

既存の日本語読解支援システムをどう評価するか —日本語教育の視点から—

豊田悦子
メルボルン大学
etsuko@unimelb.edu.au

松下達彦
東京大学
matsushita@global.c.u-tokyo.ac.jp

1 はじめに

この論文では、日本語での情報収集を援助する読解支援システムを4つ取り上げ、それぞれのシステムの既存の機能と今後望まれる機能に焦点を当てる。まず、第一に、先行研究の知見で明らかになっている読解に必要な知識やスキルについて簡潔に論じ、第二に、選んだ4つのシステムを分析評価する。そして、最後に、改善の余地があると思われる点について論じる。評価の際には、読解支援システムが学習者の目の前にあるテキストの読解を支援するだけでなく、学習者の読解力の向上に役立つかという点も評価基準に含める。

2 読解関係の先行研究

読解過程は概して文脈情報の利用などの高次レベルの処理と語彙処理といった低次レベルの処理の2つに分けられ、その両方のスキルが不可欠であると言われている (Alderson, 2000; Koda, 1994 など)。読解過程モデルは大きく分けて、単語、文、段落、文章へと解読していくことによって読解がなされるとするボトムアップモデルと、読み手が既存の知識を用いて単語を予測、検証するといった作業を繰り返すことによって理解を図るとするトップダウンモデル、そして、その両方が不可欠であることを強調する相互交流モデルがある (館岡, 2005)。しかし、日本語のような表意文字を用いる言語の場合は低次レベルの処理の重要性がかなり高いことは否めない (Ehrich, Zhang, Mu & Ehrich, 2013)。また、日本語読解においては単語以下のレベルの処理も単語認知過程において重要な役割を果たしている (Toyoda, 2007)。

2.1. 低次レベル処理の支援

読解レベルと語彙知識の間に強い相関があること、語彙知識と単語認知スキルの間に極めて高い相関があることが多くの研究で繰り返し明らかにされている (Carver, 2003; Grabe, 2004; Horiba, 2012 など)。語彙学習研究においては、語彙知識の広さ (既知語彙量) と深さ (関連知識のレベル) や、意図的語彙学習と付随的語彙学習の両方がこのような総合的なスキルの向上に重要であることが示唆されている (Laufer, 2001; Nation, 2001 など)。

特に非漢字圏の日本語学習者にとって、複雑な表記体系を有し、二字漢語を初めとする大量の複合的な語彙を有する日本語の処理は容易ではない。日本語の語彙学習負担は一般に世界の諸言語の中で相当に重いほうであると言える。欧米語中心の研究では、語彙力が占める割合には諸説あるが、おおよそ30%程度を占めると言われる (Bernhardt, 2005) が、日本語における研究では、この数字はさらに高くなり、控えめに見ても日本語読解力の40%以上は、漢字力を含む語彙力であると考えられる (小森ほか, 2004、野口, 2008 など)。目の前のテキストの意味を示すだけでなく、語彙力向上を支援する情報を提供する必要がある。また、日本語には大量の漢字語があるため、漢字認知能力 (漢字およびその構成要素についての知識と処理スキル) もまた不可欠である (Koda, 2002; Toyoda, Arief & Kano, 2013 など)。日本語読解に必要な語彙知識の深さ (関連知識のレベル) には漢字認知能力も含まれるのである。

文法知識が読解において主要な役割を果たすことを強調する研究者もいる (Jeon & Yamashita, 2014; Jung, 2012 など)。日本語はインド・ヨーロッパ語とは文構造が異なるため、名詞などの内容語だけでなく、助詞・助動詞など

の機能語の役割もまた大きい。日本語における助詞の省略がしばしば誤解のもとになる (Lee, 2002; Nariyama, 2002 など) ことは周知の事実である。語彙知識は内容語と機能語の橋渡しをするものである。一般に語彙というと内容語を指すことが多いが、この両者を含む語彙知識を提供する必要がある。

2.2. 高次レベル処理の支援

助詞の省略と並んで読解を複雑にしているものに名詞修飾節がある。日本語の名詞修飾は英語などの関係詞節のように文法と意味を結びつける関係代名詞や人や物事を示唆する動詞の活用もなく、前後関係から意味的・語用的要因を踏まえて類推しなければならないことが多い (Matsumoto, 2007)。修飾節の動詞が省略されることもあり (Matsumoto, 2007)かなり複雑になり得るが、少なくとも被修飾名詞と修飾節の始まりを示すことで、読解の負担を軽減できる。その他、テキストの中には、特に慣用句や慣用表現を含んだものの中には、歴史的、文化的背景知識を要するものがあり、これらも読解を妨げる要因となり得る (Pulido, 2007 など)。日本、日本語に関する豆知識の提供も考慮すべきである。

3 読解支援システムの評価

この論文では、日本語学習者の中で比較的良好に使われている読解支援システムとして、リーディング・チュウ太 (<http://chuta.jp/>)、あすなろ (<https://hinoki-project.org/>)、Rikai (<http://www.rikai.com/>)、WWWJDIC (<http://www.csse.monash.edu.au/~jwb/cgi-bin/wwwjdic.cgi?1C>) のサイトの4つのサイトを取り上げ、上記の観点から分析評価を試みる。

3.1. リーディング・チュウ太

広告がないので、すっきりとした印象がある、ウェブ上で機能する読解支援システムである。日本語のテキストを指定の枠内に貼り付けてボタンを押すと、新しいページが開き、いくつかの単語がハイライトされたテキストと、その横に単語一覧表がスクロールできる形で表示される。単語一覧の情報は、振り仮名 (ひらがな)、意味 (日本語、英語、その他数か国語の中から選択) で、その他、いくつかの単語には、品詞と、その単語を含む日本語の例文と (日本語以外の場合) その訳が表示される。しかし、振り仮名や例文などを非表示にすることもできる。一つの単語に数多くの意味がある場合には2-3のコア的な意味が赤字で表示されているのでわかりやすい。しかし、助詞や助動詞のような機能語、形容詞や動詞の活用形の情報は表示されない。単語以下のレベルの情報 (漢字、構成要素など) もない。単語をクリックすると自分の単語リストを作成できる。同じ単語を何回クリックしたかわかるようになっている。単語リストにある単語をクリックすると、単語一覧表の中のその単語の読みや意味の情報が見られる。メインの辞書機能の他に、単語を入れてその例文を表示するだけの機能も別につけられている。高次レベルの情報はない。(川村 2012 に制作者自身による説明がある。)

3.2. あすなろ

多言語対応日本語読解支援システム、あすなろは、日本語共起語検索システム (なつめ) 日本語学習者作文コーパス (なたね)、日本語作文推敲支援システム (ナツメグ) と並んで用意されている。あすなろのページに行くと指定の枠内にテキストを貼付けてボタンを押すと入力された日本語テキストに対して文節の係り受け解析 (CaboCha) を行い、単語の読み (カタカナ) と品詞を表示する。さらに木構造を表示することもできる。説明や表示ボタンによれば、単語の意味を英語、中国語、タイ語、マレー語、インドネシア語などで表示する機能があり、複合名詞や慣用表現にも対応するようであるが、実際にやってみると、本稿執筆時点 (2015年1月中旬) に

おいては、どの語についても情報がありませんと表示される。機能語の情報はない。単語によっては例文は豊富に出力される。しかし、例文には訳はついておらず、長い例文の意味を理解するにはかなり高度な日本語力を要する。単語に複数の意味があるときには、前後の単語からその単語の意味を推測して、順番を並べ替える機能がついているようである。また、機能もある。単語以下のレベルの情報はない。

3.3. Rikai

英語から日本語、英語からスペイン語、中国語から英語への辞書もあるが、ここでは日本語から英語へのページに焦点を当てる。ウェブ上で機能するもの他にタブレットやスマホのアプリ版もある。メインページには様々な学習情報が広告と共に現れるので見にくいだが、説明が英語なので、英語話者には使いやすい。指示通り、指定の枠にウェブサイトのURLまたはテキストを貼ると、日本語のテキスト上にマウスを置くだけで単語の読み（ひらがな）や意味（英語）の情報が得られる。単語が漢字で構成されている場合は、それぞれの漢字とその読み（カタカナで音読み、ひらがなで訓読み）と意味（英語）も現れる。しかし、助詞や助動詞のような機能語、形容詞や動詞の活用形の情報には表示されない。漢字の情報はあるが、構成要素の情報はない。単語をダブルクリックすることによって自分の単語帳が作れる。しかし、単語の使い方を示す情報に欠けている。その他、メインページからは漢字のフラッシュカードのページに行くことができ、そこでは漢字の読み（カタカナで音読み、ひらがなで訓読み）や意味（英語）、またその漢字を含む熟語のリストをフラッシュカードの形で学習することができる。高次レベルの情報はない。

3.4. WWWJDIC

WWWJDICは多機能であることと提示する情報が多いため、使い慣れるまでに時間がかかりそうである。説明が英語なので英語話者には使いやすい。読解支援は **Text Glossing** という機能で、指定の枠内に URL またはテキストを貼り付けると、テキストを文単位に分け、文ごとに出現単語の読み（ひらがな）、品詞、意味（かなり詳細な英語での説明）、関連熟語とその読みと意味を示す。多義語の場合は、複数の意味がすべて列挙されているので、学習者を混乱させるかもしれない。複合名詞や慣用表現の説明も詳しい。すべてではないが機能語も取り上げている。形容詞や動詞の活用形はごく一部を除いてはない。枠の下に、訳を表示、または非表示にしたり、カタカナまたはひらがなの言葉を抜かしたりなどを選択できるボタンがある。**Text Glossing** の他に、単語を入力して、その単語の読み（ひらがな）、品詞、意味（かなり詳細な英語での説明）、関連熟語とその読みと意味を示す辞書機能もある。単語を入れて、その単語を含む例文のリストを得ることもできる。その他、漢字を調べる機能、漢字の構成要素から漢字を調べる機能もある。部分的情報を入力して単語や用例を探す **Advanced Search** 機能もついている。高次レベルの情報はない。

4 おわりに

以上、4つの読解支援サイトを、先行文献で述べられている読解に重要な要因に関して、分析評価を試みた。その結果明らかになったことは、すべての必要機能がそろっているものはないこと、まだまだ改善の余地があることである。特に改善が望まれる点としては、多義語の場合、どの意味が目の前のテキストの文脈に沿ったものであるかと示す、その意味を示す使われ方をしている例文を示す、助詞や助動詞などの機能語の説明を充実させる、形容詞や動詞の活用形を説明する、助詞の省略の説明を加える、構文、特に名詞修飾節を示す、文化的情報を加えるなどが挙げられる。この中には、NLP 研究で可能になったものもあるであろう。今後、言語学者、語学教育者とNLP 研究者の協力でより優れた読解支援システムが作られることを望んでいる。

参考文献

- Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press.
- Bernhardt, E. (2005). Progress and procrastination in second language reading. *Annual Review of Applied Linguistics*, 25, 133-150.
- Carver, R. (2003). The highly lawful relationship among pseudoword decoding, word identification, spelling, listening, and reading. *Scientific Studies of Reading*, 7(2), 127-154.
- Grabe, W. (2004). Research on teaching reading. *Annual Review of Applied Linguistics*, 24(1), 44-69.
- Horiba, Y. (2012). Word knowledge and its relation to text comprehension: a comparative study of Chinese -and Korean-speaking L2 learners and L1 speakers of Japanese. *The Modern Language Journal*, 96(1), 108-121.
- Jeon, E. H. & Yamashita, J. ((2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64(1), 160-212.
- Jung, J. (2012). Relative roles of grammar and vocabulary in different L2 reading tasks, *English Teaching*, 67(1), 57-77.
- Koda, K. (1994). Second language reading research: Problems and possibilities. *Applied Psycholinguistics*, 15, 1-28.
- Koda, K. (2002). Writing systems and learning to read in a second language. In W. Li, J. S. Gaffney, W.L.J.S, & Packard, J. L. (eds), *Chinese Children's Reading Acquisition* (pp. 225-248). Boston: Kluwer Academic Publishers
- Laufer, B. (2001). Word focused activities and incidental vocabulary acquisition in a second language. *Prospect*, 16, 44-54.
- Lee, K. (2002). Nominative case-marker deletion in spoken Japanese: an analysis from the perspective of information structure, *Journal of Pragmatics*, 34(6), 683-709.
- Matsumoto, Y. (2007). Interaction of multiple factors in relative clause construal and acquisition. Commentary. *Studies in Second Language Acquisition* 29. 367-374.
- Nariyama, S. (2002). The WA/GA distinction and switch-reference for ellipted subject identification in Japanese complex sentences, *Studies in Language*, 26(2), 368-431.
- Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Pulido, D. (2007). The effects of topic familiarity and passage sight vocabulary on L2 lexical inferencing and retention through reading. *Applied Linguistics*, 28(1), 66-86.
- Toyoda, E. (2007). Enhancing autonomous L2 vocabulary learning focusing on the development of word-level processing skills. *The Reading Matrix*, 7(3), 13-34.
- Toyoda, E, Arief, M. F. & Kano, C. (2013). Identifying useful phonetic components of kanji for learners of Japanese, *Japanese Language and Literature*, 47(2), 235-272.
- 小森和子・三國純子・近藤安月子 (2004). 文章理解を促進する語彙知識の量的側面 —既知語率の閾値探索の試み—。 *日本語教育*, 125, 83-92.
- 川村よし子 (2012). 日本語読解学習システム『リーディング・チュー太』の取り組み. *留学交流*, 2012年10月号 Vol.19. URL: <http://www.jasso.go.jp/about/documents/kawamurayoshiko.pdf>
- 館岡洋子 (2005). ひとりで読むことからピア・リーディングへ：日本語学習者の読解過程と対話的協働学習. 東海出版会.
- 野口裕之 (2008). 試験結果の分析. 国際交流基金・日本国際教育支援協会編, 平成17年度日本語能力試験 分析評価に関する報告書, 45-111. 凡人社.