

インターネット利用による電子英文情報の即時教材化システムの開発と その教育利用

大武博*1・河本健*2・竹腰正隆*3・国村正子*4・Brian Morren*5・
竹内浩昭*6・鶴川義弘*7・藤田信之*8・金子周司*9

*1 京都府立医科大学花園学舎第一外国語教室 (ohtake@koto.kpu-m.ac.jp)・

*2 広島大学大学院医歯薬学総合研究科探索医科学講座口腔生化学 (tkawamoto@hiroshima-u.ac.jp)・

*3 東海大学医学部分子生命科学2 (mtakekos@is.icc.u-tokai.ac.jp)・

*4 福井工業高等専門学校(非常勤) (msong@po.mitene.or.jp)・

*5 福井県立大学学術情報センター (morren@fpu.ac.jp)・

*6 静岡大学・理学部・生物地球環境科学 (sbhtake@ipc.shizuoka.ac.jp)・

*7 宮城教育大学附属環境教育実践研究センター (ugawa@ipc.miyakyo-u.ac.jp)・

*8 製品評価技術基盤機構バイオ本部ゲノム情報解析課 (fujitan@nifty.com)・

*9 京都大学大学院薬学研究科医療薬理学分野 (skaneko@pharm.kyoto-u.ac.jp)

概要

タイム誌・ニューズウィーク誌・BBC など英米の主要な報道機関がインターネット上で公開している英文記事・文書などは、格好の生きた英語教材としての可能性を秘めている。しかしそれらの記事を直接教材として利用するには、難解な語句に対する注釈の準備や著作権の問題など、解決しなければならない問題がある。本研究では、インターネット上で入手可能な英文リソースの即時教材化を実現するために、電子辞書の作成と単語注釈付記の自動化システムの構築及び著作権問題回避の方策を試みた。英語教師が利用対象とする英文記事を、その記事専用の「英和単語帳・解説」とともに学生に提示可能な「オンデマンド英語教材」専用サーバを試作した。また、記事を教材利用する場合に直面する著作権の問題については、利用者各個人が記事に直接アクセスし入手するシステムを導入することで、記事のサーバへの保存を回避し、著作権問題に抵触しない工夫を試みた。

キーワード：

英語教育・教材・オンデマンド・インターネット・電子辞書

Development of an Automatic Gloss-embedding System – Converting Internet English resources to teaching materials on demand –

Hiroshi OHTAKE, Kyoto Prefectural University of Medicine (ohtake@koto.kpu-m.ac.jp)

Takeshi KAWAMOTO, Hiroshima University Graduate School of Biomedical Sciences
(tkawamoto@hiroshima-u.ac.jp)

Masataka TAKEKOSHI, Tokai University School of Medicine (mtakekos@is.icc.u-tokai.ac.jp)

Masako KUNIMURA, Fukui National College of Technology (msong@po.mitene.or.jp)

Brian MORREN, Fukui Prefectural University (morren@fpu.ac.jp)

Hiroaki TAKEUCHI, Shizuoka University (sbhtake@ipc.shizuoka.ac.jp)

Yoshihiro UGAWA, Miyagi University of Education (ugawa@ipc.miyakyo-u.ac.jp)

Nobuyuki FUJITA, National Institute of Technology and Evaluation (fujitan@nifty.com)

Shuji KANEKO, Graduate School of Pharmaceutical Sciences Kyoto University
(skaneko@pharm.kyoto-u.ac.jp)

Abstract

We have created a genre-specific electronic English-Japanese dictionary and an automatic gloss-embedding system for electronic English texts in a specific discipline (life sciences and related issues). This particular dictionary and the automatic gloss-embedding system has made it possible for English resources on the Internet to be instantaneously converted to a hyper-text type electronic file. In this way, we hope that authentic English materials that are widely available on the Internet may be readily adopted for the purpose of learning/teaching English. The provision of a gloss for a particular text is expected to greatly reduce the burden on learners arising from their vocabulary deficiencies and encourage them to read relatively difficult English materials on their own.

Key Words:

English Teaching, Teaching Materials, On-Demand, Internet, Electronic Dictionaries

1 はじめに

大学の英語教育用教材として、タイム・ニューズウィーク・BBCなど英米の主要な雑誌・新聞などの記事を「生きた教材」として利用している指導者は少なくないと思われる。時事英語を題材とした教科書の種類・数の増加や、これらの題材を使用した講義を選択受講する学生の数が多いことから、時事英語の人気は根強いように思われる。しかしながら、それらの記事を脚注・解説付の教材として利用する場合は、従来型の方式に従えば、印刷媒体で教科書として供給されているものを使用することになる。この方式では、これらの教材は、記事の発生から半年から1年遅れで教室に届くのが現状である。即時性を旨とする時事教材が、「生きた教材」でありながら、内容的には全く新鮮味を失った「生きていない素材」になってしまうというジレンマを多くの教授者は感じている。

情報通信分野の革命的な転換期を迎え、大学の英語教育も新しい時代に相応しい新たな現代化の方策が誕生してしかるべき時期にきている。インターネットを利用して、世界中の情報を即時に入手することが可能になった今、それらの英文情報を、的確に入手し処理（理解）する能力を育てることを、大学英語教育の目標の一つとして掲げることは、時代の要請に叶っていることであろう。実際、すでに英米の主要な新聞・雑誌あるいは小説でさえ、インターネットを通じて、電子媒体として購読できる時代である。これらの電子情報を大学の英語教育の中で教材として即時利用する方策を確立することで、従来の印刷媒体による教科書利用とは違った新しい英語教育の局面を開くことができる。

インターネットを介して電子媒体で供給される英文情報を、画面上でそのまま日本語注釈付き教材として利用できるシステムを構築公開することで、大学英語教育の現代化の扉を開きたいと願っている。

2 オンデマンド英語教材作成システムの構築と公開

2.1 分野別基本電子辞書の作成

インターネットなどを介して電子媒体として入手できる記事などを教材利用する際には、難解な語句に注釈を付記するなどして、学習者の情意的抵抗を低くする工夫が必要である。英単語の逐語訳をハイパーテキスト形式で埋め込むために、汎用辞書をジャンル別に3種類作成した。

2.1.1 ライフサイエンス辞書

筆者らは、電子媒体で供給される英文情報には、英語教育での利用を含めそのほか多方面での利用の可能性が秘められていることに、早くから着目していた。そこで、ライフサイエンス電子辞書作成プロジェクト（Life Science Dictionary Project = LSD プロジェクト）を立ち上げ、生命科学分野に特化した英和・和英辞書の作成に着手した。サイエンス・ネイチャーに代表される世界的に著名な欧米の雑誌より、約1万件的英文抄録を収集し、約3千万語から成る英文コーパス（LSD コーパス）を作成した。

LSD コーパスを基本資料とし、単語の出現頻度など計量的分析により、必要度の高いと思われる語を約2万語抽出し、電子英和・和英辞書「ライフサイエンス辞書 = LSD 辞書」の作成を行い、インターネット上で一般に公開し無償供与している。生命科学に関連する学術雑誌に掲載される英語論文を読む際には、有益な辞書である。

2.1.2 時事英語辞書

NHK は長年、海外向け英文ニュース記事を、国内では文字放送として放映している。インターネットで世界中の新聞・雑誌の英語ニュースが簡単に入手できることが当たり前になる以前には、電子媒体とし

て入手・保存できる貴重な英語情報源であった。また、VOA(Voice of America)も、インターネットがまだ揺籃期の時代からニュース記事などをインターネット上で公開し利用者の便宜を図ってきた。これらの記事を数年分蓄積し、頻度上位の単語約5千語を抽出し、「時事辞書」を作成した。LSD辞書がライフサイエンスの専門的な英文に特化しているのとは対照的に、この辞書は政治・経済・社会・国際問題などを扱った英文記事を対象にした辞書である。

2.1.3 ライフサイエンス一般辞書

LSDプロジェクトでは、2000年6月にReading Assistant Project (LSD-RAP) チームを編成し、ホームページにて週1回のペースで、生命科学に関連した一般記事を選び、その専用単語帳を作成・公開している。これは、オンデマンド英語教材システムを利用して、実際に英語学習を促進するのを補佐することを狙うと共に、システムの実効性を具体的にユーザに理解してもらうことを目指している。開設以来取り上げた記事の数は、本原稿執筆時点(2004年10月)で、200近くになる。開設当初はタイム誌の記事を中心に取り上げ、その後記事の有料化の波を受けて、無料利用できる英米の主要紙の記事へとシフトしてきている。このプロジェクトでは、生命科学・生命倫理などに関連した記事を多く扱い紹介していることから、これらの記事からなる英文コーパスを作成し計量的分析などにより頻度上位語を抽出し辞書登録対象とし、さらに熟語表現を含め総数約1万件から成る「ライフサイエンス一般辞書」を作成した。「ライフサイエンス辞書」が、学術誌の英語論文を対象にしている一方で、こちらは、欧米の主要な雑誌・新聞などに掲載される一般向けの記事を対象にした辞書である。

2.2 記事の自動教材化システムの開発

上記で作成した電子辞書を利用して、インターネット上で入手できる記事を、ハイパーテキスト形式に変換して出力するツールを開発し、オンデマンド教材化システムを構築した。

2.2.1 自動脚注付加ツールの開発

教材対象に選んだ特定記事(図1)について、出現単語を逐一、上記で作成した基本電子辞書(「ライフサイエンス一般辞書」)に登録済みの単語とマッチング検索を行い、登録語については自動的に単語訳をHTML形式でオリジナル記事に情報を埋めこむ脚注付加専用ツール(EtoJ Vocabulary)を開発した。これにより、学習者は必要に応じて、オリジナル記事中の単語をクリックすることで、その対訳を脚注窓で確認することが可能になる(図2)。

2.2.2 自動単語帳生成ツールの開発

学習者が英文読解中に遭遇する大きな問題のひとつに未知語との遭遇がある。しかしながら、学習者の語彙レベルは一定ではなく、どんな教材を利用するにしても、語句の注釈をどの程度付記するかは常に難しい問題である。上記EtoJ Vocabularyにより機械的に自動生成された脚注(図3)は、あくまでも基本辞書に登録済みの訳語を採用したもので、文脈上相応しくない訳語が混入することを完全に回避するものではない。特定の記事を教材として利用する際には、不適切な訳語や解説の混入は避けたいものである。そこで、脚注に付記された単語と訳語の一覧からなる単語帳を、編集用に別ファイルとして自動生成するツールを開発した。教授者は、事前にこの単語帳の訳語の妥当性をチェックした上で不備を修正し、また必要な解説を追記することで、特定記事の専用単語帳(図4)に仕上げるができることになる。

For the first time, a treatment has delayed the development of Alzheimer's disease in people with mild cognitive impairment, a form of memory loss that often precedes dementia, scientists from the Mayo Clinic reported yesterday.

Individuals taking the drug donepezil developed Alzheimer's an average of six months later than those with the same symptoms who took a placebo, according to a study presented at the Ninth International Conference on Alzheimer's Disease in Philadelphia. Donepezil, also known as Aricept, is approved by the Food and Drug Administration for treating symptoms of Alzheimer's disease.

"This is the first treatment to demonstrate any effect on mild cognitive impairment," said Zaven Khachaturian, senior science adviser to the Alzheimer's Association, which sponsored the conference but was not involved in the research. He said the results are likely to spur the search for other early treatments that can slow the brain degeneration that is the hallmark of Alzheimer's.

(図1：記事オリジナル。Boston Globe紙 'Drug found to delay Alzheimer's development' July 19, 2004 原文の一部。)

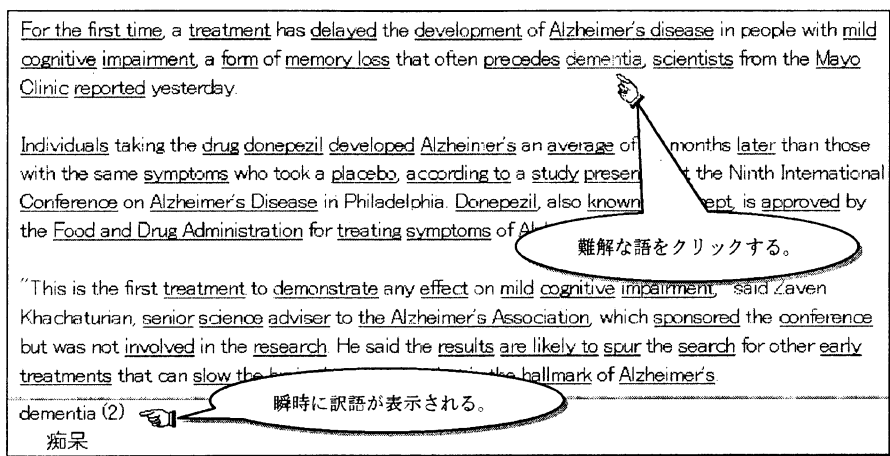
2.3 教材オンデマンドシステムの雛形運用

英語教材オンデマンドシステムの具体像を提供するため、LSDプロジェクトでは専用ホームページを立ち上げ一般公開をしている (<http://lsd-rap.pharm.kyoto-u.ac.jp>)。インターネット上で公開される生命科学関連の英文記事を中心に取り上げ、毎週1記事を目処にその特定記事専用単語帳の作成・確認・修正・追記の作業を行い、大学など教育機関での教材利用と個人学習者の利用に便宜を図っている。ホームページを公開して以来、3年を経過した現時点で、利用可能記事としてタイトルリストで紹介してきた数は190篇を超えている(図5)。

2.4 著作権への配慮

インターネット上で公開されているリソースとはいえ、著作権で保護されている電子情報を教材利用する場合、それを侵害しない配慮が必要である。この見地から、開発したサーバシステムにおいては、対象記事そのものをサーバに取得保存することをせず、利用者の責任において直接公開されているサイトから記事を取得してから利用する方策を講じた。このシステムにより、利用者は、自己の責任において、提供される対象記事を取得し、本サーバにおいて脚注付記事に変換することになり、個々人が著作権を了解の上利用することで、オリジナル記事の著作権を守りつつ教材利用することが可能になった。具体的には、利用者がオリジナル記事の著作権を了解の上で記事を利用することに同意する確認画面を、脚注付画面に移行する直前に挿入することとした(図6)。

ただし、利便性に配慮して、対象記事を開いた後、利用者側でその記事をコピー・ペーストし本システムに取り込むという一連の複雑な作業をサーバが代行する仕組みになっており、利用者は「同意する」とうボタンを押すことで即脚注付画面に移行できる。



(図2：学習者に提示される記事画面。上の窓枠に原文、下の窓枠に単語訳。EtoJ Vocabulary 利用による元記事変換後画面)

2.5 システムの汎用性

本プロジェクトで立ち上げたホームページでは、インターネットを介して入手できる多量の英文電子情報の中から、生命科学関連の記事を中心に選択・採択し、英語読解の一助とするべく、専用単語帳を独自に作成公開し、オンライン上での英文電子情報の教材化の具体例を提示している。しかしながら、開発したシステム自体は、汎用性が高く、本プロジェクトで公開している電子辞書（ライフサイエンス辞書・時事英語辞書・ライフサイエンス一般辞書）以外の辞書を利用したり、あるいは独自に辞書・単語帳を作成することで、ほぼあらゆる英文電子情報について、HTML形式で語句の注を埋めこんだ形態で、オリジナルの英文電子情報を提示することを可能にしている。

ライフサイエンスに関連した記事であれば、本プロジェクトが既に作成し公開している辞書を利用して、任意の記事の URL を入力するだけで即時に簡易教材化が出来るようホームページ上で入力画面を設定している（図7）。ただし、この場合はあくまでも、これらの電子辞書に登録されている単語について自動的に訳語をHTML形式でオリジナル記事に付記するだけであるので、文脈上相応しくない訳語が提示される場合があることを利用者側で承知しておく必要がある。教材利用を想定しているのであれば、機械的に生成された単語帳を確認し、修正・加筆する作業が必要となる。

また、利用者が独自に特定の英文電子情報に特化した専用単語帳を作成し教材利用できるように、単語帳作成の際の基本的な規則に関する解説を公開し、積極的な利用を促進している（図8）。

自分自身の教育理念に叶った教材を利用したいと願っている英語教師は少なくないと思われる。教材用の望ましい素材が電子媒体で入手できれば、本システムを利用することで、オリジナル素材を脚注・解説付英語教材へと変換することが可能になる。

2.6 システムの教育的評価

本システムにより供給される英語教材が、従来型の印刷媒体の教科書で提示される英語教材と大きく異なる点は、語・句などの注が画一的に学習者の目に触れる訳ではないという点である。語彙力は個人差が

```

# dic=http://lsd-rap.pharm.kyoto-u.ac.jp/dic/010612_postPlacebo_gloss.txt
# title=Reassessing Our Investment in the Placebo Effect (偽薬効果の真偽)
Washington_Post_010612
# Copyright 2002 Lsd Reading Assistant Project. All rights reserved.

Alternative Medicine 代替医学・医療
analyze              分析する
anyway              とにかく
article             記事
aspirin             アスピリン
assistant           アシスタント, 助手
asthma              (病名) 喘息
bear                (株式で) 弱気筋
belief              信念
bottom              底 (に達する)
bull                (株式で) 強気筋

```

(図3：自動単語帳生成ツールにより作成された記事専用単語帳)

```

###Medical bears      医学分野の弱気筋
###ended up          結果--になる
###belief in         ---
feel-good bonus      甲
bulls note           強気筋のノ

###Now the bottom has fallen out of the placebo market
e n t (投資) という表現が使われていることから類推できるように、株式市場の動向に例えている。「偽薬の効果に関する評価は、底値が落ちてしまった」くらいの意味か。
###There's a lot of evidence that points in the other direction
「そうではない方向を指し示す多くの証拠がある」文脈から、「プラシーボ効果があるとする証拠が沢山ある」と読める。
###not that the placebo doesn't do anything, but that it doesn't do everything
not A but B の骨組みで、「AではなくBである」。さらに、前半の文章が全部否定、後半が部分否定だから、「偽薬が全く何もしないというのではなく、偽薬が何もかもするわけではない=偽薬が全く効果が無いというのではなく、偽薬が必ずしも例外無く効くわけではない」くらいの意味。

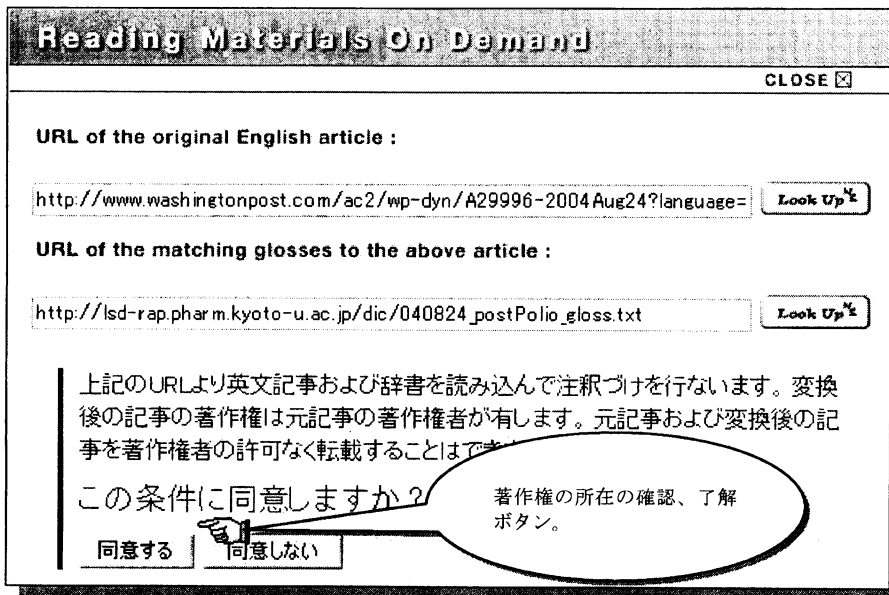
```

単語訳だけでなく、文脈の説明など長い解説を単語帳に追記することも可能。

(図4：上記単語帳の訳語を吟味した上で更に必要な解説を追記)

大きく、どの程度語・句の注を付記するかは、教材を作成する場合、常に悩ましい問題である。今回開発したシステムの利点は、単語帳の全容を画一的に提示するのではなく、学習者の必要に応じて、必要な語・句の注や解説のみが提示される点である。単語帳を作成する場合に多少の配慮をするだけで、広範囲の語彙レベルの学習者に利用可能な脚注を付記することが可能である。

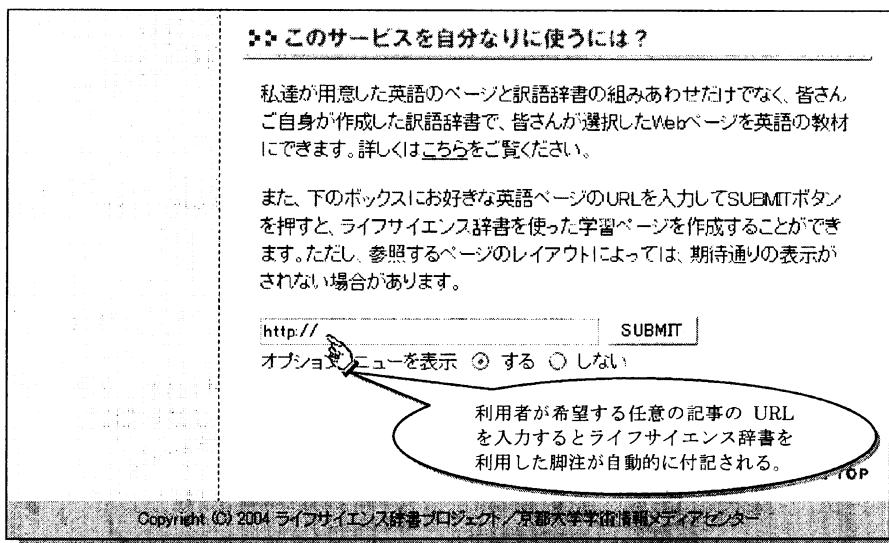
このシステムにより作成された教材を、大学教養過程の英語授業で実際に利用して学生の評価を調査した。「利便性」「市販辞書との利用頻度の比較」という2項目において、5段階評価で評定4以上を選び、本システムを積極的に支持した受講生は、いずれの項目においても80%強であった。自由記述欄の感想の中には、「このシステムがなければ、記事を読む気が起こらなかっただろう」と具体的に、単語脚注があるおかげで、多数の未知語が含まれる記事を読む場合に生じる抵抗感が軽減されたことについて言及した回答があった。学習者の情意面で、本システムが良い影響を与えていることを窺い知ることができる。ただ、画面上の細かい文字を集中して読むことを要求されることになり、「目が疲れる」と印刷媒体の教科書の利点に言及する回答も含まれていた。



(図6：著作権を了解の上自己責任でオリジナル記事を取得・利用することに関して同意を確認する画面)

参考文献

- [1] Hiroshi Ohtake, Takeshi Kawamoto, Masataka Takekoshi, Masako Kunimura, Brian Morren, Hiroaki Takeuchi, Yoshihiro Ugawa, Nobuyuki Fujita, Shuji Kaneko. Development of a genre-specific electronic dictionary and automatic gloss-embedding system — Converting Internet English resources to teaching materials on demand — . 445-449. Dictionaries and Language Learning: How Can Dictionaries Help Human & Machine Learning (ASIALEX '03 Tokyo Proceedings) 2003.
- [2] Shuji Kaneko, Nobuyuki Fujita, Yoshihiro Ugawa, Takeshi Kawamoto, Hiroaki Takeuchi, Masataka Takekoshi, Hiroshi Ohtake. Life Science Dictionary: A Versatile Electronic Database of Medical and Biological Terms. 434-439. Dictionaries and Language Learning: How can Dictionaries Help Human & Machine Learning (ASIALEX '03 Tokyo Proceedings) 2003.
- [3] 鶴川義弘、大武博、河本健、藤田信之、竹腰正隆、金子周司. インターネット情報の英語教材化の試み. 教育システム情報学会第26回全国大会講演論文集. 325-326、2001
- [4] 大武博、鶴川義弘、久保田章、河本健、藤田信之. 大学英語教育の現代化に向けて、インターネット利用による電子英文情報の即時教材化システムの開発. 電気通信財団研究調査報告書 No.17 166-171、2002.



(図7：任意の記事を脚注付英文に変換するための URL 入力画面)

辞書の作り方

辞書としてはタブ区切り形式のテキストファイルを準備します。各行はタブで区切られた2つもしくは3つのフィールドで構成されます。

第1フィールド	英語見出し	単語だけでなく、2語以上からなる熟語や用法を見出し語にすることもできます。基本的に、文章中に出てくる通りに記述する必要があります。
第2フィールド	日本語訳	単純な訳だけでなく、用法の解説や背景説明などを自由に記載することができます。
第3フィールド	注(オプション)	変更履歴などのメモを記載することができます。

以下の行は特別な働きをします。

- #で始まる行(ファイルの先頭)
注釈と解釈されます。辞書の版權等を記載しておくといいでしよう。
- #で始まる行(ファイルの途中)
英語見出しの先頭に#がついている項目は、文章中で赤色のリンクで表示されます。重要な単語や用法を強調表示したい時に使用します。
- #note#行以下(ファイルの最後)
#note#とだけ書かれた行があると、それ以降は解説文と解釈されます。訳語欄に書ききれないような長い文章を別ウィンドウに表示させることができます。

(図8：利用者が任意の英文電子情報に独自の自作辞書を作成し教材利用するための単語帳作成手順解説)