも、コンピュータ専門家以外にも、いわゆるシステムのブラック・ボックスは少なくなる。これは、入力媒体、出力媒体、作業長表が同一媒体で統一されており、この点では、コンピュータとのインターフェースは作業者自身の目で確実に視覚化された形となるためである。

これらシステムの利点は次の通りである。

- (1) 第一次入力データを減少させることが可能。
- (2) システムのモジュール化が可能。
- (3) 外部データとの接続が容易。
- (4) 中間段階でのデータ再入力のみスマッチング防止可能。
- (5) データの再パンチ量を減少させることが可能。
- (6) データの現場での作成が可能。
- (7) 同語異語判別処理が通常の付加情報処理と同一時点で処理可能。
- (8) データの転記作業が省略可能。
- (9) 用例カードによる情報検索用カードとして使用可能。

4. 結論

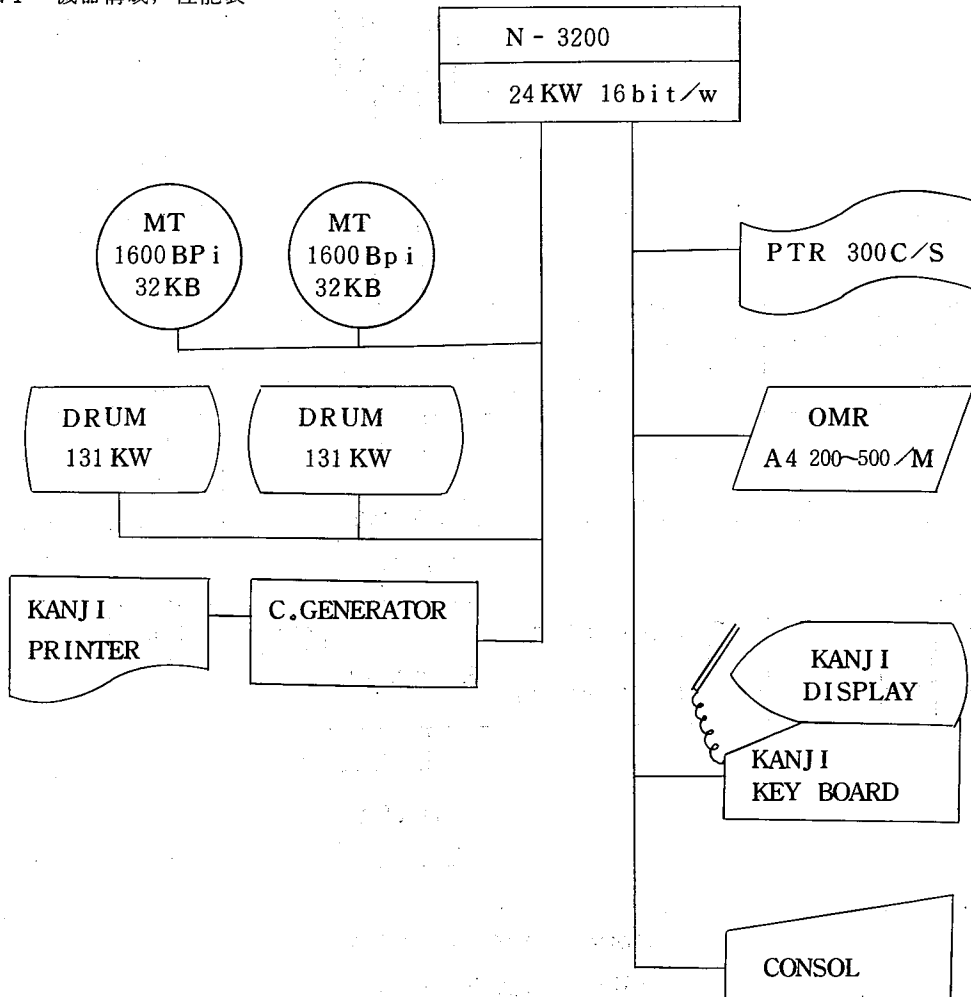
カード及びシートによる調査システムの概要について述べてきたが、本システムによる問題点の一つは、漢字の読みの処理について、まだ未解決な点があることである。

これは、同語異語判別処理の中で活用形処理、同音処理また語配列の作表等に影響する。システムとしては、資料から得られる範囲での処理と、それらを鍵として、さらに発展した形での自動化処理の二部分に分けてあるが、今後漢字カナ変換処理の自動化を進める必要がある。

最後に、このシステムの開発に当たっては、ファコム・ハイタック中島保行氏、大日本印刷 坂巻照夫氏、日本電気官庁システム部の方々に大変お世話になった。深く感謝の意を表する。

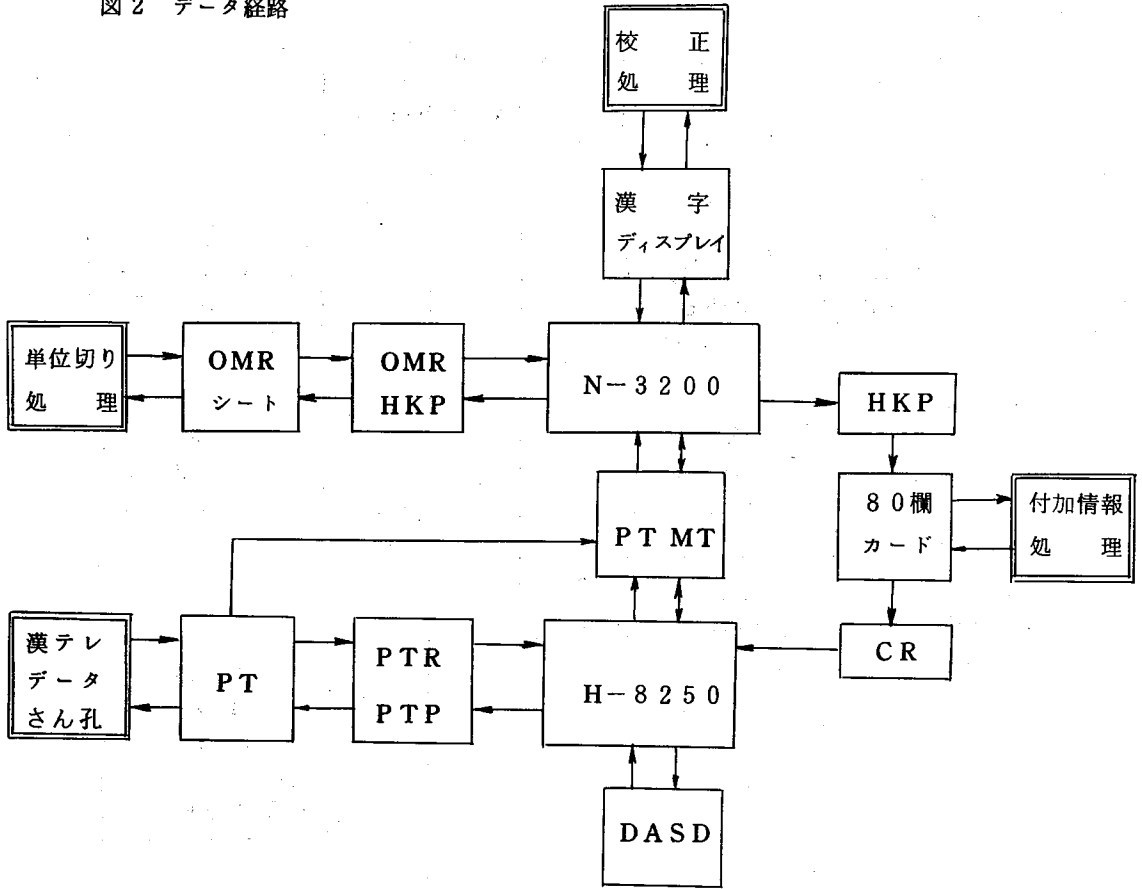
- (1) 斎藤 秀紀「漢字プリンタを使用したターンアラウンドシステム」国立国語研究所報告51,
(1974)
- (2) 斎藤 秀紀「漢字プリンタを使用したターンアラウンド システムⅡ」国立国語研究所報告54, (1974)
- (3) 斎藤 秀紀「国立国語研究所における高速漢字プリンタシステムの概要」ドキュメンテーション研究・UoI.25, No.8, (1975)
- (4) 「C-5210 高速漢字プリンタ取扱説明書」日本電気KK (1975)

図1 機器構成、性能表



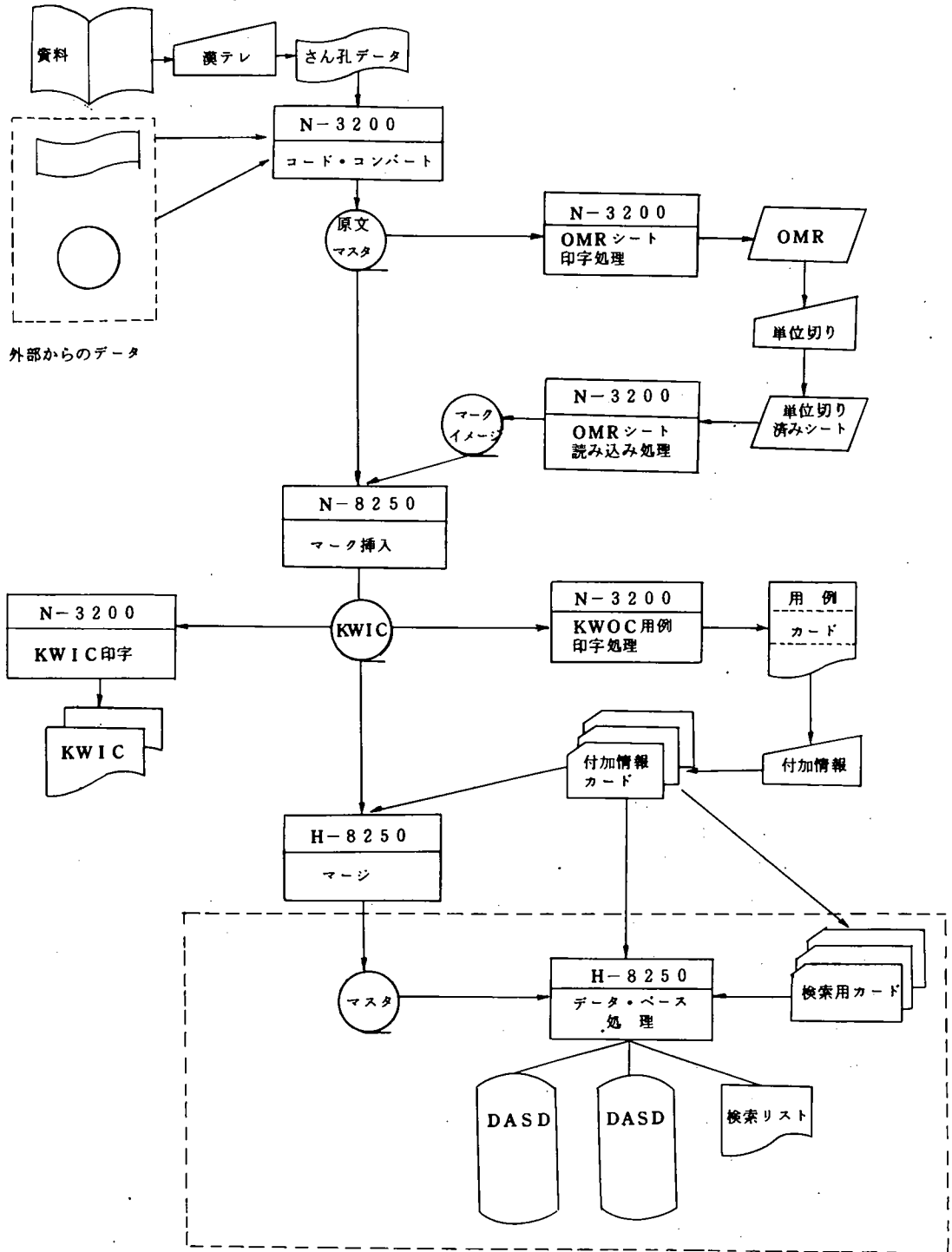
収容文字 (基本+外字)	4977+3584	記録方式	多針電極 湿式静電記録
文字発生方式	ドット(24×24)	印字速度	6行/インチ 1800行/分 3行/インチ 1200行/分
文字の大きさ	9,12 P	用紙巾	200~457 mm
文字記憶素子	ワイヤメモリ	用紙厚	55 kg~135 kg
印字行数	80字×48行 (9,12共)	用紙送り	スプロケット
文字間隔	9 P 7/1インチ 12 P 5/1インチ	用紙	静電記録紙
行間隔	3/6行インチ 4/8行インチ	コピー枚数	同一ページ繰返印字
字 体	ゴチック タイポス		

図2 データ経路



- OMR 光学記号読み取り装置
- HKP 高速漢字プリンタ
- CR カードリーダー
- PTR 紙テープ読み取り装置
- PTP 紙テープさん孔装置
- DASD 磁気ディスク装置
- MT 磁気テープ装置

図3 ゼネラルフローチャート



AB0039 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0

案内に立つ。三人共小倉袴に紺足袋で、朴歯の下駄をがらつかせて出る。上野
 の山から根岸を抜けて、通新町を右へ折れる。お歯黒溝の側を大門に廻る。吉
 原_↑を縦横に溜歩する。軟派の生徒で出くはした奴は災難だ。白足袋がこそ
 そと横町に曲るのを見送つて、三人一度にどつと笑ふのである。僕は分れて、
 今戸の渡を向島へ渡つた。

図4 OMRシート印字例

	INDEX NO.									情報I	情報II	情報III	情報IV
さまざま	0	1	2	3	4	5	6	7	8	0	0	0	0
***の形をした植物や動物 が、基本的にはすべて同じ構 造の細胞からなり、その本質 的な化学成分もほぼ等しいこ とから、細胞が生物のからだ を構成する基本単位ではない かという予想がうまれる。	2	2	2	2	2	2	2	2	2	2	2	2	2
	3	3	3	3	3	3	3	3	3	3	3	3	3
	4	4	4	4	4	4	4	4	4	4	4	4	4
	5	5	5	5	5	5	5	5	5	5	5	5	5
	6	6	6	6	6	6	6	6	6	6	6	6	6
	7	7	7	7	7	7	7	7	7	7	7	7	7
	8	8	8	8	8	8	8	8	8	8	8	8	8
	9	9	9	9	9	9	9	9	9	9	9	9	9

付加情報処理カード

	INDEX NO.									情報I	情報II	情報III	情報IV
の	0	0	0	0	0	0	0	0	0	0	0	0	0
さまざま***形をした植物 や動物が、基本的にはすべて 同じ構造の細胞からなり、そ の本質的な化学成分もほぼ等 しいことから、細胞が生物の からだを構成する基本単位で はないかという予想がうまれ る。	1	1	1	1	1	1	1	1	1	1	1	1	1
	2	2	2	2	2	2	2	2	2	2	2	2	2
	3	3	3	3	3	3	3	3	3	3	3	3	3
	4	4	4	4	4	4	4	4	4	4	4	4	4
	5	5	5	5	5	5	5	5	5	5	5	5	5
	6	6	6	6	6	6	6	6	6	6	6	6	6
	7	7	7	7	7	7	7	7	7	7	7	7	7
	8	8	8	8	8	8	8	8	8	8	8	8	8
	9	9	9	9	9	9	9	9	9	9	9	9	9

付加情報処理カード

図5 80欄カード用例

小説△は／たくさん／読む／。／新聞△や／雑誌△を／見る／とき△は／、／議論△なんぞ△は／
 見△な／い△て／、／小
 説△を／読む／。／しかじ／もし／なん△と／思っ△て／読む△か△と／いう／こと△を／作者△が
 ／知っ△たら／、／作者△
 は／憤慨△するだ△ろ△う△。／芸術△品△と／し△て／見る△の△て△は／ない／。／金井△君△は
 ／芸術△品△に△は／非常△に／
 高い／要求△を／し△て／いる△から／、／そこいら△じゅう△に／ある／小説△は／この／要求△
 を／充たす△に／足
 り△ない／。／金井△君△に△は／、／作者△が／どう／いう／心理△的△状態△で△書い△て／い
 る△か△と／いう／こと△が／
 おもしろい△の△て△ある／。／それ△だ△から／金井△君△の／ため△に△は／、／作者△が／悲