

テキスト校正ツールを活かすための  
過指摘回避辞書構築 報告書

2010年3月

一般財団法人テクニカルコミュニケーター協会

過指摘回避辞書構築ワーキンググループ



## はじめに

当協会では、これまでさまざまな使用説明（製品カタログ、マニュアル、製品に組み込まれた操作支援情報、Web上の製品情報などの総称）の作成および発信に関する標準化活動として、マニュアル評価ガイドライン、外来語（カタカナ）表記ガイドライン、電子マニュアルマークの標準化、地上デジタル放送対応テレビの取扱説明書における記載項目のガイドラインなど、さまざまな提言を行ってきました。

2008年7月に、技術研修会「校正工程の品質安定と効率向上を徹底追求する～ツールの機能の効果的活用～」を実施しました。この場において、電子校正ツールの利用ガイドライン策定を推進することについて多くの賛同が得られたことを契機に、会員有志によるワーキンググループを発足させ、文章校正ツールの活用に際して過剰な指摘（過指摘）を回避して、校正作業の円滑な遂行に役立つ辞書構築について検討を重ねてきました。

その結果、TC協会の会員有志より提供された取扱説明書のテキストデータ解析を進め、文章校正支援ツールに登録することで過指摘の回避に役立つ単語のリスト抽出を行いました。このリストが多くの使用説明の提供にかかわる人々を中心に、全国的に利用されることにより、文章校正支援ツールの活用が円滑に行われることを期待します。

また、このリストは、TC協会会員をはじめ、使用説明の制作・発信に携わる関係者からの意見や要望を反映することで、今後も改訂を重ねてまいります。改訂されたリストは、適宜、当協会のWebサイトにおいて公開する予定です。皆様からのコメントを歓迎します。

2010年3月  
一般財団法人テクニカルコミュニケーター協会

## 目次

|                           |           |
|---------------------------|-----------|
| <b>1 概要</b> .....         | <b>3</b>  |
| 活動の背景.....                | 3         |
| 活動の目的.....                | 3         |
| 体制.....                   | 4         |
| 活動期間.....                 | 4         |
| <b>2 活動の経緯</b> .....      | <b>5</b>  |
| 1) 過指摘の実態調査.....          | 5         |
| 調査の目的.....                | 5         |
| 調査の方法.....                | 5         |
| 調査の結果と検討.....             | 7         |
| 2) 過指摘回避語句の収集.....        | 9         |
| 語句の収集方法.....              | 9         |
| 語句の選定方針.....              | 9         |
| <b>3 今後の活動と課題</b> .....   | <b>10</b> |
| 用語数の充実.....               | 10        |
| ツールベンダーへの働きかけ.....        | 11        |
| <b>付録 用語リスト(抜粋)</b> ..... | <b>12</b> |

## 1 概要

### ◆ 活動の背景

校正支援ツールとは、「日本語テキストデータ（欧文で表記される単語などを含む）を対象として、文法的誤り、表現の不統一、用字用語の表記のゆれ、不適切表現などを指摘するツールである。

マニュアル制作現場においても、対象機器やシステムの複雑化に伴う執筆の分業化や、制作期間の短縮化などを背景にして、校正支援ツールを実際に導入している、あるいは導入を検討しているところが増えている。

確かに校正支援ツールの利便性は高い。しかし、ツールが標準で装備している辞書を使用した場合、TC 分野では一般的に用いられる表記にもかかわらず、ツールから表記間違いの可能性を指摘されることがある。このように、制作者側が意図的に使用している表記に対して校正支援ツールが指摘してくる状況を、「過指摘」と呼ぶ。

テクニカルコミュニケーター協会が 2008 年 7 月に開催した技術研修会「校正工程の品質安定と効率向上を徹底追求する～ツールの機能の効果的活用」において、こうした過指摘を減らし、マニュアル制作における校正支援ツールの利便性向上を推進する活動を提案したところ、技術研修会参加者から多くの賛同が得られた。

### ◆ 活動の目的

このワーキンググループ（以下、WG）の目的は、マニュアル制作において過指摘と判断される用語を調査・収集し、校正支援ツールに取り込むことによって過指摘を回避する用語リストをまとめることにある。

WG の最終成果物としては、個別の校正支援ツールに特化した辞書ファイルではなく、種々の校正支援ツールで利用されることを前提として、用語リストという形で、汎用性のある csv ファイルを公表することとした。マニュアル制作現場においては、用語リストの内容を確認いただき、各現場の判断により、各々が使用している校正支援ツールで利用可能な辞書ファイルに変換したうえで利用していただきたい。

## ◆ 体制

WG のメンバーは以下のとおりである。

### 主査:

白輪 琢郎 東芝テック画像情報システム株式会社

### 副主査:

若山 陽介 株式会社情報システムエンジニアリング

### メンバー:(五十音順)

貝島 良太 有限会社アトリエ・ワン  
 金子 泰規 株式会社ニコンシステム  
 柴崎 茂 横河電機株式会社  
 徳田 直樹 株式会社パセイジ  
 前川 純子 株式会社サン・フレア  
 蓑田 正俊 NEC デザイン&プロモーション株式会社  
 村尾 昌浩 株式会社ジャストシステム  
 山口 昌美 富士ゼロックスエンジニアリング株式会社

### 事務局:

雨宮 拓 テクニカルコミュニケーター協会事務局

## ◆ 活動期間

|            |                                          |
|------------|------------------------------------------|
| 08年9月10日   | WGメンバー募集案内                               |
| 08年11月10日  | 第1回会合(キックオフ)                             |
| 09年1月      | 過指摘の実態調査に利用する校正支援ツールの検討                  |
| 09年02月06日  | 第2回会合(各校正支援ツールの概要確認、実態調査に利用する校正支援ツールの決定) |
| 09年02月~03月 | 実態調査準備                                   |
| 09年04月~05月 | 実態調査実施                                   |
| 09年05月~06月 | 実態調査結果の分析                                |
| 09年07月09日  | 第3回会合(実態調査結果の確認、中間報告に向けた検討)              |
| 09年07月~08月 | 中間報告準備                                   |
| 09年08月11日  | 第4回会合(中間報告内容の事前確認)                       |

|               |                              |
|---------------|------------------------------|
| 09年08月26日     | 中間報告実施（TC シンポジウム）            |
| 09年09月～10月    | 過指摘語句の収集に使用するマニュアルの募集        |
| 09年11月        | 過指摘語句のリストアップ・分析              |
| 09年12月08日     | 第5回会合（過指摘語句の、用語リストへの登録方針の確認） |
| 09年12月～10年01月 | 用語リストへの登録・登録した内容のチェック        |
| 10年01月～10年02月 | 最終報告書作成                      |

## 2 活動の経緯

### 1) 過指摘の実態調査

WGメンバーが制作しているマニュアルにおいて生じる過指摘を調査した。

#### ◆ 調査の目的

この調査の目的は、マニュアルを実際に校正支援ツールにかけたときにどのような過指摘が発生するのかを確認し、それらを分類整理し、その後の調査や活動の方針を検討するための材料とすることにある。

#### ◆ 調査の方法

調査対象には、WGメンバーが自社で制作していて、すでに市場に出ているエンドユーザー向けのマニュアルを使用した。

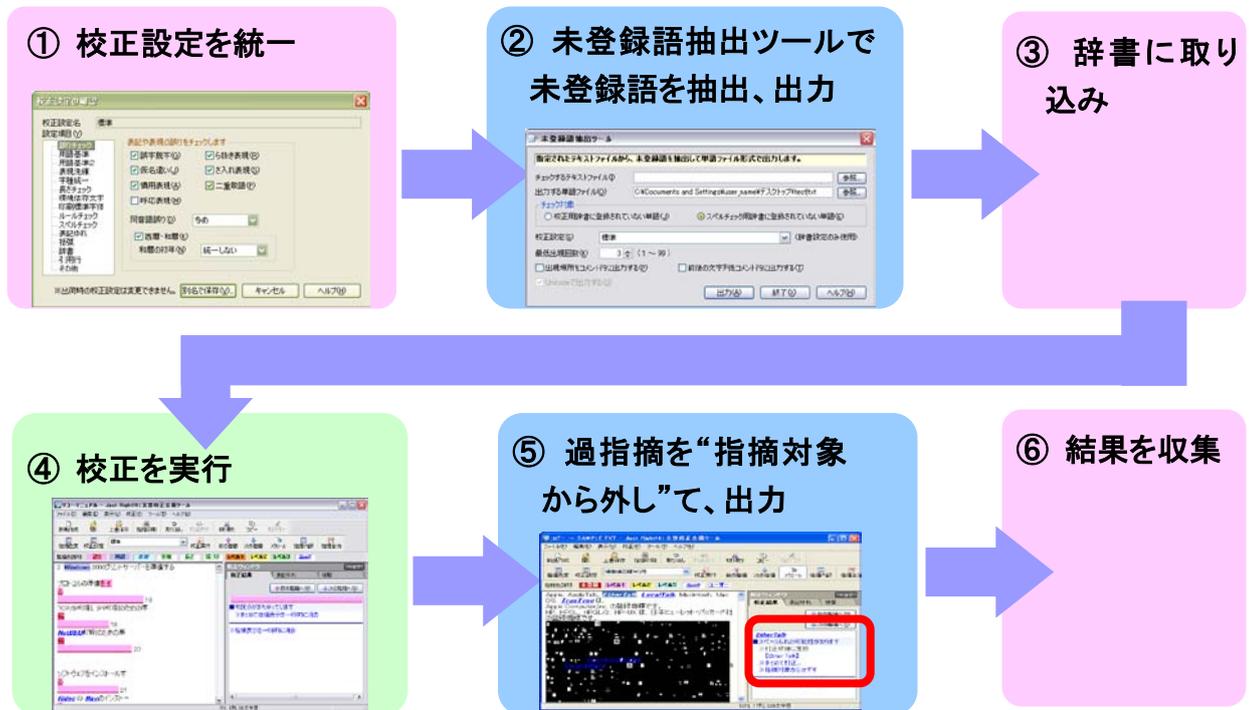
調査に使用するツールとしては、市販されている以下の3つを候補として検討した。

- Acrolinx IQ（株式会社 サン・フレア／アクロリンクス株式会社 販売）
- Just Right!4 Pro CE（株式会社ジャストシステム 開発・販売）
- SuperHT<sup>3</sup>（有限会社アトリエ・ワン 開発・販売）

WGメンバーにおける利用数を考慮して、調査には「Just Right!4 Pro CE」を使用することとした。

過指摘抽出の具体的な手順を以下に示す。

図 1 過指摘抽出作業の流れ



- ① 本ツールの校正設定群から、TC ドキュメントの校正になじまない項目を除外し、WG メンバー間で校正設定を統一する。
- ② 対象ドキュメントを、本ツールに付属の「未登録語抽出ツール」<sup>1)</sup>にかけ、「辞書に登録されていないために生じた過指摘」を抽出する。
- ③ 抽出した未登録語を、本ツールの辞書に取り込む。
- ④ 対象ドキュメントを、本ツールで校正する。
- ⑤ 過剰な指摘を「指摘対象外」に指定して、ファイルに出力することにより、「辞書には登録されているが、別の理由で生じた過指摘」を抽出する。

上記の手順により、校正用辞書に登録されていない単語とスペルチェック用辞書に登録されていない単語を抽出した。

<sup>1)</sup> 「Just Right!4」の-Pro CE/-CE版に付属するツール。辞書に登録されていない以下の単語を、テキストファイルからまとめて抽出・出力する。

- 校正用辞書に登録されていない単語
- スペルチェック用辞書に登録されていない単語

### ◆ 調査の結果と検討

抽出した過指摘を、以下のように4つに分類した。4分類の特徴と、それらを検討して導出した過指摘回避方法を以下に示す。

図2 過指摘の4分類



#### ① [辞書要因]－[一般的]

##### 特徴:

- TC分野に広く共通して発生する
- あらかじめ大量に語句を収集しておけば、幅広い過指摘回避ニーズに応えられる

##### 例:

漏液、防爆、落射、傾角鏡筒、ターレット、RoHS、RGB、CMYK、など

このグループは、TC業界では一般的に使われている語句にもかかわらず、ツールが指摘してく

るものである。あらかじめこれらの語句を大量に集めておき、校正支援ツールの辞書に登録することにより指摘を回避できる。

## ② [辞書要因]－[専門的]

### 特徴:

- 一部のドキュメントにのみ発生する
- 現場ごとに辞書に登録することで回避できる

### 例:

自社製品名、特定のファイル名やディレクトリー名

このグループは、校正支援ツールが備えている標準辞書には登録されていない。また、TC 業界全体としても一般的であるとは言えず、制作現場ごとに語句を辞書に登録して回避すべきものである。

## ③ [ツール要因]－[一般的]

### 特徴:

- ツールの仕様により、現状では回避できない
- 指摘を回避できるよう、ツール側で対応が進められるのが望ましい

### 例:

- リーダー罫、片括弧、= (等号) や ¥ (ディレクトリーの区切り記号) の前後の語、16 進表記
- 数字とそれに続く単位 (15mm など)
- 拡張子とその前の語 (○△□.exe)
- 英単語に数字が続く語 (AUX2 など)
- http : // . . .

このグループは、現状のツールの仕様では指摘を回避できないため、指摘回避のための機能の盛り込みなどを、ツールベンダーに求めていく。

## ④ [ツール要因]－[専門的]

### 特徴:

ドキュメントのデータや作り方などに起因する

**例:**

- PDF ファイルをチェック対象にした場合に生じる、途中改行や余分なスペース
- 字どり組みを実現するために文字間にスペースを入れたことにより、ツールが適切に文を解析できなかった部分

このグループは、DTP データの作り方を見直すことで回避できるものもあるが、現状のツールの仕様では回避できないものもある。指摘回避のための機能の盛り込みなどを、ツールベンダーに求めていく。

以上の検討から、マニュアル業界全体で共通の対策が可能なのは、①「辞書要因」－「一般的」のグループであり、その対策は業界で広く使われている語句を大量に収集することによって可能になると考えられる。

## 2) 過指摘回避語句の収集

### ◆ 語句の収集方法

TC 業界で広く使われているが、校正支援ツールが備える辞書には登録されていない語句を収集するため、TC 協会の法人会員各社に依頼し、すでに市場に出ているマニュアルの PDF ファイルを提供いただいた。これらの PDF ファイルについて、校正用辞書に登録されていない単語と、スペルチェック用辞書に登録されていない単語を、株式会社ジャストシステムに依頼して抽出を行った。こうして得られた語句について、用語リストに登録すべきものを WG メンバーが選定した。

### ◆ 語句の選定方針

選定に当たっては、以下の条件に該当する語句を除外することとした。使用頻度の判定に際しては、調査対象マニュアル 85 冊のうち、複数メーカーまたはマニュアルに合計で数回しか出現していない語句を「低い」として基本的に除外している。なお、単一メーカーまたはマニュアルで使用頻度が高い用語であっても、汎用性を考慮して除外している場合がある。

校正用辞書に登録されていない語句のうち、

- 使用頻度が低いと思われる社名、製品名、サービス名などの固有名詞
- 使用頻度が低いと思われる外来語（人名・地名を含む）
- 使用頻度が極端に低いと思われる技術用語、専門用語（「名詞サ変」は造語を増やすことになるので、個別に判断し、「名詞」として登録する）

- 送り仮名本則に従っていない漢語（合成語を含む）
- 常用漢字でない漢字を、ひらがなで表記した漢語
- TC 協会の「外来語（カタカナ）表記ガイドライン」に準拠していないカタカナ語
- ひらがなで表記されたカタカナ語

スペルチェック用辞書に登録されていない語句のうち、

- 括弧記号、ハイフンなどを含む語
- 明らかにスペルミスと思われる語
- 使用頻度が低いと思われる社名、製品名、サービス名などの固有名詞
- 指摘理由が「辞書に未登録」以外の語

また、以下の点にも留意した。

#### **カタカナ語の複合語:**

2つ以上の語から成る複合語は、校正支援ツールが備える辞書には登録されていない場合が多い。その場合は、複合語自体を登録対象とするのではなく、まず単語レベルに分解し、その中の未登録の単語部分を登録対象とする。

例) 「トップテザー」 = 「トップ」 + 「テザー」・・・「テザー」で登録。

#### **カタカナ語の長音記号:**

「外来語（カタカナ）表記ガイドライン」にならう。

最終的に用語リストに登録した語数は、日本語 181 件、スペリング 469 件となった。また、用語リストは TC 協会 Web サイトにて公開する。用語リストの一部を本報告書の末尾に掲載したので参照いただきたい。

### **3 今後の活動と課題**

WG としては活動を終了するが、以下に見るよういくつか課題も残った。今後、用語リストの改訂作業とともに、新たな WG を編成して対応策を検討していく。

#### **◆ 用語数の充実**

今回の調査において、用語リストへの登録語句の選定には、TC 協会の法人会員各社に提供いただいたマニュアルの PDF ファイルを使用させていただいた。あらゆる TC 分野からまんべんなく用語を収集するのが理想的だが、調査に使用したマニュアル数は限られており、あらゆる分野を網羅することはできなかったため、収集した用語数は充分ではなく、偏りも見られる。

#### ◆ ツールベンダーへの働きかけ

TC 分野で一般的に使われる用語を集めて校正支援ツールの辞書に登録するだけでは回避できない過指摘については、これを回避するための対策をツールに盛り込むよう、ツールベンダーに提案していく必要がある。

今回の活動を通して、使用する校正支援ツールによっては、PDF ファイルを校正対象にすることに起因する過指摘が存在することが分かった。しかし、実際の制作現場では PDF ファイルを校正対象にするケースが多いため、これを前提としたツールの改良が望まれる。

\*記載されている会社名および製品名は、各社の登録商標または商標です。

## 付録 用語リスト(抜粋)

### ◆ 日本語

| 表記          | 品詞   | 出現数 | 出現マニュアル数 |
|-------------|------|-----|----------|
| アッテネーター     | 名詞   | 55  | 9        |
| アッテネート      | 名詞サ変 | 10  | 3        |
| アッテネート      | 名詞   | 10  | 3        |
| アルペジエーター    | 名詞   | 10  | 2        |
| アンプリチュード    | 名詞   | 44  | 3        |
| エクスターナル     | 名詞   | 27  | 3        |
| オルタネート      | 名詞   | 6   | 3        |
| カチツ         | 副詞   | 37  | 15       |
| クオンタイズ      | 名詞サ変 | 43  | 2        |
| クオンタイズ      | 名詞   | 43  | 2        |
| クラビノーバ      | 固有一般 | 91  | 2        |
| スレッシュホールド   | 名詞   | 45  | 4        |
| セーム革        | 名詞   | 6   | 4        |
| セブンス        | 名詞   | 17  | 3        |
| ソステヌート      | 名詞   | 15  | 3        |
| ディマー        | 名詞   | 56  | 6        |
| テザー         | 名詞   | 34  | 3        |
| ドルビーラボラトリーズ | 固有組織 | 30  | 15       |
| ピーツ         | 副詞   | 15  | 3        |
| ヒーテッド       | 名詞   | 10  | 2        |
| ピツ          | 副詞   | 34  | 10       |
| ピピツ         | 副詞   | 36  | 9        |
| ファミリンク      | 固有一般 | 491 | 2        |
| フェーザー       | 名詞   | 9   | 2        |
| フェーダー       | 名詞   | 228 | 5        |
| フッテージ       | 名詞   | 13  | 2        |
| ベンド         | 名詞   | 56  | 4        |
| ポルタメント      | 名詞   | 69  | 3        |
| ラウンチャー      | 名詞   | 56  | 3        |
| ロガー         | 名詞   | 10  | 2        |
| ワウ          | 名詞   | 25  | 2        |
| 押鍵          | 名詞サ変 | 20  | 3        |
| 横縦          | 名詞   | 27  | 6        |
| 給排紙         | 名詞   | 10  | 3        |
| 重送          | 名詞サ変 | 4   | 3        |
| 瞬停          | 名詞サ変 | 4   | 3        |
| 瞬停          | 名詞   | 4   | 3        |
| 消画          | 名詞サ変 | 68  | 2        |
| 色差          | 名詞   | 13  | 4        |

◆ スペリング

| 表記            | 出現数 | 出現マニュアル数 |
|---------------|-----|----------|
| %RH           | 18  | 6        |
| %rh           | 10  | 2        |
| +RW           | 8   | 2        |
| A2DP          | 3   | 2        |
| AAC           | 197 | 9        |
| AACS          | 8   | 1        |
| AC100         | 20  | 15       |
| AC100V        | 37  | 17       |
| ACK           | 20  | 3        |
| Acrobat®      | 2   | 2        |
| ActiveX       | 44  | 5        |
| ADF           | 191 | 4        |
| Adobe®        | 8   | 3        |
| Adobe®Reader® | 2   | 2        |
| AdobeRGB      | 11  | 1        |
| ADPCM         | 6   | 2        |
| ADSL          | 58  | 12       |
| AES           | 17  | 7        |
| AFS           | 26  | 1        |
| aif           | 8   | 3        |
| AIFF          | 7   | 3        |
| AMD           | 4   | 1        |
| AMEI          | 4   | 4        |
| AOSS          | 9   | 5        |
| APIPA         | 9   | 2        |
| APOP          | 5   | 3        |
| Arial         | 10  | 2        |
| ARIB          | 11  | 6        |
| Audiobooks    | 5   | 2        |
| AUTH          | 7   | 3        |
| autoexec.bat  | 2   | 2        |
| AUX           | 196 | 11       |
| AUX1          | 38  | 2        |
| AUX2          | 38  | 2        |
| AVC           | 18  | 5        |
| AVCHD         | 160 | 3        |
| AVCREC        | 176 | 1        |
| AVI           | 19  | 6        |
| AWM           | 3   | 2        |