

TR-039

新聞における校正・校閲の
実データによる調査

石 井 暁

1983. 12

©1984, ICOT

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03) 456-3191 ~ 5
Telex ICOT J32964

Institute for New Generation Computer Technology

新聞における校正・校閲の実データによる調査

石井 暁

要約

文章作成における校正作業の計算機による代行の可能性を調べ、また、もし可能なら、どのような研究を行う必要があるかを調べるため、校正作業の調査を行った。

対象としたのは朝日新聞社の実データ約4000行であり、制作局における校正と校閲部における校閲を含んでいる。このデータ量は朝刊（24ページ）の記事の大体3割に過ぎないが、それなりの価値はあると考えている。

計算機化の困難度により修正を分類し、易しいと思われる順にそれに属する修正の割合と共に列挙すると次の様になる。

現状の技術、用語集等の知識により大体可	13%
用語集の質・量的拡充により可	19%
固有名詞データベースとの結合により可	12%
対象文の十分な理解により可	45%
記事全体の理解により可	0%
不可（常識が必要）	11%

現在の技術でメドがたつのは、初めの3点および4番目（文の理解）の簡単な部分であろう。

目次

1. はじめに	1
2. 調査方法	2
2. 1. 新聞社における校正の流れ	2
2. 2. 使用した資料	3
2. 3. 修正の数え方	4
2. 4. 修正の種類による分類	4
2. 5. 修正の困難度による分類	10
3. 調査結果	12
3. 1. 修正の種類による分布	12
3. 2. 修正の困難度による分布	12
4. おわりに	18
文献	19

1. はじめに

文章の作成には校正作業^{*}が伴う事が多い。この作業は、書類（あるいは本、論文など）の構想を練る様な作業に比較すると、著者にとっては面倒な、避けたい作業である。また、完全に行うのは至難である。一方、専門の校正者による校正は、時間的または経済的理由から充分には求められない事も多い。そこで校正作業を計算機に代行させる事ができれば望ましい。^[5]

更に、積極的に校正作業は計算機に向いている面もある。即ち、誤字の検出においては、文字が正しいかを機械的に確実に見ていかなければならない。特に、例えば、人の名前が何文字か一体となって目に入ると、人間はその中の1文字が誤っていてもなかなか解らないので、1文字1文字に神経を集中しなければならない。^[6, p. 19]計算機によれば、この種の確実性は容易に実現できる。

勿論、高度な判断を必要とする校正もあるので、どの程度の校正が計算機によって代行可能かを実データによって調べるのが、本調査の目的である。

実データとしては新聞の記事を用いた。これは、新聞の文章は、難しさが大体一定であり、また漢字の使い方等について規準が比較的是っきりしている事による。^{[8][9][12]}また新聞社側に、校正作業の計算機化の動きがあり、^{[13][14, pp. 144-158]}参考になる事にもよる。

事務書類ならまだしも、小説や随筆においては筆者の個性が強く出るため、漢字の使い方等に必ずしも統一的な規準なく、^{[1][6][10]}更には、一つの作品の中でも複数の表記が存在し得る。^{[1, p. 72][6, pp. 61-62]}そのため、校正の計算機化は当面メリットが少いであろう。

本文において、通常の英文の綴りの誤り検出^[2]と同様に、計算機は文字を表すコード情報によって表わされた文章のみを用いて校正作業を行うものとする。即ち、文字認識技術において複数の候補文字の中から正しい文字を選んだり、必要に応じて文字パターンを参照する場合とは状況が異なる。また、計算機による誤りや正しい表記の提示のし方が、ディスプレイターミナルを用いた対話型であるか、また印刷された紙によるバッチ型であるかは、特に本調査に関係ない。

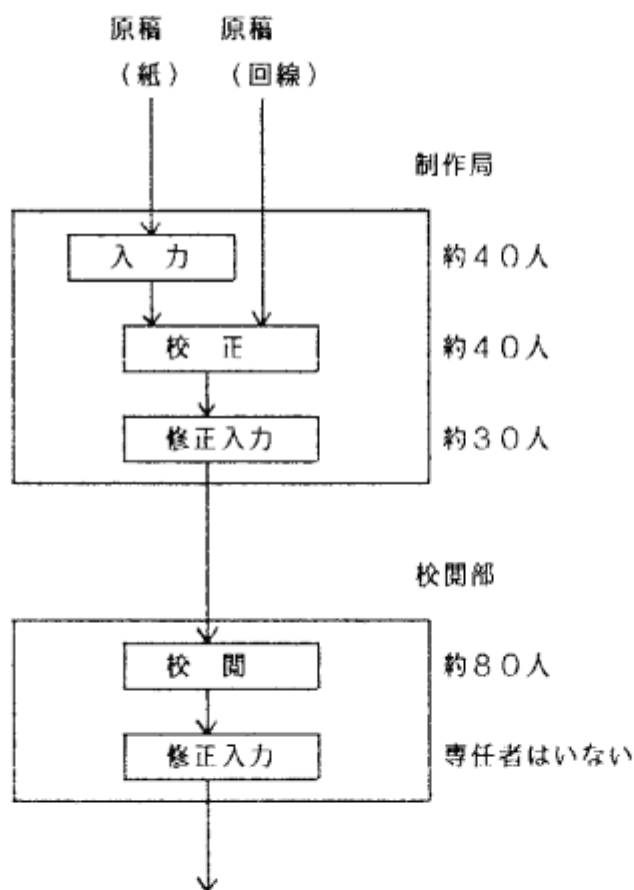
* ここでは文字（言葉）についての校正のみ考える。

2. 調査方法

2. 1. 新聞社における校正の流れ*

データを提供して頂いた朝日新聞社東京本社の校正の流れは、第1図に示す様になっている。

校正の仕事は、制作局と校閲部が行っている。制作局では紙の形の原稿は入力した上、また支局や通信社から回線経由で送られてきた原稿はそのまま、一度紙に印刷し、校正作業をする。ここでの校正は、原則として原稿通りの入力を得るために行うが、用語の誤り等、^{[8][9][12]}明らかな誤りも訂正する。また、入力用の端末は漢字の字数を少なくしてあ



第1図 朝日新聞社の校正の流れ

* 2章以降、原則として“校正”と“校閲”を朝日新聞社の体制に従って使い分ける。

る（約2000字）ため、ゲタ*が生じる事が有り、その修正も行う。

一方、校閲部の校閲は、制作局の校正漏れを調べる事もするが、次の3点が主な仕事である。第1に事実との相違が無いが、職員録等の資料により調べる。これは時間が掛り、完全に調べる事は困難である。また調べても誤りの発見は少ない。しかし、記事の信頼性を高めるために、重要な仕事である。

第2に記事中の矛盾の発見である。紙面の1つの記事は、多くの細切れの記事及び別のルートで作られる標題や表から成り、かつ新しい情報が得られたためや、紙面の編集のため、大きな加除が行われる事も多い。そのため1つの記事内で食い違いや、不自然な事が無いかを調べる。

第3に言葉の問題であり、より良い文にするための修正を行う。この際、修正の規準が辞書等の資料に有る事も、無い事もある。また、それが何面に載るかに依っても規準は異なる。

第1図の右側の数字は、それに従事している人の数である。支局での入力作業専任者は合計約20名である。

この他の校正作業としては入力オペレータによる修正や出稿部（記事を書いた部）、整理部（記事をどの面に載せるかを決める部）による物がある。これらの修正は調査の対象にしていない。しかし後に述べる修正の例の中にはこれらの校正による物を含んでいる。

2. 2. 使用した資料

使用した資料は次の3種である。

①：制作局で校正に使用したゲラ刷り 121枚、 1,639行であり、制作局から入手した物である。制作局の修正が88箇所入っている。

内容は地方面であり、選挙運動関係の記事が大部分で、教員異動の記事がそれに次ぐ。

②：校閲部で校閲に使用したゲラ刷り79枚、 1,109行であり、校閲部から入手した物である。校閲部の修正が24箇所入っている。このうち2箇所は復旧されている（イキになっている）が、それも数えた。

* 入力用端末で入力できない字が原稿に有った時、その字の代りに■の記号を仮に入力しておく。この記号のこと。活字を使った印刷においては、活字を拾った後、それを句読点、改行等を整えながらならべる。この際、何らかの理由で、必要な活字が拾われていない場合、あり合わせの活字を上下さかさにしてならべ、スペースを確保しておく。すると活字の底面にあたるところの真中には溝がほってあるので、下駄の歯の様に■または■■と印刷される。これを伏字またはゲタと呼ぶ。[6, pp.23-24] JIS (C 6226-1978, 情報交換用漢字符号系の附属書表1, 6) でも■記号の名称は“げた記号”となっている。

内容は社会面である。

③：校閲部で校閲に使用したゲラ刷りで、一面の記事である。次の様な物が混ざっている。

③-a：56枚、829行に校閲部の修正が8箇所入った物である。

③-b：21枚、345行に制作局の修正が39箇所入った物である。

いずれの資料においても同じ文章（電文）のコピーが複数ある場合は、修正の少ない方を取り除いた。また、同じ事件を扱い、同内容の文章であっても、異なる文章であれば、調査の対象にした。上記の数字はその様にして得た物である。

以上により、制作局の校正は①および③-bの資料合計1984行、127箇所の修正（約16行に1箇所の修正）となる。^{*} 一方校閲部の校閲は②及び③-aの資料合計1938行、32箇所の修正（約61行に1箇所の修正）となる。

尚、校正は1984行、校閲は1938行を調べたので、量に2%程度の差がある。後の議論で、この差の影響を無くすため、正規化する必要があるが、本文ではこの差を重視していない。

使用した資料は合計3922行となり、かなり少く、朝刊（24ページ）の記事の量^[15]の32%程度である事に注意して見て頂きたい。

2. 3. 修正の数え方

修正は、文字数等では無く、修正箇所により数えた。1箇所の修正が1文字の修正である事が多いが、単語や文節単位の修正もあった。

文字数により修正を数えなかった理由は、計算機により誤り修正ができる様になった場合、数文字からなる1箇所の修正が一度に行なわれるだろうと考えたためである。しかし、現状ではそれを完全に行うことは不可能であるため、計算機化の効果を、修正入力作業がどの程度削減されたかで判断する事になると思われる。そのため、修正入力を行う際のストローク数を評価する必要がある。しかし、これは装置やソフトウェアに依存するため、ここでは行わなかった。

2. 4. 修正の種類による分類

修正は次の様に6種類に大分類し、各々を更に分類した。各分類の説明をする。

（1）文の変更

修正前でも一応（一見）正しい文であるが、修正を加えたものである。

^{*} 朝日新聞社の以前の調査では、校正は1.8箇所／300字だそうである。これは約12行に1箇所の修正となる。

(1-1) 読み易さ向上(大規模)

文節単位以上での修正である。計算機による代行は、当面難しいと考えられる。

(1-2) 読み易さ向上(小規模)

(1-3) 同上 (助詞のみ)

(1-4) 同上 (、のみ)

小規模(大体文節の範囲内)の修正である。助詞のみの修正と読点のみの修正は別に数えた。

(1-2)の例としては次の様な物がある。

住宅一戸を全焼、 → 住宅一戸を全焼し、

クジ運はツイてない → クジ運はない

〇〇市長 → 〇〇××市長

(〇〇は人の姓、××は市の名)

この中で → その中で

(1-3)の例としては次の様な物がある。

先輩の四候補 → 先輩四候補

定数七のところ九人が立候補 → 定数七のところに九人が立候補

内容としてほしい → 内容にしてほしい

文全体の解析及び単語の使い方、つながり方の詳細な調査により計算機による代行も可能と思われる。ただし、以上の修正については、なぜそう修正した方が良いのか、筆者には解らない物が多い。専門家の“知識”を充分知る必要がある。

(1-5) 事実に反する物の修正

(1-6) 同上 (固有名詞、データベースに有り)

(1-7) 同上 (固有名詞、データベースに無し)

いずれも事実に反する物の修正である。その中で(1-5)は

市役所の駐車場 → 市役所前の駐車場

の様な物、隣に行へすり替った物等の外、法令の内容などを知らないと全く解らない物等がある。計算機化は困難である。

固有名詞の誤りについてはデータベース(職員録や人名録など)に記載が有る物が、無い物かを区別した。データベースにどの程度まで記載されているかはまだ調査が終わっていないので、一応次の様な組織、その長、その長に次ぐ者がデータベースに有りとした。

政府機関、地方自治体（その部課などの組織を含む）

学校

裁判所（全裁判官名を含む）

データベースに有る物については、文章の解析及び計算機に格納されたデータベースにより計算機化が可能であろう。しかしデータベースに無い物については計算機化は困難であろう。

（２）表記の修正

言葉の表記の修正などである。

（２－１）漢字、かなの使い分け（用語集に有り）

（２－２） 同上 （ ” 無し）

（２－３）送りかな （ ” 有り）

（２－４） 同上 （ ” 無し）

（２－５）かな使い （ ” 有り）

（２－６） 同上 （ ” 無し）

（２－７）固有名詞表記 （ ” 有り）

（２－８） 同上 （ ” 無し）

かな使いは、“ず”と“づ”の使い分けであり、これのみ、用語集に無い物については辞書^[7]によった。固有名詞の表記は、外国の国名などのカタカナによる表記や会議名の日本語訳である。その他は、特に説明の必要はない。

本調査すべてについて用語集としては朝日新聞社の物^[8]を用いた。新聞協会の物^[12]によると、外来語のカタカナ表記は、すべて用語集に無しとなる他、送りかなについて１つ用語集に無しとなる。また共同通信社の物^[9]によると、次に述べる約束ごとについて、用語集に有無が変わる。即ち、選挙関係の語（候補者など）は用語集に無い。漢数字の単位語（一〇〇〇二九等）について記述が少なく、下の例は明示されていない。一方、敬称のつけ方の例は用語集に有りとなる。その他は、次に（３）で述べる漢字の誤りについてを含めて、どれに依っても、同じ結果となる。

尚、用語集と異なる（例外扱い）表記をとっている物は、他の種類の修正についても、用語集には無しの扱いとした。また漢字、かなの使い分けには用語集の原則等の記述を基に考えれば判断できる物もあったが、直接記述されていなければ用語集に無しの扱いとした。

（２－９）約束ごと（用語集に有り）

（２－１０）同上 （用語集に無し）

例えば

立候補七人の → 候補者七人の （用語集に有り）

一ツル 二十九ルド → 一ツル 二九ルド (")
 ナップザック → 小型リュック (用語集に無し)
 ○○党新人 → ○○党新顔 (")
 (○○は政党名)
 ○○ → ○○さん (")
 (○○は女優の名、社会面)

である。

以上の(2-1)～(2-10)の訂正は、用語集に有る物は現在の技術で計算機化が人体可能であり、用語集に無い物についても用語の収集を続ける事により可能であろう。

尚、以上で用語集に無しとした物でも、各新聞社内には定まったルールが存在すると思われる。しかし、公開されていない様なので、用語集に無しとした。他の種類の修正でも、同様な事情があり得る。

(2-11) 記号の表記(用語集に有り)

カッコの使い方は用語集に有るが、それに従っての修正は無かった。

(2-12) 記号の表記(用語集に無し)

意味の無い空白行の削除

開きカッコが有るが、対応する閉じカッコが無い事の修正

びの前に数字が無い事の修正

勢ぞろい 写真。 → 勢ぞろい 写真。

等である。

大体は常識的なルールに基づいた修正であり、ルールをはっきりさせる事により計算機化が可能と思われる。尚、一や一が誤った場所に入った例を一応ここに入れたが、これは、もっと困難である。

(3) 漢字(熟語)の誤り

(3-1) 単語分けできない

例えば

全車国協議会 → 全国協議会

市場の立動き → 市場の動き

の様に単語分け途中に“変だ”と計算機が解る様なエラーである。次の“辞書に無い”エラーとの区別は必ずしも明確ではない。しかし、いろいろな単語分けの可能性に基づいて

辞書を調べてそれに無い事を知り、次に新語や固有名詞であるか否かの判定を行うことは効率の悪い処理であり、辞書を調べる前にエラーだと解る“べき”であると考えて別の分類とした。文字の接続情報のチェック^{[3][4]}など、異なった手法を併用することが特に有効であろうと考えたのも別の分類とした理由の1つである。

(3-2) 辞書・用語集だけで解る

その語が用語集に誤りや言い換えた方が良い語として載っている場合は問題が無い。

(その様な例は無かった。) 辞書^[7]に載っていない語の例としては、

共調 → 協調

政府交官 → 政府高官

冒陳 → 冒頭陳述

等がある。この種の修正は、この様に同音または字体が似ている等の理由により、解析できる物が多いと考えられる。

(3-3) その文の解析で解る

例えば

〇〇宮司らからお払いを受けた → 〇〇宮司らからおはらいを受けた

(〇〇は人の名)

顔を合わせた → 顔を会わせた

の様に漢字の誤りの修正の他、

農林地帯 → 農村地帯

の様に一応正しい言葉の修正や

出陣式の各事務所を予定 → 出陣式を予定

という例がある。最後の例はここに示した文面だけでは誤りとは解らない難しい例である。

(別の分類に入れるべきかも知れないが、一応ここに入れておいた。)

(3-4) その文の解析と用語集で解る。

同音意義語の使い分けの問題などである。上記(3-3)と同様な物もあるが、用語集にその言葉の使い方が載っている物はここに入れた。通常(3-3)以上に前後の語の理解が必要となろう。例をいくつか挙げる。

強行 → 強硬
合わせ検討 → 併せ検討
競争油種 → 競合油種
鎮静化 → 沈静化

(3-5) 他の文も解析すると解る

前後の文により記事の主題を理解して初めて誤りと解るものであり、

ミサイルの完全配備に至る → ミサイルの完全廃棄に至る

の様な修正がある。

これの計算機化には前後の文を併せた理解が必要で、現在、ほとんど研究が進んでいない。また、このエラーと(1-5)“事実と反する物の修正”は区別が難しい場合もある。

(4) かなの誤り

漢字の誤りとほぼ同様であり、

(4-1) 単語分けできない

(4-2) その文の解析で解る

(4-3) 他の文も解析すると解る

に分類した。

この様に分類したのは、前の漢字の誤りの分類と同様の理由からである。ただしかなの誤りの修正に、辞書は余り役に立たない。例えば助詞を別の助詞と誤った場合、辞書だけでは全く役に立たない。そのため漢字の誤りの項と比較すると“辞書により解る”等の分類は無い。

例を挙げる。(4-1)には

発症ほ可能性 → 発症の可能性

「××協議会」は(○○会長)は → 「××協議会」(○○会長)は^{*}
(○○は人の名、××は会の名)

また(4-2)には

政党別によると → 政党別にみると

それこれ → それぞれ

* 原資料は縦書きなので、カッコ類は-、一、へ、ゝとなっている。

の様に自立語の例もあるが、多くは次の様に助詞の誤りであって、文節だけでは誤りとは解らない物である。

常習とばくの二人を → 常習とばく、客二人を
擁立の難しく → 擁立も難しく
販売と監督する → 販売を監督する

最後の（４－３）には

この終業式で → この日終業式で

がある。この修正の様に漢字の誤りか、かなの誤りか分類が困難な物も多い。

（５）ゲタ

次の様に分類した。

（５－１）人名、データベースに有り

（５－２）人名、データベースに無し

（５－３）地名、組織名

（５－４）その他の漢字

（５－５）その他

（５－６）作字

データベースの範囲は（１－６）で述べた物である。地名、組織名はデータベースに載っていると考えた。また、その他はㄱ、ㄴ、CC等である。

２．５．修正の困難度による分類

前項にも一部述べた様、修正の分類ごとに計算機化の困難度を考察した。それをつぎの６通りに分類した。

A：現状の単語分け技術及び用語集・辞書の知識に基づき、少しの改良で実現し得る。

B：用語集・辞書の知識の拡張及び記号の使い方に類する知識の整備により実現し得る。

C：固有名詞のデータベース整備及びそれに対する検索文の作成機能により実現し得る。

D：対象とする文全体について単語のつながり方を解析し、文全体の意味を理解する事により実現し得る。

E：記事全体の各文の意味の理解に基づき、文と文の関係を理解する事により実現し得る。

る。

F：幅広い常識を必要とする修正で、現状では一応計算機化が無理と考えられる。

大体、この順に困難さが増しており、Dの難しい方（変えなくても意味は変わらないが、助詞を変える物等）以下は現状ではまだかなり困難な問題である。

尚、計算機化に当たっては修正すべき点の指摘のみを行わせるのか、あるいはどう修正すべきか（の案）の提示まで行わせるのかを区別すべきであるが、本文では特に区別を行っていない。

3. 調査結果

各資料の修正の数を分類して数えた物を表1に示す。以下、これについて考察する。

3. 1. 修正の種類による分布

修正をすでに述べた種類によって分類したグラフを第2図に示す。上が制作局の校正、下が校閲部の校閲に対応する。^{*} 棒の高さが修正の数を表わす。棒の上下の数は校正、校閲各々の中での割合を示す。棒の中は種類を表わす語と校正、校閲を合わせて数えた割合である。

校正ではゲタが39%で最も多く、その他については文、表記、漢字、かなの順で修正が多い。尚、校正の調査の対象とした資料のうち97行(約6%)が教員異動の記事であり、その中で8箇所のゲタ(16%)が発生していた。これは無視できない偏りである。しかし、この程度では第2図の傾向に変化が無い。また人事の記事は政治面などに多く、特に問題にはならないと考えられる。

一方校閲については、校正より修正が少ない。当然、作字を除いてゲタが無くなっている他、文の変更、漢字の誤りおよびかなの誤りが大きく減少している。その中で表記については、それ程減少していない事が解る。

3. 2. 修正の困難度による分布

第3図に、困難度A～Fによって分類した分布を示す。図の書き方は前の第2図と同様である。

第1に、困難度D(その文の理解により計算機化が可)の物が、最も多い。文章の理解の研究により、大きな成果が期待できる。ただし、この中には、特に意味が変わらない助詞の使い分け、読点の使い方、同音異義語の使い分け、一般語のゲタ等、かなり研究を要するものが多い。

困難度E(記事全体の理解により計算機化が可)のものは1%以下で、少なく、記事全体の理解の研究は、現状ではそれ程重視しなくて良い事が解る。ただし上に述べた困難度Dの物の中には、前後の文を理解する必要のある(または、理解した方が簡単に解る)修正も含まれている可能性はある。

困難度F(現状では計算機化が不可)の物も1割余り存在する。文章としては全く正しいが事実と反する物の修正やデータベースに無い固有名詞の修正等、多くの修正が人間の介入を欠かせないと思われる。

^{*} 各々調査の対象とした資料は異り、量も約2%の相違がある。

一方、計算機化が容易と思われるものについて検討する。困難度 A（現状の技術及び用語集・辞書の知識を基に計算機が大体可）の物及び困難度 C（職員録等に記載された固有名詞の修正であり、データベースとの結合により計算機化が可）については校正の段階でほとんど全部が修正されている。

また、困難度 B（現在の用語集には無いが、その改定・整備として考えられる情報によって可）については、校正・校閲共に修正しているが、校閲によるものが倍近く多い。

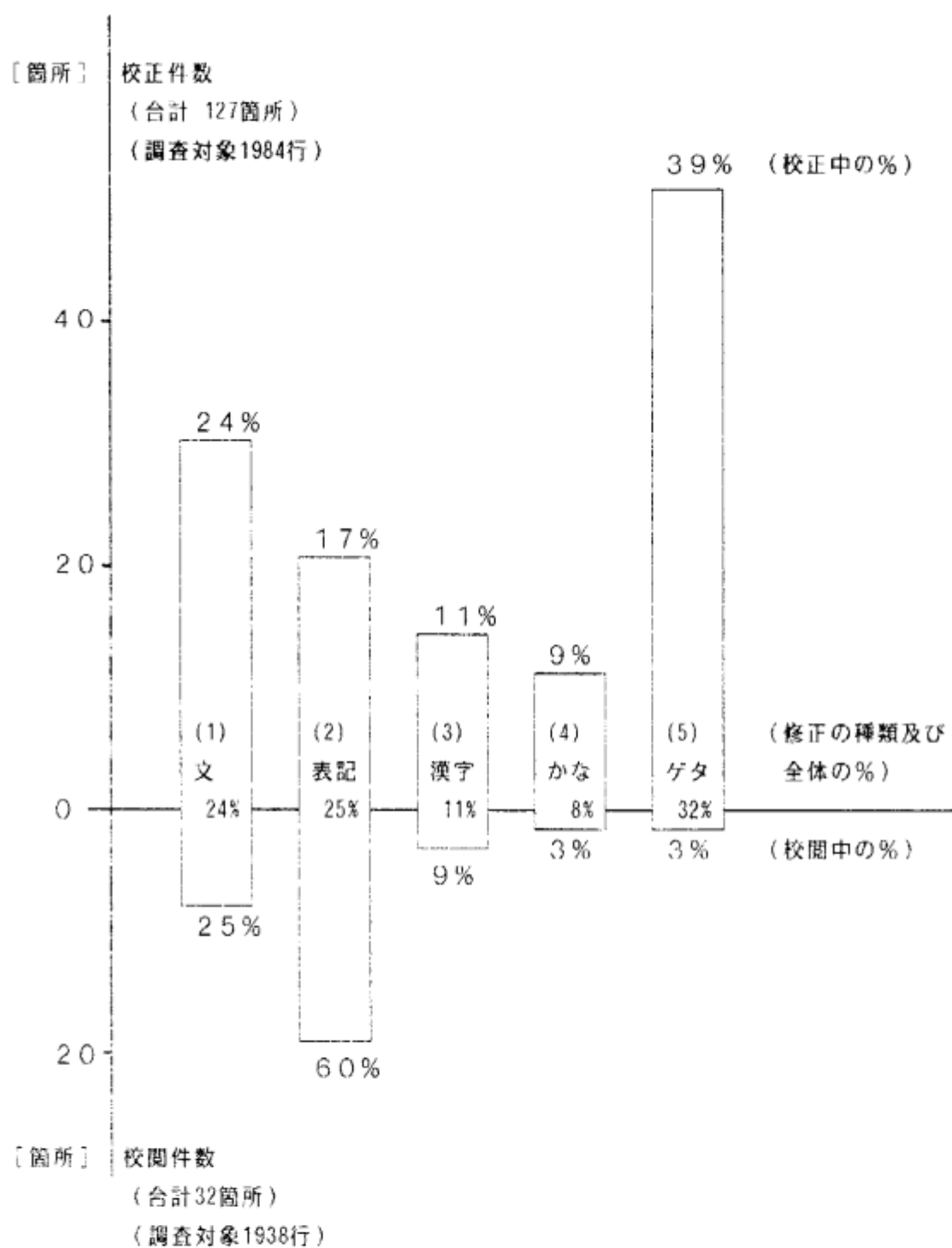
困難度 A～C の計算機化が比較的容易な物の合計は、44%である。

表1. 修正の分布（その1）

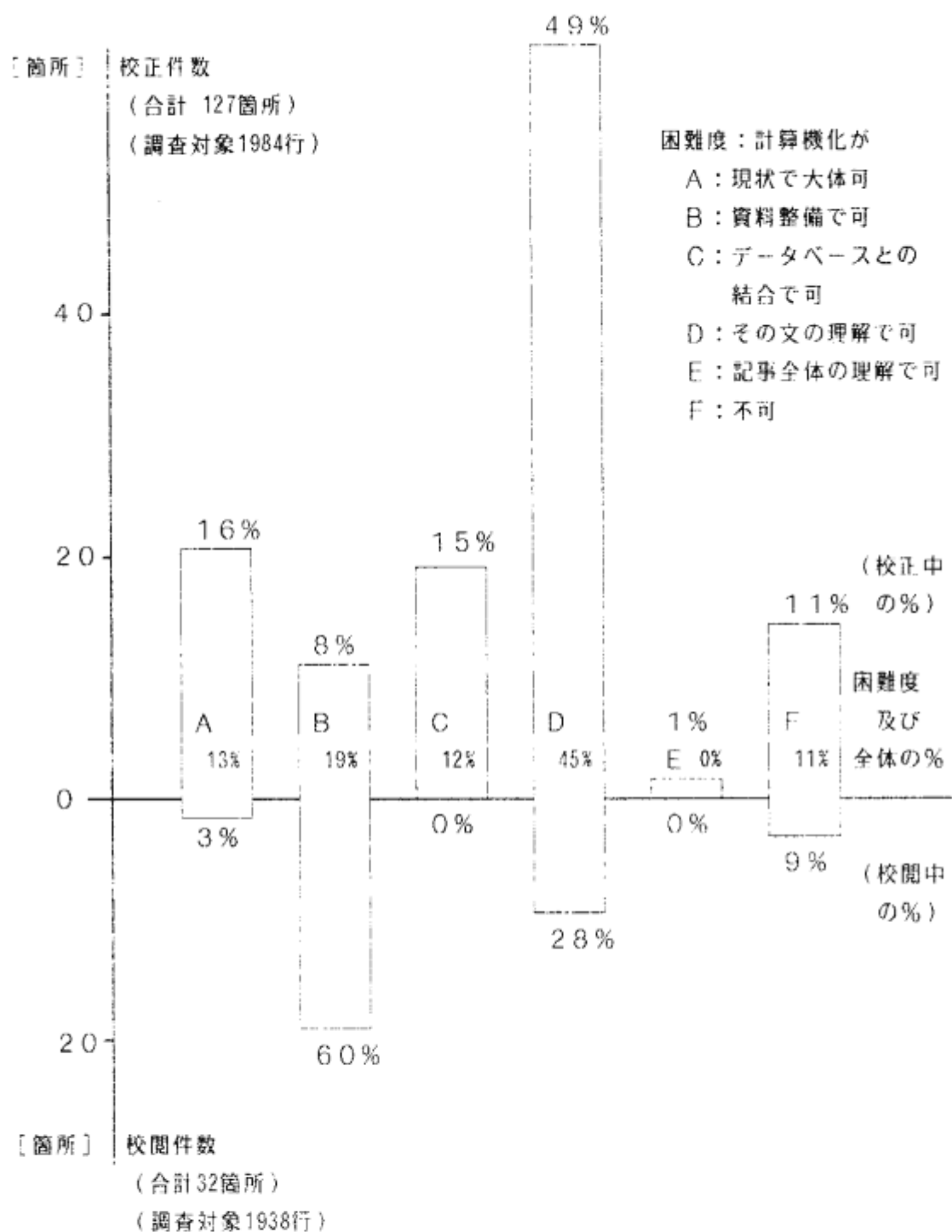
資料 修正	制作局の校正		校閲部の校閲		困 難 度
	① 1,639行	③-b 345行	③-a 829行	② 1,109行	
1. 文の変更					
1-1. 読み易さ向上（大規模）		2			F
1-2. " （小規模）	6	1	1	2	D
1-3. " （助詞のみ）	6	1	1	1	D
1-4. " （、のみ）	1	1		1	D
1-5. 事実に反する物の修正	4			1	F
1-6. " （固有名詞、 データベースに有り）	6				C
1-7. " （固有名詞、 データベースに無し）	2			1	F
小計	25	5	2	6	計38
2. 表記の修正					
2-1. 漢字、かな（用語集に有り）	2				A
2-2. " 無し				2	B
2-3. 送りがな 有り	2	1			A
2-4. " 無し					
2-5. かな使い 有り		1（辞書による）			A
2-6. " 無し					
2-7. 固有名詞表記 有り		1			A
2-8. " 無し		2	4	1	B
2-9. 約束こと 有り	2	1			A
2-10. " 無し	2	1		5	B
2-11. 記号の表記 有り					
2-12. " 無し	5	1	1	6	B
小計	13	8	5	14	計40

表 1. 修正の分布 (その 2)

資料 修正	制作局の校正		校閲部の校閲		困 難 度
	① 1,639行	③-b 345行	③-a 829行	② 1,109行	
3. 漢字 (熟語) の誤り					
3-1. 単語分けできない		2		1	A
3-2. 辞書・用語集だけで解る		3			A
3-3. その文の解析で解る	5	2		1	D
3-4. その文の解析と用語集で解る		2	1		D
3-5. 他の文も解析すると解る					
小 計	5	9	1	2	計 17
4. かなの誤り					
4-1. 単語分けできない	2	3			A
4-2. その文の解析で解る	5	1		1	D
4-3. 他の文も解析すると解る	1				F
小 計	8	4		1	計 13
5. ゲタ					
5-1. 人名、データベースに有り	6				C
5-2. 人名、データベースに無し	6				F
5-3. 地名、組織名	7				C
5-4. その他の漢字	17	3			D
5-5. その他	1	10			D
5-6. 作字				1	F
小計	37	13		1	計 51
合計	88	39	8	24	計 159



第2図 修正の種類による分類



第3図 修正の困難度による分類

4. おわりに

新聞社の校正・校閲の実データに基づいて、校正・校閲においてどのような修正が加えられているか調査した。

修正の種類により分類すると、校正ではゲタの修正が最も多く（約4割）、文の修正、表記の修正、漢字の修正、かなの修正と続く。校閲ではゲタがほとんど無くなり、表記の修正が最も多くなる。

一方、修正の計算機化の可能性を見るため、計算機化の困難度により修正を分類し、易しいと思われる順に、それに属する修正の割合と共に列挙すると次のようになる。

現状の技術、用語集等の知識により大体可	13%
用語集の質・量的拡充により可	19%
固有名詞データベースとの結合により可	12%
対象文の十分な理解により可	45%
記事全体の理解により可	0%
不可（常識が必要）	11%

現在の技術でメドがたつのは、初めの3点および4番目（文の理解により可）の簡単な部分であろう。それによって修正の数の約半分を自動化し、固有名詞については現在以上に確実に確認する事ができよう。

本調査は、少量の実データによって、計算機化の可能性の面から修正の傾向を調査した物である。必要に応じて、出稿部等の校正も併せて、なぜエラーを生じるか^[11, p. 51]等を含めて、多量のデータを用いて更に調査する必要があるだろう。

最後に、校正の流れをお教え頂き、また資料を提供して頂いた関係各位に対して謝意を表す。

文献

- [1] 加藤康司：“辞書の話”再版、中公新書 425、中央公論社(1976)
- [2] 川合 慧：“英文綴り検査法”、情報処理Vol. 24、No. 4、pp.507～ 513、情報処理学会(1983)
- [3] 栗田他：“かな単語入力における誤入力訂正”、自然言語処理技術シンポジウム論文集、pp.125～ 130、情報処理学会(1983)
- [4] 杉村他：“文字接続情報を用いたリジェクト文字の判定処理（文字認識への応用）”、電子通信学会研究会資料PRL81-105、pp.73 ～79、電子通信学会(1982)
- [5] 田中他：“日本語文の入力誤りと自動抽出”、情報処理学会第26回（昭和58年前期）全国大会講演論文集（II）1H-9、pp.1161 ～1162、情報処理学会（1983）
- [6] 長谷川鑑平：“本と校正”7版、中公新書83、中央公論社(1970)
- [7] 久松他：“角川国語辞典”63版、角川書店(1982)
- [8] 朝日新聞社用語幹事編：“朝日新聞の用語の手びき”第10刷、朝日新聞社(1982)
- [9] 共同通信社編：“記者ハンドブック”第4版第3刷、共同通信社(1982)
- [10] 講談社校閲局編：“日本語の正しい表記と用語の辞典”、講談社(1983)
- [11] 日本新聞協会工務委員会編：“新聞印刷・CTS編”、日本新聞協会(1974)
- [12] 日本新聞協会新聞用語懇談会編：“新聞用語集”、日本新聞協会(1981)
- [13] “新聞製作の新時代②”、文化通信1982年8月9日号、p.5、文化通信社(1982)
- [14] “先端的小コンピュータに関する調査研究報告書（適用分野・波及効果編）”、日本機械工業連合会／新世代コンピュータ技術開発機構(1983)
- [15] “朝刊1部の文字の量”、あなたの朝日新聞、p.11、朝日新聞社