

構文的な分かりやすさを評価する可読性評価技術

祖 国威、加納 敏行

東芝ソリューション株式会社 IT 技術研究所

{so.kokui, kano.toshiyuki}@toshiba-sol.co.jp

1. 業務文書における品質問題

1.1. 業務文書の品質

企業の活動において、大量の業務文書が作成されている。これらの文書は業務遂行を目的として作成されており、記述された内容が正確にかつ素早く伝わる必要がある。そのため、業務文書の品質を確保することは、業務の正確な遂行に必須である。

業務文書の品質は大きく分けて、内容品質(何を書くか)と表現品質(どのように書くか)からなる[1]。内容品質は、どのような事実を、どのレベルの詳細度で記述するか、という文書内容を設計する際に決まる品質である。それに対して、表現品質は、表現の正確性、整合性、可読性など、文書を執筆する時に決まる品質である。

業務文書の品質を確保するために、日々の業務ルーチンのなかで業務文書の評価し、不適切な記述があれば、適切に修正することが必要である。通常は、このような評価作業は人手で実施されているが、多大な時間を要してしまう。従って、ソフトウェアで文書品質を自動的に評価するシステムが求められている。

1.2. 業務文書の可読性問題

業務文書を読むとき、限られた時間内に文書から情報を正しく把握することが重要である。文書が読みにくくなると、貴重な業務時間を無駄にするだけでなく、意味を誤解するリスクも発生する。業務文書の可読性が与えるビジネスへの影響は主に人間向けと機械向けの 2 つの側面から考えられる。

①人間向けの文書理解

近年の企業活動の国際化により、日本語を母語としない人々が、日本語文書を読む機会が増加している(例えば、オフショア開発)。これらの読者にとって、日本人と同等な言語力と背景知識を持つわけではないので、文が分かりにくくなると、文書を理解するための時間がかかり、Q&A

が頻繁に発生する。

筆者らは、不適切な表現を自動的に発見する仕様書チェックによる中国技術者向けの文書読解支援を行った[2]。しかし、試行運用結果により、語彙レベルの文書チェックだけでは、致命的な悪文を発見することができない。例えば、例文 1 のような目的語と述語が離れすぎた文の場合、意味を理解するため、かなり時間がかかってしまう。

例文 1: 駅情報を、駅コードと業務日付をキーに、駅マスタ、駅マスタ・サマータイム情報、駅マスタ・スポット、電車運用時間情報から取得する。

一方、日本人の場合でも、分かりにくい文や曖昧な文があると、その意味を解明するため、読解時間がかかることがある。特にお客様に提出する文書の品質が悪いと、会社のイメージが低下する、あるいは内容が正しく伝わらないことによりジョブロスに繋がることもある。

従って、読者にとって分かりにくい文を事前に発見する必要がある。しかし、文書の作成者は、文書の内容が事前に分かるので、読者にとって分かり易いかどうか、客観的に評価するのは困難である。そのため、文書の可読性を客観的に評価できるツールが求められる。

②機械向けの文書処理

近年、IT の進歩につれて、大量の文書に対する自動化処理(翻訳、要約、分類、検索など)が幅広く利用されている。但し、これらの技術は、正しい(分かりやすい)文書を処理の前提とするので、文の可読性が落ちると、処理自体をいくら工夫しても、正しい結果を得られない。

例えば、機械翻訳技術を用いる多言語横断検索[3]の場合、例文 2 のような修飾関係が混乱した文を自然文検索のクエリとして用いると、正しく翻訳されないので、検索精度が落ちる。この場合、翻訳や検索結果だけ確認しても、精度が落ちる原因は発見できない。

例文 2: このような構造に発泡したプラスチックで一体に形成

もし、このような悪文を事前に発見し、それを修正することができれば、機械処理の精度は向上できるだろう。

1.3. 従来の研究

文書可読性の自動評価は、さまざまな分野で試みられている。建石らは、日本語テキストの難易度を測定するため、文字種類や文長などを用いた複数の公式を提案している[4]。また、川村は、語彙と文章の難易度との関係を分析し、日本語学習者のための教材レベル判定システムの開発を進めている[5]。さらに、柴崎らは、小学校の国語の教科書を基準テキストとして採用し、ひらがなの含有率、文の総数に対する単文数、内容語の漢語率から、テキストの読みやすさを算出する公式を提案している[6]。

これらの研究の殆どは、教科書などのテキストを選択することを目的として、1つの数値で文書全体の可読性を表すものである。業務文書の品質評価の場合、文書全体だけではなく、文単位あるいは表現単位で可読性が低い箇所を検出する必要がある。

また、従来の可読性評価は、文字種、文字数などの文字・語彙レベルの情報に基づいている。それ以上の要因(構文、文脈など)に基づく評価は、行われていない。国語教科書など一般文書を分析対象としており、業務文書への適用については考慮されていない。

本研究は、業務文書における可読性の要因を分析し、そのうち、構文的な要因に着目し、文の構造による悪文(分かりにくい文)を自動的に検出することを目的としている。

2. 業務文書の可読性要因分析

読者が文書を読むときの過程に基づき、文書が分かりやすいかどうかを判断するのが、可読性評価の最終目的である。しかし、実際には、文書が分かりやすいかどうかは読者の感覚によってかなり異なる。このようなあいまいな感覚をソフトウェアにより再現するのは、現状の技術では難しい。そのため、事前に何らかの客観的な基準で文書の可読性を評価する必要がある。

2.1. 評価基準

図1は、Kintsch [7]に示された人間の文生成と文理解

のプロセスである。文の理解は、以下の(ア)～(ウ)の3段階に分けられる。

- (ア) 表層構造の理解: 単語や句や文の記憶
- (イ) テキストベースの理解: 文章の記述内容に則して形成される意味の要約(命題抽出)
- (ウ) 状況モデルの理解: テキストから得た情報により心内に構築された状況や場面

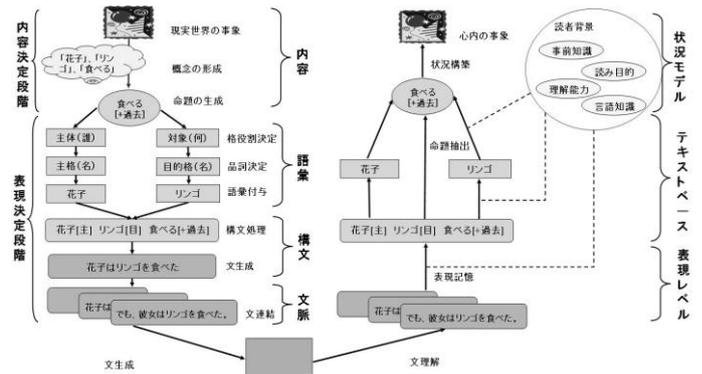


図1. 文生成と文理解のプロセス

通常、人間の文書理解は、表層構造を1個ずつ記憶し、それを1個ずつ翻訳(還元)するわけではなく、文の意味を抽出(要約)し、それを心内の知識に結び付け、状況を還元するのが基本である。特にビジネス文書の場合、文書を精読する時間ではなく、通読した上、意味を要約するのが大切である。

文の意味を表す最小単位は命題である。従って、文章の可読性を評価するとき、命題が正しく、かつ素早く抽出できるかどうか、重要な評価指標となる。

2.2. 文書要因

文理解の容易さは、読者の知識、能力、読書目的などの外部環境に関連する要因(読者要因)と文書内の記述内容に関連する要因(文書要因)に依存する。

本研究の目的は、文書の書き方を原因とする読みにくい文を発見することであるため、文書要因に着目する。

図1で示した文生成のプロセスより、命題から文書の形成するまで、語彙、構文、文脈などの複数の処理段階を経過する。従って、文書の可読性も、さまざまなレベルの要因によって決定される。

実際の業務文書に基づいて、可読性に影響する要因を調査した結果を表1に示す。

表 1. 文書の可読性要因 (文書要因)

大分類	中分類	小分類	タイプ	
文書要因	語彙的	語彙の数	長文	
		語彙の難易度	専門用語	
		語彙の多義性	多義語	
	構文的	命題要素		文に複数の命題がある(タイプ1)
				命題を構成する主格か目的格が欠落(タイプ2)
				命題の構成要素間の距離が長すぎる(タイプ3)
				命題の主語と動作間の関係が矛盾(タイプ4)
	修飾関係		修飾語と修飾先は離れすぎる(タイプ5)	
			修飾部が長すぎる(タイプ6)	
	あいまいさ		構文的あいまい(タイプ7)	
文脈的	文の接続関係		接続詞の欠落	
	文の位置		重要文が目立たない	

これらの要因のうち、語彙的要因は、従来から研究しており、本研究は、主として、構文的要因を対象として可読性評価手法を検討した。

3. 構文的要因の評価手法

構文的要因、特に表 1 で示した命題要素、修飾関係、あいまいさの 3 種類、7 タイプの要因に関する評価手法を述べる。

3.1. 命題要素要因の評価

言語の基本的単位として重視されるのは命題である。命題の構成が複雑になると、文の可読性も低くなる。

我々の考察結果、命題要素が文の可読性に影響する要因は表 1 で示した 4 タイプ (タイプ 1 - タイプ 4) がある。1.2 節で掲示した外国人に分かりにくい例 (例文 1) は、タイプ 3 による問題である。

命題は述語と対応している。述語とその項を取得できれば、命題関係を要因とする可読性を評価することができる。

本研究では、以下の 5 種類の情報を採用した。

述語の主格、述語の目的格、主格と述語間の距離、

目的格と述語の距離、主格と述語動詞の意味関係

評価するとき、まず文単位で述語を抽出し、述語毎に属性を取得する。次に、抽出された属性に対して、予め決められた基準値及び評価ルールによって、命題関係の可読性を評価する。

例えば、1 文に基準値以上の述語命題セットが存在する場合、タイプ 1 と判定する。また、主格、又は目的格が取得できない場合、タイプ 2 と判定する。主格と述語間の距離、又は目的格と述語の距離が基準値以上である場合、タイプ 3 と判定する。さらに、主格と述語動詞の意味関係が

矛盾とするとき、タイプ 4 と判定する。主格と述語動詞の意味関係が矛盾しているか否かは、主語の有情属性と動詞の意志性の対応関係によって決定する。

3.2. 修飾関係要因の評価

文の構成要素には命題要素を修飾する修飾語も重要な構成部分である。修飾関係が複雑になると、文の可読性も低下する。1.2 節の例文 2 の場合、修飾語「このような構造に」は、被修飾語「形成した」と離れすぎるため、誤って翻訳されてしまう。

本研究の場合、以下のような係り受けを修飾関係と判定する。

指示関係、名詞に修飾する形容詞と形容動詞、助詞

「の」で接続する係り受け、副詞・句、連用中止表現、連体動詞とその係り先

修飾属性は、被修飾語、修飾距離、修飾部の深さとする。修飾距離は、修飾語と修飾先のためのノード数である。修飾距離が基準値を超える場合、タイプ 5 と判定する。

また、修飾部の深さは、修飾先から構文木のリーフノードまでのノード数である。この情報によって、修飾語が長すぎる場合、タイプ 6 と判定する。

3.3. 構文あいまいさの評価

人間が文の構造を理解するとき、複数の解析候補を生成でき、候補の可能性が近いとき、理解に迷い易い。このような文はあいまいであると考えられる。

例文 3 : 科学的な眼球運動の実験調査

例文 3 の場合、「科学的な」「眼球運動」か、それとも「科学的な」「調査」か、読者の解釈が揺れる可能性がある。

このようなあいまい文については、本研究では構文解析の仮想アーク間の類似度によって判定が行える。

本研究で使っている構文解析エンジンは「最適解探索」技術に基づく解析ツールである [8]。係り受けを解析するとき、まず、係り受け候補 (仮想アーク) を生成し、係り受けの信頼度スコアによって、最適な解析結果を出力する。

この原理を利用し、出力結果と候補 (仮想アーク) の信頼度の差によって、構文のあいまいさを評価する。

図 2 は例文 3 の構文木を示している。仮想アーク「眼球運動」→「実験」と主アーク「眼球運動」→「調査」の信

頼度の差が大きいため、あいまいさは低い。それに対して、仮想アーク「科学的な」→「調査」と主アーク「科学的な」→「眼球運動」の信頼度の差が小さいため、この2つの解釈の文では、あいまいさが高いと判断できる。

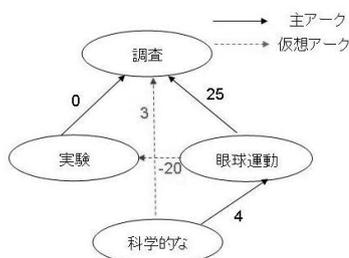


図 2. 構文のあいまいさの評価

4. 検証

今後は、本研究で提案した可読性評価技術の効果を検証するため、以下の妥当性検証及び適用効果検証を実施する。

4.1. 妥当性検証

3章で述べた構文的な可読性評価の方式について、実際のチェックシステムへ組み込み、その出力結果に対して、評価の妥当性を検証する予定である。

評価対象データは、いままでの業務において収集した分かりにくい例文や、機械翻訳で誤訳された例文である。

検証の目的は、3章で提示した3種類の構文要因による可読性問題が、正しく評価できるかどうかを確認することである。

4.2. 適用効果検証

妥当性検証は、提案した手法の精度を確認できるが、業務文書の品質向上に役に立つかどうか、1.2節で述べた課題を改善できるかどうか、について確認できない。そのため、業務現場の利用者が直接使用し、その適用効果を評価する。

但し、どのような場面に設定すれば、適用効果を発揮できるかが、また、どのような仕組みを提供すれば、利用者が評価しやすくなるかがなど、今後明確にしなければならない課題がある。

現状では、特許文献を対象とする機械翻訳の前処理、及び開発仕様書を読者に提出する前の品質検査という2つの検証場面を想定している。

5. 終わりに

本研究では、業務文書を対象として、文書の構文的な可読性を自動的に評価し、分かりにくい悪文を自動的に検出する手法を検討している。本論文では、命題構成要素、修飾関係、および構文のあいまいさの3つの問題に着目する可読性評価手法を紹介した。今後は、提案技術の精度を確認する妥当性検証、及び、実ユーザが試用してもらう適用効果検証を実施する予定がある。

今回検討した課題は、構文的要因による可読性問題だが、文と文の繋がり、文の位置を考慮する文脈的要因による可読性問題はまた着手していない。また、読者の背景や読書目的などの読者要因も考慮していない。これらの問題は今後の研究課題である。

参考文献

- [1] 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 31(7)、pp.1051-1062、1990
- [2] Guowei ZU, et al. “The Supporting Technology of Business Document Proofreading based on Intercultural Differences”, CEC’07 and EEE’07、pp.91-98、2007、Tokyo
- [3] 真鍋俊彦 “母国語で外国語の情報を収集できる言語横断検索技術”、東芝レビュー Vol.64 No.8、2009.8
- [4] 建石由佳、小野芳彦、山田尚勇 “日本文の読みやすさの評価式” 情報処理学会研究報告 1988-HI-018、pp.1-8、1988
- [5] 川村よし子 “語彙チェッカーを用いた読解テキストの分析” 早稲田大学日本語研究教育センター講座日本語教育、第34分冊、pp.1-22、1998
- [6] 柴崎秀子、沢井康孝 “国語教科書コーパスを応用した日本語リーダビリティ構築のための基礎研究”、信学技報 NLC2007-32(2007-10)、pp.19-24、2007
- [7] Kintsch, W. “Learning from text”, American Psychologist 49、pp.294-303、1994
- [8] 平川秀樹 “最適解探索に基づく日本語意味係り受け解析”、情報処理学会論文誌、43(3)、pp.696-707、2002