

# マニュアルの校閲作業における文書推敲支援ツールの実適用評価

笠原 健成<sup>†</sup> 小林 栄一<sup>†</sup>  
荒井 真人<sup>†</sup> 絹川 博之<sup>††</sup>

計算機ソフトウェア製品では、ソフトウェアのみならずユーザマニュアルも商品としての品質が求められる。高品質なマニュアル商品の開発に際しては、マニュアル原稿を脱稿した段階で、執筆担当のテクニカルライターおよび上級ライターによる校閲作業が必須である。この校閲作業を効率化し、目視による校閲をより高度なものとするために、単純な表現上のミス等の検出作業を軽減することが必要である。この検出作業軽減を目的に、日本語ワードプロセッサ用の文書推敲支援ツールを基に、用字・用語の統一、表現スタイルの統一、文表現指標の算定、分野特有表現への統一、の4つに大別できる処理機能からなる校閲作業支援ツールを開発している。校閲作業支援ツールの開発に続いて、ソフトウェアマニュアルの2,800ページ以上の原稿に実適用して実績評価している。その結果、本ツールを使用せず、全数目視作業する場合と比べて、本ツールを校閲作業に使用する場合は、項目別の不良表現の検出・修正件数について、42%から84%を削減し、所要工数について、執筆担当のテクニカルライターの内容確認作業を80%以上、上級ライターの校閲作業の20%以上を削減する効果があり、本校閲作業支援ツール導入が、当初の目的を実現していることを確認している。

## Performance Evaluation of Applying a Computer-assisted Writing Tool to Proofreading Manuals

TAKENARI KASAHARA,<sup>†</sup> EIICHI KOBAYASHI,<sup>†</sup> MAKOTO ARAI<sup>†</sup>  
and HIROSHI KINUKAWA<sup>††</sup>

Software, when viewed as a total product, requires not only high quality in the software itself, but also in the user manuals delivered with it. Meticulous proofreading is an essential step in developing high quality manuals, first by the technical writer who wrote the manuscript, and then by an editor when writing is finished (or when the manuscript is completed). To make proofreading efficient and effective, it is important to make it easier to detect mistakes in expression. To lighten this proofreading burden, we developed a computer-assisted proofreading tool, based on Japanese word processing software, which features four classes of functions: consistency check of terminology, consistency check of wording, calculation of indices of expression and application of a dictionary defined by writers. We applied this tool to the proofreading of more than 2,800 pages of manuscript and evaluated its performance. As a result, we found that this tool detected from 42% to 84% of mistakes and reduced proofreading time by more than 80% for technical writers and by more than 20% for editors. We confirmed that applying this tool is effective in proofreading manuals.

### 1. はじめに

計算機ソフトウェア製品では、ソフトウェアそのものが使いやすく、所与の目的の遂行に必要な十分な機能と性能を有することのみならず、そのソフトウェアのユーザマニュアルも、使いやすく分かりやすいという商品としての品質の具備が求められる。商品としての

マニュアルは、当該ソフトウェアの使用に際しての内容の適切さとともに、表現的的確さ、分かりやすさが要諦である。これらの要諦を確保するには、テクニカルライターのライティング技術の向上がまず必要であり、種々の作文技術<sup>1)</sup>を基に開発された教材を使用して、ライティング教育が施されている。しかし、商品としての、具体的なマニュアル開発においては、マニュアル原稿を脱稿した段階で、執筆担当のライターおよび上級ライターによる校閲作業が必須である。この校閲作業を効率化し、目視による校閲作業をより高度なものとするには、単純な表現上のミス等の検出作業を軽減することが必須であり、このための校閲作業支援ツールは、用字・用語の統一、表現スタイルの統一、文表現

<sup>†</sup> 株式会社日立製作所ソフトウェア事業部  
Software Division, Hitachi, Ltd.

<sup>††</sup> 株式会社日立製作所システム開発研究所  
Systems Development Laboratory, Hitachi, Ltd.  
現在、東京電機大学工学部情報通信工学科  
Presently with Department of Information and Communication Engineering, Tokyo Denki University

指標の算定、分野特有表現への統一、の4つに大別できる処理機能を有することが必要である。

関連するツールとして、日本語文書の推敲支援ツール<sup>2)</sup>がある。これらの推敲支援ツールは、用字・用語に対する処理機能が十分でない<sup>3)~5)</sup>、常用漢字と仮名の用字の処理機能の具備が明確にされていない<sup>6)</sup>、日本語形態素解析で処理可能な誤りの検出と修正に限定され文表現指標を算定していない<sup>7)</sup>、上記4機能の具備を視野に入れているが、プロトタイプ開発<sup>8)</sup>や、可読性評価が中心<sup>9)</sup>の報告にとどまっており、校閲作業への運用評価は報告されてない。我々はマニュアル開発の校閲作業への適用も視野に入れ、ワードプロセッサ用の推敲支援ツールを開発した<sup>10)</sup>。開発したのは、用字・用語の統一、表現スタイルの統一、文表現指標の算定、分野特有表現への統一、の4機能である。

これらの機能は、

- 人による目視作業を支援できる契機となればよいこと、
- コンパクトで高速に処理できること、
- 出現新語に対する形態素解析用情報の逐次付与が困難であること、
- 構文意味解析の処理漏れや誤処理を皆無にできないこと、

を考慮し、用字・用語、表現スタイル、分野特有表現に関する情報を収録した辞書を照合する方式とした。

本報告では、日本語ワードプロセッサ用に開発されたこの推敲支援ツール<sup>10)</sup>を基に、校閲作業支援ツールを開発し、約2,800ページのソフトウェアマニュアルの原稿作成に実際に適用して校閲作業への効果を評価している<sup>11)</sup>。さらに、1,000冊以上の商品マニュアルの開発に適用し実績をあげている。これらの内容を、以下、2章で推敲支援ツール適用の背景と目的、3章で校閲作業支援ツールの開発、4章で校閲作業への実適用評価、5章で校閲作業支援ツールの機能改善、6章で得られた結果と今後の課題のまとめ、なる構成で論述する。

## 2. 推敲支援ツール適用の背景と目的

### 2.1 適用の背景

ソフトウェア製品は、当該ソフトウェア本体だけでなく、それを扱うためのマニュアルとあわせて提供される。ソフトウェア本体が大規模かつ複雑になると同時に、ネットワークを前提として複数のミドルウェアやアプリケーションプログラムを組み合わせ使用することが一般的となっている昨今、それらを解説するマニュアルの果たすべき役割は大きくなっている。ま

た、提供形態は、従来の紙媒体に加え、計算機のディスプレイ上で参照するためのCD-ROM媒体やWWWやLAN, WAN等ネットワークを経由して提供するオンラインマニュアル等多岐にわたり、ユーザの利用環境にあわせてよりきめ細やかな対応が必須となっている。このようなきめ細やかな対応のためには、ドキュメントの開発効率を向上させる作業プロセスを整備すると同時に、短期間で商品としての品質を確保する施策の確立が重要である。

我々の所属する組織には、マニュアルをはじめとするユーザドキュメントを専任で開発する部署が設けられている。対象製品はメインフレーム、ワークステーション、パーソナルコンピュータ上で稼動するオペレーティングシステム、ミドルウェア、アプリケーションプログラム等多岐にわたる。さらに、それらを組み合わせシステムを構築する際に必要な製品横断のマニュアルも開発する。開発量は年間約60,000ページに及ぶ。これらのユーザドキュメントは約170名のテクニカルライターで開発している。在籍するテクニカルライターは、テクニカルライティングに必須の技術を習得し、マニュアル作法に関して一定水準の技術レベルに達している。具体的には、情報処理関連の基礎技術に加え、客観的文章表現や用字・用語の使用法、ソフトウェアの概念を記述する図の作成方法等を習得済みである。また、マニュアルの開発手順や遵守事項は数百ページに及ぶ規準書と用字・用語の使用法の手引書で標準化されており、テクニカルライターはその内容を学んでいる。

一定の技術水準に達している作業者が標準化された手順に沿って開発することを前提としても、商品マニュアルの開発という見地では、第三者による客観的な品質評価が必須である。マニュアル開発にはそのような評価工程をいくつか設けている。マニュアルは大まかに次の工程に沿って開発し、品質評価は以下の(2)、(4)、(5)、(6)の工程で実施する。

#### (1) プランニング

マニュアル開発担当のテクニカルライター(以下、マニュアル開発ライターと略す)が、マニュアル全体の構成、開発方針等の計画を立案する。

#### (2) プランニングレビュー

マニュアル開発ライターと、対象製品の開発担当者および製品検査の専任者として、(1)のプランニングの妥当性を評価する。評価時にあがった問題点は、マニュアル開発ライターが是正する。

#### (3) 原稿執筆

マニュアル開発ライターが、対象製品の仕様書を参照

しながらマニュアル原稿を作成する。作成にあたっては、規準書に規定されているマニュアル原稿執筆規準に従う。

#### (4) マニュアル開発部署校閲

マニュアル開発ライター自身、および上級ライターが(3)で作成した原稿の品質を評価する。この段階では主に、記述内容の一貫性や表現の分かりやすさといった観点を中心に評価する。評価時にあがった問題点は、マニュアル開発ライターが是正する。評価にあたっては、規準書に規定されているマニュアル校閲規準に従う。マニュアル開発ライターおよび上級ライターはこの規準を習得済みであり、かつ上級ライターは校閲作業に習熟している。なお、ここであがった問題点に基づいて、マニュアル原稿執筆規準、および校閲規準の内容は定期的に見直される。

#### (5) 原稿レビュー

マニュアル開発ライターと、対象製品の開発担当者および製品検査の専任者が、(4)で校閲した原稿の品質を評価する。この段階では主に、プランニング時の方針に沿って実際マニュアル開発がなされているか、対象製品の仕様上記述に誤りがないか、マニュアルとしての使用性に問題はないかといった観点を中心に評価する。評価時にあがった問題点は、マニュアル開発ライターが是正する。なお、ここであがった問題点に基づいて、マニュアル原稿執筆規準、および校閲規準の内容は定期的に見直される。

#### (6) 検査

製品検査の専任者が、(5)での問題点が是正されているかを中心に最終確認する。製品検査の専任者は、マニュアル開発者とは別の部署に属し、商品としてのマニュアルの品質を最終評価する。ここでマニュアルの内容に不良表現があった場合、マニュアル原稿はマニュアル開発部署に差し戻される。マニュアル開発部署には該当する不良表現の是正だけでなく、類似不良表現の有無の見直しおよび観点を変えた不良表現の有無の見直しが義務付けられる。マニュアル開発部署は、これらの見直しを完了後、再度検査を依頼する。なお、ここであがった問題点に基づいて、マニュアル原稿執筆規準、および校閲規準の内容は定期的に見直される。

### 2.2 適用の目的

マニュアルの品質評価のレベルは以下のように分類できる。

- 表記・統語レベル
- 意味・文脈レベル

品質評価は、あらかじめ作成した評価項目リストに沿って実施する。主な評価の観点は、マニュアル原稿

が対象製品の仕様を誤って記述していないか、文意は一意に解釈できて誤解を生じさせないか、マニュアル全体の構成は妥当か、等である。このような評価項目は、上記の意味・文脈レベルに相当する。これらは高次の知的活動を必要とするものであり、自動化や機械化は現時点ではきわめて難しい。

一方、表記・統語レベルについてはマニュアル開発ライターの不注意なミスに起因する誤りが多く、かつそれらを検出するための機械処理が比較的容易である。これらの不良表現を開発工程の早い段階で機械的に検出して修正することで、校閲者による品質評価の作業工数を低減でき、かつ評価観点が絞り込まれることで校閲者の心理的負担を軽減させて高次の評価作業に専念しやすくてできる。開発工程でいえば、原稿執筆が完了した直後に、機械的な検証処理に基づく表現修正を施すことが有効と思われる。

商品マニュアルでは「検出すべき不良表現を見落とす」という第1種の誤りは限りなく0に近づけることが必須である。逆に「検出すべきでない表現を指摘してしまう」という第2種の誤りは、目視による校閲工数を増大させるものの、ある程度許容することとした。

構文解析や意味解析等の自然言語解析を応用した推敲支援ツールが一般的に実用に供されているが、その方式では解析処理の不備に由来する処理漏れや誤処理により、思いもよらぬ表現が検出されたり、計算機処理時間が長くなったりし、商品マニュアルを対象とした大量文書処理には向いていない。これに対して、辞書照合の方式は、不良として混入しやすい字句や表現を辞書項目として網羅することで、確実に不良表現を検出できるため有効である。

前述の内容をまとめると以下ようになる。これらを適用に際しての方針とした。

- 校閲者による目視作業を支援する契機となればよい。
- コンパクトで高速に、かつ大量データを一括で処理できること。
- 辞書照合の方式を採用する。

上記の方針のもと、表記・統語レベルの校閲作業を支援する機能として以下の4分類を設定した。

#### (1) 用字・用語の統一

記述の一貫性を保つため、説明の対象となる事物や事象を指し示す用語は統一しなければならない。表記揺れを防ぐとともに、日本語入力システムの誤変換等を検出することが必要となる。これに該当する検出処理の観点として、具体的には常用漢字等、使用する漢字のコード範囲の確認、同訓異字、送り仮名、英字やカタカ

表 1 校閲作業支援ツールの機能

Table 1 Functions of the computer-assisted writing tool.

機能分類	#	機能項目	検出内容
用字・用語の統一	1	常用漢字	常用漢字外の漢字を検出する。
	2	漢字の使い分け	誤った同音異義語を使用している可能性のある記述を検出する。
	3	同訓異字	誤った同訓異字語を使用している可能性のある記述を検出する。
	4	仮名・漢字の使い分け	仮名と漢字の使い分けの誤りや、表記の不統一を検出する。
	5	送り仮名	送り仮名が誤っている記述を検出する。
	6	誤り語	誤った漢字を使用している可能性のある記述を検出する。
	7	キーワード一覧	指定した文字種(英字, カタカナ)の用語を一覧表示し, 表記の不統一を検出する。
表現スタイルの統一	8	接続詞・接続助詞	接続詞・接続助詞の用法が誤っている可能性のある記述を検出する。
	9	文体	丁寧表現(です・ます調)あるいは通常表現(だ・である調)の不統一を検出する。
	10	文語的表現	古めかしい表現がされている記述を検出する。
	11	指示語	指示語を検出する。
	12	受け身表現	受け身表現を検出する。
	13	括弧整合性	開き括弧・閉じ括弧の対応関係の不整合を検出する。
文表現指標の算定	14	文の長さ	設定した数値より長い文を検出する。
	15	読点の数	読点の多い文を検出する。
	16	段落の長さ	文の多い段落を検出する。
	17	文字列の長さ	長い特定文字種(平仮名, カタカナ, 漢字)から構成される文字列を検出する。
分野特有表現への統一	18	分野別ユーザ辞書	ユーザが独自に定義した分野別辞書で, 分野特有の表現に合致しない記述を検出する。

ナで構成される技術用語の揺れの検出等が該当する。

## (2) 表現スタイルの統一

用字・用語の一貫性に加え、マニュアルでは各文の表現形式や文間のつながりが一意に伝わらなければならない。接続詞・接続助詞や文体の統一の検証、指示語の検出や、持って回ったような表現に陥りやすい受け身表現の検出等が、この機能に該当する。

## (3) 文表現指標の算定

商品マニュアルには明快かつ簡潔な文章表現が要求されるとともに、客観的な事実の叙述や1センテンス1トピック記述の徹底が必須である。具体的には、文の長さ、段落の長さや用語を構成する文字列の長さ等の評価の機能が有効である。

## (4) 分野特有表現への統一

マニュアルの説明対象となるソフトウェアの性質や対象ユーザ、マニュアルの種類(入門書、操作説明書、リファレンス等)等の分野の相違によって、使用する文章表現が異なる。分野に特有な表現への統一化を支援するのがこの機能である。マニュアルの文章表現を分野に特有な表現に統一するための辞書(これを分野別ユーザ辞書という)を作成すれば、上記(1)、(2)、(3)の検出処理条件は同一にしたまま、適用対象文書の分野にあわせた最適な検出結果を得ることができる。さらに、マニュアル開発の規準書および用字・用語の

手引書での規定内容を反映させることで、定められた記述規則との不整合を検出できる。

上記の4分類からなる検出機能を実装して実際のマニュアル開発の校閲作業に適用すべく、日本語ワードプロセッサ用の推敲支援機能<sup>10)</sup>をベースとして、校閲作業支援ツールを開発した。校閲作業支援ツールの開発状況について、次章で説明する。

## 3. 校閲作業支援ツールの開発

校閲作業支援ツールの機能概要、および同機能に使用する辞書(以下、これを校閲作業支援用辞書という)の最適化について説明する。

### 3.1 校閲作業支援ツールの機能概要

表1は、校閲作業支援ツールの機能の一覧を示している。18種類から構成されるこれらの機能は、前節で記述した4機能分類、すなわち用字・用語の統一、表現スタイルの統一、文表現指標の算定、分野特有表現への統一に大別できる。表1の#1~#5、#10、および#18の機能項目については、校閲作業支援用辞書との照合によって記述内容を検証する。校閲作業支援用辞書には、検出すべき誤表記とそれに対する正表記、および表記上の注意事項を一对にして登録している。校閲対象文書で誤表記が検出された場合、それに対する正表記と表記上の注意事項とが対になって検出

[No]	[頁/行]	*****	[ 不良表現 ]	*****
1	4/12	ジャーラ環境定義	: または	: スケジュールコマ
2	7/2	的に実行されます。	: あらためて	: セッションを開始す
3	8/26	組み合わせによって	: 色々	: なプログラムファイ
4	13/18	示の定義」を参照し	: て下さい	: 。このコマンドを使

①

②

③

または	【No】 1	【正表記】 又は	【注意事項】 選択を意味する場合。
あらためて	【No】 2	【正表記】 改めて	【注意事項】 副詞の場合。
色々	【No】 3	【正表記】 いろいろ	【注意事項】 「色々」は使わない。
て下さい	【No】 4	【正表記】 てください	【注意事項】 補助動詞の場合。

図1 検出結果リストの出力例

Fig.1 An example of error list.

結果リストに出力される。たとえば「～のみならず」という表現はマニュアル原稿執筆規準では「～だけでなく」と書き改めるよう規定されており、表1の#10「文語的表現」の機能で検出される。「～のみならず」という表現がマニュアル原稿中に出現した場合、この表現が誤表記であること、またそれに対する正表記が「～だけでなく」であること、およびこの表現を使用する場合の注意事項が検出結果リストに出力される。検出結果リストの出力例を図1に示す。

図1は、表1の#4「仮名・漢字の使い分け」の検出機能についての出力を示す。図中の①、②、③は以下の説明の便宜上付けている項番である。

- 図中の①は機能項目名の見出しを示す。不良表現は表1に示す各機能単位に検出され、結果として出力される。
- 図中の②は、検出された表現を示す。[No]は検出された表現の通し番号を、[頁/行]は該当する表現のマニュアル原稿内での記述位置を、[不良表現]は不良として検出された表現（およびその前後の記述）を、それぞれ示す。
- 図中の③は、検出された表現についての説明を示す【No】は図中②内の通し番号[No]と対応する。【正表記】は検出された表現に対する正表記を【注意事項】はこの表現を使用する場合の注意事項をそれぞれ示す。たとえば、検出された表現「または」を選択の意味で使用する場合の正表記は「又は」であることを示す。

マニュアル開発ライタは、この検出結果リストを参照しながら、マニュアル原稿内の該当する不良表現を修正する（たとえば、選択を意味する原稿内の誤表記「または」を正表記「又は」に修正する等）。

また、表1の#14～#17は文表現指標の算定に関する機能項目であり、検出の閾値を数値で設定する（たとえば、#14に示した機能項目「文の長さ」でいえば、1つの文の長さとして許容する文字数は最大60文字

であると設定する等）。これら18種類の機能項目は任意に組み合わせて使用できる。

校閲作業支援用辞書の保守機能も提供している。マニュアル開発の規準書および用字・用語の手引書で規定している規準に従って、辞書内の語の追加、変更、削除といった再編成が可能である。校閲作業支援用辞書の再編成による最適化実績については、次節で説明する。

これらの機能はUNIXワークステーション上に実装した。複数文書を一括処理して、検出結果のリストを印刷またはファイル出力する。処理性能はプロセッサMC68030 25MHzで7,500文字/分であった。

校閲作業支援ツールは、マニュアル原稿を脱稿した後のマニュアル開発部署校閲の工程で適用する。本ツールを適用した場合のマニュアル開発部署校閲の作業の流れは以下のとおりとなる。

- (1) 校閲作業支援ツールによる検出結果リスト出力  
マニュアル開発ライタがマニュアル原稿に校閲作業支援ツールを適用し、検出結果のリストを出力する。
- (2) 検出結果リストの確認およびマニュアル原稿の修正

マニュアル開発ライタが検出結果リストの内容を確認し、マニュアル原稿の不良表現を修正する。その後、不良表現がなくなるまで校閲作業支援ツールの適用および原稿修正を繰り返す。

- (3) マニュアル校閲規準に基づく校閲

マニュアル開発ライタおよび上級ライタが、マニュアル校閲規準に従って、目視でマニュアル原稿を校閲する。

- (4) 校閲結果に基づく原稿の修正

マニュアル開発ライタが校閲結果の内容を確認し、マニュアル原稿の不良表現を修正する。その後、不良表現がなくなるまで目視による校閲および原稿修正を繰り返す。

### 3.2 校閲作業支援用辞書の最適化

校閲作業支援ツールの母体となった、日本語ワードプロセッサ用の推敲支援機能は、公用文、法令、技術文書、手紙等、様々な文書の記述内容を汎用的に検証することを目的としたものであった。照合用の辞書は以下の文献に基づいて抽出した用語から構成されていた。

- 常用漢字表（内閣告示第一号・内閣訓令第1号：昭和56年10月1日）
- 公用文における漢字使用等について（事務次官等会議申合せ：昭和56年10月1日）
- 法令における漢字使用等について（内閣法制局総発第141号：昭和56年10月1日）
- 学術用語集（文部省）

当初の辞書に登録されていた用語の中には、一般文書で使用するものでも、ソフトウェアマニュアルではほとんど使用しないものが含まれていた。また、通常の文書では誤りとなる表現をソフトウェアマニュアルでは意識的に使用する場合もある。たとえば、「取消」は「取り消し」の誤表記であるが、GUIのプルダウンメニューの表示として用いる場合がある等である。

本校閲作業支援機能では、検出すべき誤表記を見出し語として辞書に登録している。辞書から用語を削除した場合、その用語は検出されなくなり、本来検出すべき表現が検出漏れとなってしまう。たとえば、正表記「取り消し」で記述を統一するために「取消」や「取消し」が誤表記として辞書に登録されていたとする。これらの誤表記を辞書から削除してしまうと、該当する表現は検出されなくなり、「取り消し」以外の異表記（不良表現）が検出できなくなる。逆に辞書に用語を登録した場合、その用語は検出対象となるため、本来検出すべきでない表現がノイズとなって検出されてしまう。たとえば、GUIのメニュー項目では「取消」を使用することから、それに対する誤表記「取り消し」を検出させるために「取消」と「取り消し」の双方を辞書に登録したとする。その場合、「[取消]ボタンは操作の取り消しを意味します。」といった表現では「取消」と「取り消し」の双方が検出されてしまい、いわば検出ノイズとなる。しかしながら、校閲作業支援の目的は検出漏れを限りなく0に近づけることにあり、それによって検出ノイズが増加することはある程度許容することとした。前述の「取消」と「取り消し」の例でいえば、相互に正表記、誤表記として登録することとしている。また、マニュアル開発ライタは、用字・用語の使用方法の規準について一定の技術レベルに達していることから、当初の辞書に登録されている用語の

うちで明らかにマニュアルでは使用しない用語を削除することとした。

このように当初の推敲支援機能用の辞書内容を見直して校閲作業支援用辞書として再編し、ヒット率の向上および検出時のノイズ低減を図った。見直しの対象は表1に示した機能項目のうちの#1～#5、#10、および#18であり、いずれも辞書との照合による機能項目である。

再編に際しては、対象製品、種別（入門書、操作説明書、リファレンス等）、およびマニュアル開発ライタの経験年数に偏りがないように既存マニュアルを選定し、校閲作業支援ツールの検出結果およびマニュアル原稿への反映状況を詳細に分析した。分析対象としたマニュアルは11冊、約3,000ページであった。分析の結果をふまえ、不要な用語を削除するとともに、不足している用語を追加した。削除した語は、校閲作業支援ツールで検出された内容のうちでマニュアル原稿にまったく反映した実績がないもの、およびソフトウェア製品のマニュアルでは明らかに使用しないものであった。また、追加した語は、マニュアル開発の規準書および用字・用語の手引書に記載されていて既存の辞書にないものであった。

表2は再編の実績を示している。再編後の校閲作業支援用辞書は再編前の辞書と比較して、約2/3の語数となった。辞書の最適化による効果を実際のマニュアル原稿を用いて評価した。辞書再編時に分析対象としたマニュアルとは別の10冊、約3,000ページを評価対象とし、再編前、再編後のそれぞれの辞書を適用して、ツールによる処理時間および検出結果リスト量と比較した。その結果、校閲作業支援ツールの処理時間が約1割、検出結果のリスト量が約4割削減されることを確認した。なお、校閲作業支援用辞書は、実際のマニュアル開発での校閲、製品検査の結果を基に、定期的に収録用語を見直し、精度の向上を図っている。

## 4. 校閲作業への実適用評価

開発した校閲作業支援ツールは、これまで延べ1,000冊以上のマニュアル開発に実際に適用している。ここでは、校閲作業支援ツールの適用による効果およびそれに対する考察を記述する。なお、評価対象としたデータは39冊のマニュアルであり、原稿枚数はA4版換算で2,843ページであった。

### 4.1 不良表現の機械的検出の評価

マニュアル開発ライタおよび上級ライタが校閲規準に基づいて実施する校閲項目のいくつかは、校閲作業支援ツールの機能項目と重複する。これらの機能項目

表2 校閲作業支援用辞書の最適化内容  
Table 2 Details of optimizing the computer-assisted writing tool's dictionary.

#*	確認項目	語数(語)**	語例	辞書語数(語)	
				再編前	再編後
1	常用漢字	24	矩形, 罫線などを常用漢字扱いに変更	1,938	1,962
		0	-		
2	漢字の使い分け	82	こうもく(項目/綱目), かいそう(階層/改装)など	457	375
		0	-		
3	同訓異字	55	おこる(起こる/興る), かえす(帰す/返す)など	119	64
		0	-		
4	仮名・漢字の使い分け	536	相々傘, 皆既蝕, 昏睡など	1,112	659
		83	御願(お願い)など		
5	送り仮名	1,316	空地, 稲光り, 押問答など	1,775	460
		1	申込み(申し込み)		
10	文語的表現	1	より	60	72
		13	か否か(かどうか)など		
18	分野別ユーザ辞書	77	仮名・漢字の書き換え, 文語的表現辞書に移動	258	203
		22	定型ジョブ(定形ジョブ)など		
計		2,091	-	5,719	3,795
		119	-		

※表1に示した機能項目の項番に対応している。

※※上段は削除語数を, 下段は追加語数を, それぞれ示している。

表3 不良表現検出実績  
Table 3 Result of detecting errors.

校閲項目		校閲者不良表現検出・修正件数 (件/ページ)		不良表現検出率 (%) (a・b)/a×100
分類	確認の観点*	(a)ツール非適用	(b)ツール適用	
用字・用語	マニュアル開発の規程書に従っていない。 (キーワード一覧, 分野別ユーザ辞書, 仮名・漢字の使い分け, 文語的表現, 常用漢字)	0.38	0.06	84
専門用語	用語が不統一。 (キーワード一覧)	0.05	0.01	80
文章	日本語としておかしい。調子が硬い。 (文の長さ, 仮名・漢字の使い分け, 文語的表現, 指示語, 送り仮名, 受け身表現, 常用漢字, 接続詞・接続助詞)	0.21	0.05	76
	簡潔でない。 (文の長さ, 文語的表現, 受け身表現, 読点の数, 文字列の長さ, 段落の長さ)	0.06	0.03	50
誤り	記述に誤りがある。 (文体, 同訓異字, 漢字の使い分け)	0.05	0.02	60
	誤字・脱字がある。 (キーワード一覧, 漢字の使い分け)	0.12	0.07	42
計		0.87	0.24	72

※括弧内は, 各校閲項目に対応する校閲作業支援ツールの機能項目を示す。ツールで検出した不良表現を, 校閲規程に基づく校閲のみを実施した場合の不良表現検出・修正の内容と照らし合わせて, 機能項目を対応付けているため, 一部重複する機能項目がある。

について, 本ツールを適用していないマニュアル原稿を校閲規程に基づいて校閲した場合と, 本ツールで不良表現を機械的に検出して修正を加えた後に校閲規程に基づいて校閲した場合とを比較した。表3に比較結果を示す。

校閲者不良表現検出・修正件数とは, マニュアル開発ライターおよび上級ライターによる不良表現の検出・修正件数の実績であり, マニュアル原稿1ページあたりの検出・修正件数の平均値で示す。

下の式において, (a) は校閲規程に基づく校閲対象不良表現数を意味する。(b) はツールを適用した後に

残されている校閲対象不良表現数を意味する。不良表現検出率は, 本来人手で検出・修正しなかった不良表現のうち, 本ツールで機械的に検出できた割合を示す。すなわち,

$$\text{不良表現検出率}(\%) = \frac{(a) - (b)}{(a)} \times 100$$

である。なお, 校閲規程に基づく人手による不良表現検出・修正では, 見落としの可能性が0ではない。しかし, マニュアル開発ライター, および上級ライターによる校閲を経た後に実施される製品検査部門での差し戻し不良表現の実績は平均でマニュアル原稿100ページ

表 4 抽出・修正不良表現の内訳  
Table 4 Classification of detected errors.

#*	機能項目	不良表現抽出・修正件数(件)	全体に占める割合(%)
7	キーワード一覧	457	25.4
14	文の長さ	369	20.5
18	分野別ユーザ辞書	271	15.1
4	仮名・漢字の使い分け	136	7.6
10	文語的表現	126	7.0
11	指示語	89	4.9
5	送り仮名	69	3.8
12	受け身表現	65	3.6
9	文体	52	2.9
1	常用漢字	40	2.2
8	接続詞・接続助詞	39	2.2
15	読点の数	37	2.1
3	同訓異字	19	1.1
17	文字列の長さ	18	1.0
2	漢字の使い分け	9	0.5
16	段落の長さ	1	0.1
6	誤り語	0	0.0
	計	1797	100

※表 1 に示した機能項目の項番に対応している。

表 5 作業工数削減実績  
Table 5 Result of reducing proofreading time.

作業項目	作業工数(分/ページ)		工数削減率(%) (c-d)/c×100
	(c)ツール非適用	(d)ツール適用	
マニュアル開発ライターによる校閲	40.9	7.1	83
上級ライターによる校閲	11.7	9.3	21
計	52.6	16.4	69

あたり 0.03 件であり、無視できると見なした。

表 3 の結果は、校閲項目によってばらつきはあるものの、本来抽出・修正すべき不良表現の 42%から 84%、全体で 72%を校閲作業支援ツールで機械的に検出できることを示している。

また、抽出・修正した不良表現を校閲作業支援ツールの機能項目ごとに分類し件数を算出した。表 4 に結果を示す。

不良表現抽出・修正件数とは、校閲作業支援ツールが検出した不良表現に基づいて、マニュアル開発ライターが実際にマニュアル原稿を修正した件数を示しており、機能項目ごとに件数の多い順に配列している。全体に占める割合とは、不良表現抽出・修正件数の総数 1,797 件を 100 として、校閲作業支援ツールの各機能項目によって検出された不良表現の比率を示している。

キーワード一覧、文の長さ、および分野別ユーザ辞書の 3 機能項目によって検出された不良表現が、全不良表現件数の約 60%を占めている。キーワード一覧、文の長さの機能項目は、キーボード入力ミスや日本語入力システムでの変換ミスに起因する不良表現、1 文が長くなりすぎる不良表現を検出できる。原稿執筆時に、これらの不良表現を作り込まないようにマニュアル開発ライターがたえず意識することは、心理的負担が大きい。

このような用字・用語の不統一や、長すぎる文を検出するには、目視よりも、ツール適用による機械的処理の方がはるかに効率が良い。また、分野別ユーザ辞書の機能項目は、マニュアル開発の規準書や用字・用語の手引書との不整合を機械的かつ的確に検出できる。

校閲作業支援ツールは、表記・統語レベルの不良表現の検出に主眼を置いていることは前述したとおりである。このような不良表現を目視による校閲作業の前段階で機械的に検出して是正できることは、校閲作業の負担軽減とマニュアル原稿品質の早期作り込みにおいて効果がある。

#### 4.2 作業工数削減の評価

校閲作業支援ツールを適用した場合としない場合とで、マニュアル開発ライター自身、および上級ライターの校閲規準に基づく校閲作業の所要工数を比較した。表 5 に比較結果を示す。

作業工数とは、マニュアル開発ライターおよび上級ライターの作業時間の実績であり、マニュアル原稿 1 ページあたりの校閲作業に要した時間の平均値で示す。

下の式において、(c) は校閲作業支援ツールを適用せずに校閲規準に基づく校閲のみを実施した場合の実績値を意味する。(d) はツールを適用した後に校閲規準に基づいて校閲した場合の実績値を意味する。工数



削減率は、本来人手で校閲する場合に必要な作業工数のうち、本ツールの適用によって削減できた割合を示す。すなわち、

$$\text{工数削減率 (\%)} = \frac{(c) - (d)}{(c)} \times 100$$

である。

1冊あたりのマニュアル開発工数に占める校閲作業の割合は必ずしも大きくはないが、表5の結果は、マニュアル開発ライターおよび上級ライターの校閲作業工数のそれぞれ、83%および21%を削減できることを示している。ツール自体が表記・統語レベルの不良表現の検出に主眼を置いたため、マニュアル開発ライターの不注意なミスは抽出・修正に効果的である。一方、上級ライターによる校閲作業は意味・文脈レベルの不良表現抽出・修正に重点が置かれるため、削減率はマニュアル開発ライターの場合と比較してそれほど大きくはない。作業工数の削減という観点でも、校閲作業支援ツールの適用によって一定の効果を得ているといえる。

#### 4.3 考察

マニュアル39冊、計2,843ページ(A4版換算)に対する校閲作業支援ツールの適用結果について、以下の2つの観点から考察する。

##### ● 実用性

校閲作業支援ツールの適用により、抽出・修正すべき不良表現の42%から84%、全体で72%が機械的に検出できる。特に用字・用語の不良表現の検出については高い値を示している。これは、校閲作業支援用辞書の最適化に負うところが大きいと考える。ツール適用後にマニュアル開発ライターおよび上級ライターが校閲規準に基づいて抽出・修正した不良表現は、意味・文脈レベルでの判断を必要としているものであった。また、本ツールを適用する目的が「校閲するのはあくまで人間であり、校閲作業支援ツールはその契機および校閲の手がかりを提供できればよい」であることを勘案すれば、疑わしい表現は漏れなく検出することが、校閲作業を支援するうえで有効に働いていると考える。

また、校閲作業全体に要する工数の約69%が、校閲作業支援ツールの適用により削減できた。これは、年間あたりの総開発ページ数が60,000ページであることから換算すれば、年間36,200時間の工数削減に相当する。マニュアル開発ライターおよ

び上級ライター1人あたりでは年間200時間強の工数削減となる。マニュアル開発の効率を向上させるという観点で、この効果は大きい。

また、不良表現は開発工程の上流で抽出・修正するほど効果的である。マニュアル開発ライター自身が脱稿直後に、校閲作業支援ツールを適用して一定量の不良表現を抽出・修正する。このことによって副次的な作業工数も軽減できる。具体的には、上級ライターが校閲で発見した不良表現をマニュアル開発ライターへ申し送るためのドキュメント作成工数、作業者間の連絡工数、マニュアル開発ライターによる類似不良表現の見直し工数等の削減である。

##### ● 操作性、使用性

現状一般的に実用に供されている推敲・校閲支援ツールは対話型のインタフェースを採用しているものが主流である。それに対して、本ツールは複数文書を一括して処理するバッチ方式を採用している。ツールによって検出された表現すべてを修正するわけではないこと、およびマニュアル開発ライターのマシン操作の手間が省けるという点で効果的である。しかし反面、ツールが出力する検出結果のリストをマニュアル開発ライターが精査して、それを手がかりに修正すべき不良表現を検索するといった手間が発生する。マニュアル原稿の編集ツールと連動させて、不良表現の表示および修正作業の対話処理を実現する等の改善が必要であると考える。

また、マニュアル開発ライター自身が校閲作業支援ツールを適用して校閲前に不良表現を抽出・修正することで、上級ライターの作業工数を低減できる。しかし、マニュアル開発ライター自身が検出結果のリストを結果的に目視で全件確認する際の不良表現見落としの可能性は抑止できていない。たとえば、表記の揺れの検出等で同様の不良表現が複数ある場合には、総件数を不良表現とともに提示する等、全体の傾向を統計情報として与えるといった工夫が必要である。また、同じ不良表現が複数ある場合は、マニュアル原稿の編集ツールと連動して不良表現を一括修正する等の機能も効果的であると考える。

## 5. 校閲作業支援ツールの機能改善

近年、商品マニュアルは複数メディアへの対応のニ

36,200時間 = ページあたり削減工数(表5参照)(52.6 - 16.4)分/ページ × 年間総開発ページ(60,000) ÷ 60

200時間強 = 年間あたり削減工数(36,200時間) ÷ ライター総数(170名, 2.1節参照)

ズが高まっており、そのようなマニュアルを高効率で開発することが必要となっている。そのための手法として、SGML ( Standard Generalized Markup Language ) をはじめとするマークアップ言語の使用があげられる。タグと呼ばれる制御コードを機械処理の手がかりとして、単一の文書データから、複数メディアに対応した商品マニュアルが開発できる。当該ドキュメント開発専任部署でもマニュアル開発の前提としてSGMLを採用している。また、計算機の低価格化を背景として、現在では各マニュアル開発ライターに専用のパーソナルコンピュータが与えられており、マニュアル開発自体にもそれを使用している。当初校閲作業支援ツールは、複数のマニュアル開発ライターが共用するUNIXワークステーション上に実装した。しかし、上記のようなマニュアル開発環境の変遷から、パーソナルコンピュータへの移植が必要となった。以下、パーソナルコンピュータ版の開発に際しての機能改善内容、および今後の課題を記述する。なお、機能改善による効果の算定は、継続調査課題となっている。

### 5.1 機能改善の概略

マニュアル原稿は市販のワードプロセッサソフトを使用して作成し、それをSGML形式にデータ変換する。パーソナルコンピュータ版の校閲作業支援ツールは、ワードプロセッサソフトのアドオンとして実装した。UNIXワークステーション版と同等の機能を実装するとともに、作業効率を向上させるために以下の改善を行った。なお、パーソナルコンピュータ版はPentium 120 MHzのプロセッサで実用上問題なく動作し、130,000文字/分の処理性能である。

#### ● 検出した不良表現の表示

校閲作業支援ツールで検出した文書内の不良表現に電子的なしおりを挿入し、不良内容の説明とともに一覧表示する。校閲作業支援ツールの使用者は検出結果のリストを机上ではなく画面上で確認できるとともに、修正が必要な不良表現へ直接ジャンプしてその場で編集できる。

#### ● 同件の指摘の件数表示

キーワード一覧のように、表記の揺れを検出する機能項目については、同じ不良表現を累計の件数で表示する。これにより、ツールの使用者が確認すべき検出結果のリスト量を削減でき、かつ検出結果全体の傾向が把握しやすくなる。

#### ● タグテキストへの対応

マニュアル原稿は基本的にワードプロセッサソフトを用いてWYSIWYGで作成し、それをSGML形式にデータ変換する。しかし、一部のマニュアル

はマークアップ言語のタグをテキストエディタで直接入力・編集して作成する場合がある。また、他社から導入した製品のマニュアルデータが、マークアップ言語で記述されている場合もある。タグは、その意味を解釈して積極的に機械処理に活用するのであれば、校閲作業支援ツールにとって単なるノイズとなる。このような状況に対応するため、マニュアルデータを処理する前にタグの有無を判定し、マークアップ言語で記述されている場合には、タグを無視して処理するよう実装している。これにより、SGML、HTML ( Hypertext Markup Language )、XML ( Extensible Markup Language ) に代表されるマークアップ言語全般のマニュアルデータを処理できる。

図2にパーソナルコンピュータ版で実装した処理方式を、図3に校閲作業支援ツールの画面例を、それぞれ示す。

### 5.2 今後の課題

マークアップ言語のタグを利用することにより、文書要素特有の表現方法に適合しない不良表現を機械的に検出することが可能である。パーソナルコンピュータ版の校閲作業支援ツールは、タグをノイズとして無視するよう実装している。しかし本来、タグは文書構造を示す道具であり、文章内の記述要素を判定する有効な手がかりとなる。特定の記述要素と、それが記述されるべき表記の形式とを相互に関連付けて不良表現を機械的に検出することも容易である。たとえば、文末の形式は「だ・である調」で記述することを基本とするが、注意事項の記述は「～しないでください」というように例外的に「です・ます調」で記述することが取り決められている状況を仮定する。文体の不統一の検出機能で、本文は「だ・である調」で記述されているかを確認し、注意事項の記述はタグを識別して正しく「です・ます調」で記述されているかを確認する、といった処理の振り分けが可能となる。このようにタグやその属性等と不良表現検出条件とを組み合わせることで、よりきめ細やかな記述内容の検証が実現できる。また、タグと不良表現検出条件とを組み合わせることで定義ファイルとして保存すれば、複数のタグセットや、対象となるドキュメントの性質、種別によってツールの検出処理を可変にできる。これは一例にすぎないが、静的な不良表現検出条件を実装するだけでなく、対象ドキュメントの特性に応じて、より柔軟な不良表現検出方式を継続して検討することが重要である。

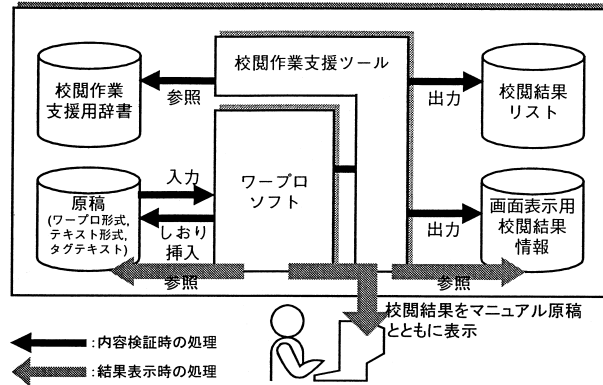
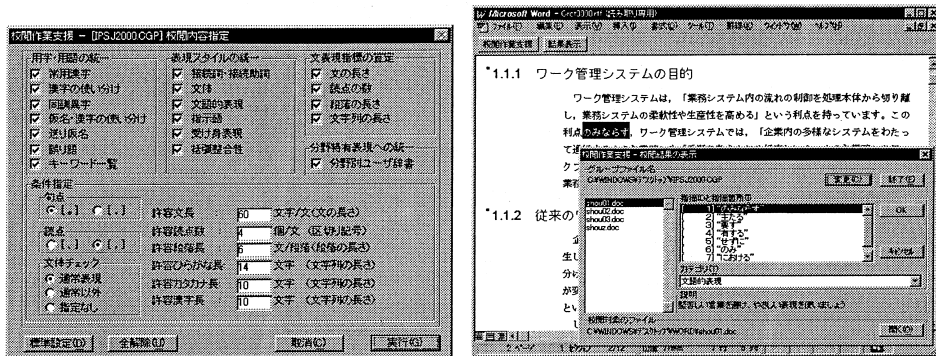


図2 パーソナルコンピュータ版での処理方式

Fig. 2 Processing method of the personal computer version.



校閲内容指定ダイアログ

(デフォルトの指定状態を示す。検出対象とする・しないを機能項目単位に指定でき、かつ検出の数値条件等を指定できる。)

校閲結果表示ウィンドウ

(検出結果及びそれに該当する記述表現が表示された状態を示す。前面ダイアログには検出結果が、背面ウィンドウには該当する記述表現が反転して表示される。本例は、文語的表現「のみならず」が検出された状態を示す。)

図3 パーソナルコンピュータ版の画面例

Fig. 3 An example of a window in the personal computer version.

## 6. おわりに

高品質なマニュアル製品の開発に際しては、マニュアル原稿脱稿時、マニュアル開発ライターおよびその上級ライターによる校閲作業が必須である。この校閲作業を効率化し、目視作業をより高度なものとするために、単純な表現上のミス等の検出作業を軽減することが必要である。この検出作業軽減のため、用字・用語の統一、表現スタイルの統一、文表現指標の算定、分野特有表現への統一、の4つに大別できる18の処理機能からなる校閲作業支援ツールを開発した。これらの機能は、人による目視作業を支援できる契機となればよいこと、コンパクトで高速に処理できること、出現新語に対する形態素解析用情報の逐次付与の困難さ、構文意味解析の処理漏れや誤処理を皆無にできないこと、か

ら、校閲すべき表現を登録した辞書との照合方式で実現し、UNIXワークステーション(プロセッサMC68030 25 MHz、処理性能7,500文字/分)およびパーソナルコンピュータ(プロセッサPentium120 MHz、処理性能130,000文字/分)上に実装した。また、この校閲作業支援ツールを1,000冊以上のソフトウェアマニュアルの開発に実際に適用している。そのうち、マニュアル39冊、2,843ページを使用して実績評価した。その結果、本ツールを使用せず、全数目視作業する場合と比べて、本ツールを校閲作業に使用する場合は、校閲項目別の不良表現摘出・修正件数について、42%から84%を削減し、校閲作業の工数については、マニュアル開発ライターの内容確認作業を83%、上級ライターの校閲作業の21%を削減する効果があり、本校閲作業支援ツールの導入が、当初の目的を実現していること

を確認した。今後に残された課題としては、SGML、XML等のマークアップ言語で記述されたマニュアル原稿を対象として、文書構造の定義内容を利用した新たな不良表現検出機能の開発がある。

謝辞 本研究の機会を与えていただくとともに、推敲支援ツールのマニュアル校閲作業支援への適用を図っていただいた(株)日立製作所ソフトウェア開発本部マニュアル開発部長の松永芳之氏〔当時、現(株)日立テクニカルコミュニケーションズ〕に感謝いたします。

### 参考文献

- 1) 木下是雄：理科系の作文技術，中央公論社(1981)。
- 2) 大用昌之：次世代ワープロの決め手となるか校正支援/可読性評価ツール，日経バイト，No.43，pp.96-104(1988)。
- 3) 牛島和夫，日並順二，尹志熙，高木利久：日本語文章推敲支援ツールのプロトタイプング，コンピュータソフトウェア，Vol.3，No.1，pp.35-46(1986)。
- 4) 福島俊一，大竹暁子，大山裕，首藤友喜：日本語文章作成支援システムCOMET，電子通信学会，オフィスシステム研究会，OS86-21，pp.15-22(1986)。
- 5) 鈴木恵美子，武田浩一：日本語文書校正支援システムの設計と評価，情報処理学会論文誌，Vol.30，No.11，pp.1402-1412(1989)。
- 6) 池原悟，安田恒雄，島崎勝美，高木伸一郎：日本語訂正支援システム(REVISE)，NTT研究実用化報告，Vol.36，No.9，pp.1159-1167(1987)。
- 7) 小山紀子，斎藤裕美，小林賢一郎，小山和雄：文章校正支援機能における日本語解析，情報処理学会自然言語処理研究会，Vol.69-2(1988)。
- 8) 空閑茂起：文書作成・校正支援システムWISE，電子通信学会，オフィスシステム研究会，OS86-28，pp.13-18(1986)。
- 9) 高橋善文，吉田哲三：計算機マニュアル推敲・査読支援システムMAPLEの開発と運用，情報処理学会論文誌，Vol.31，No.7，pp.1051-1062(1990)。
- 10) 光行博志，松岡慎二，佐藤裕，山本昭彦，松田泰昌，絹川博之：日本語ワードプロセッサ「ワードバル」の新シリーズ，日立評論，Vol.70，No.9，pp.25-30(1988)。
- 11) 笠原健成，絹川博之：文書推敲支援ツールのマニュアル開発への適用事例，情報処理学会デジタル・ドキュメント研究会，Vol.98，No.23，pp.19-26(1998)。

(平成11年10月26日受付)

(平成13年2月1日採録)



笠原 健成(正会員)

1989年早稲田大学教育学部教育心理学専修卒業。1992年学習院大学大学院心理学専攻博士前期課程修了。同年(株)日立製作所入社。現在同社ソフトウェア事業部にて顧客提供ドキュメント制作環境の開発に従事。日本心理学协会会员。



小林 栄一(正会員)

1985年早稲田大学法学部卒業。同年(株)日立製作所ソフトウェア工場入所。DBMSの開発を経て、1988年よりソフトウェアマニュアル制作環境の研究開発に従事。現在、ソフトウェア事業部テクニカルインフォメーション部主任技師。



荒井 真人(正会員)

1972年東北大学工学部電気工学科卒業。同年(株)日立製作所ソフトウェア工場入所。1985年よりマニュアル開発部においてソフトウェアマニュアルの開発に従事。著書「ソフトウェアマニュアルの開発」(日科技連)(共著)等。現在ソフトウェア事業部大型プロジェクト推進部長。



絹川 博之(正会員)

1947年生。1970年東京大学理学部数学科卒業。同年(株)日立製作所入社。以来、漢字・日本語情報処理、仮名漢字変換、文書処理、自動インデクシング、情報検索、自然言語処理、自然言語インタフェース、電子図書館、知識管理等の研究開発に従事。同社システム開発研究所研究主幹、東京工業大学大学院客員教授を経て、1999年4月より、東京電機大学工学部情報通信工学科兼大学院工学研究科教授、理学博士。2000年11月より、国立情報学研究所客員教授を併任。昭和61年度情報処理学会論文賞、平成7年度関東地方発明表彰発明奨励賞受賞。ACL、言語処理学会、電子情報通信学会、ヒューマンインタフェース学会各会員。