

## A-1

## ケモインフォマティクス法を用いる視覚障害者支援テクノロジーのための化合物名の読み上げ手法 A reading method of the chemical formula for the support technology for visual disabilities using cheminformatics method.

○上條 治夫<sup>1</sup>, 赤澤 寛行<sup>1</sup>, 平野 真央<sup>1</sup>, 深津 誠<sup>1</sup>, 小泉 公志郎<sup>2</sup>,  
豊岡 尚樹<sup>3</sup>, 梅寄 (多田) 雅人<sup>4</sup>, 守井 清吾<sup>5</sup>, 廣林 茂樹<sup>6</sup>  
\*Haruo Kamijo<sup>1</sup>, Hiroyuki Akazawa<sup>1</sup>, Mao Hirano<sup>1</sup>, Makoto Fukatsu<sup>1</sup>, Koshiro Koizumi<sup>2</sup>,  
Naoki Toyooka<sup>3</sup>, Masato Umezaki(Tada)<sup>4</sup>, Shingo Morii<sup>5</sup>, Shigeki Hirobayashi<sup>6</sup>

Abstract:As the way the visual disabilities recognize the chemical formula graphics, various tactile methods have been tried until now. Based on our experience, we investigated that listening to the compound name in document is a more effective method for the visual disabilities getting the chemical information in the chemical structure. However, such trials have never been known. This research is constituted on the cheminformatics method as fundamental technology. And it aims at base construction and establishment of basic requirements of the system by which the visual disabilities recognize the structure drawn in the chemical formula by listening to a compound name.

### 1. 概要

視覚障害者が画像である化学構造を認識する方法として、これまで様々な触読法が試みられてきた。著者は、視覚障害者が化学構造を含めた化学情報を取得するより有効な方法として、著者自身の経験から資料中の化合物名を聞き取ることを考えた。しかしながら、そのような試みはこれまで皆無である。

本研究はケモインフォマティクス法を基盤技術として、視覚障害者が化学構造式を「化合物名を聞き取る」ことによって描かれている構造を理解するシステムの基盤構築と基本要件の確立を目的としたものであり、本稿では聞き取りの前提としての読み上げについて報告する。本稿では、本システムの目的と方法について述べ、次いで、開発した基本システムの構成と実行例を示す。実行例として、化学構造式を化合物の IUPAC 名で読み上げるシステムと墨字 (点字に対して普通に書かれた文字や印刷された文字) と化学構造が混在する文章の読み上げシステムについて実験し、化学構造式の読み上げに求められる基本要件について考察した結果を報告する。

### 2. 実験方法

本研究は教科書や論文などに描かれている化学構造式を認識し、化学構造式名を墨字と併せて読み下すための手法の検討を目的としているが、そうした資料の多くは図表やイラストなども記述されているので、本稿では化学構造式と墨字だけを含む資料を作成し検討した。

LaTeX/XyMTeX または Microsoft Word/CS ChemDraw で原本を作成し、読みとり用の画像データとして、bmp ファイルおよび jpg ファイルに変換した。変換して得た化学構造式及び墨字が記述された画像を、化学構造式が存在する領域と墨字が存在する領域に分割し、抽出した化学構造式画像は ChemAxon 社の MarvinS2N を用いて mol ファイルを経由して name ファイルに変換してスクリーンリーダーに渡した。

画像の抽出からスクリーンリーダーによる読み上げまでの一連のプログラムを C# を用いて作成した。プログラム作成は守井清吾氏に準じた。

### 3. 結果と考察

筆者らは既に、LaTeX/XyMTeX [1] 法を用いて、化学構造式と墨字が混在する文章から化学構造式を抽出し、化学構造式の IUPAC 名と墨字を併せて読み下す方法を見いだして発表した [2]。

更に本稿で述べるシステムの概要は以下の通りである。

<sup>1</sup> 日大短大・教員・化学 <sup>2</sup> 日大理工・教員・一般 <sup>3</sup> 富大・教員・生融 <sup>4</sup> 富大・教員・和漢 <sup>5</sup> 富大・研究員・工 <sup>6</sup> 富大・教員・工

1. 本システムは、視覚障害を持つ高等教育における化学教育の受講者あるいは化学情報に興味を持つ人々を対象にしている。
2. 墨字と化学構造式を含む文章の一括読み上げを想定している。
3. 墨字と化学構造式が共存する文章を取り込み、墨字領域と化学構造式領域を識別し、それぞれを読み上げ可能なデータに変換した後に再結合した上で文章全体を読み上げるという一連の動作を行うことによって、使用者の化学情報獲得を補助する目的で使用される。
4. 本稿で取り扱うシステムは、墨字と化学構造式が混在する文章と、それを取り込んで読み上げるためのソフトウェアからなる。
5. 使用者は、墨字と化学構造式が混在する文章画像データをソフトウェアに入力し、ソフトウェアが、入力された画像を処理し、読み上げ音声として出力する。

上述の手法に従って処理した結果の一例を下記に示す。化学構造式と墨字が混在する資料（左図）と本プログラムを用いて抽出した化学構造式図とその IUPAC 名（右図）を示す。

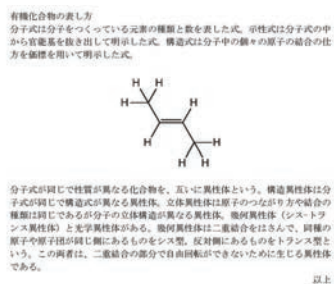


Fig. 1: Original Document

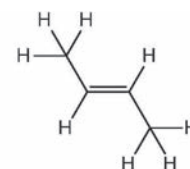


Fig. 2: Extracted. (2E)-but-2-ene

LaTeX/XyMTeX または Microsoft Word/CS ChemDraw で作成した化学構造式を含む文章では、本プログラムを用いた場合、化学構造式画像が 90 % 以上の割合で識別できたが、正しい化合物名を示した割合は約 50 % となった。これは化合物を構成する元素の誤読や結合線の数の誤読に起因していたので、引き続き誤読への対応について検討が必要であることが判った。

本稿で提案したソフトウェアを用いることで、資料中の化学構造式と墨字の文理が可能となり、それぞれを音声情報に変換し、視覚障害者に利用可能なソフトウェアとして化合物名と墨字全体の読み上げが可能であることを確認した。総じて、本研究による化合物名読み上げは、XyMTeX/LaTeX 法による構造式描画との併用により視覚障害者による化学構造式の「読み書き」を可能とするものである。

#### 4. 謝辞

本研究は平成 25 年度～平成 27 年度の日本学術振興会科学研究費助成事業（挑戦的萌芽研究、課題番号:25590293）の助成を得て進めている。

## 参考文献

- [1] 藤田眞作：湘南数理科学研究所 <http://homepage3.nifty.com/xymtex/fujitas3/xymtex/index.html>
- [2] 上條治夫、第 57 回日本大学理工学部学術講演会 2013 年 12 月 7 日