

以下のスライドをご覧になる方は、先に以下のビデオをご覧ください。

国立国語研究所「講義「[どうすれば語彙学習負担を減らせるかーコーパスとテストから考える語彙の学習と教育ー](#)」（松下達彦）／言語学レクチャーシリーズ Vol.25」

以下では、このビデオをご覧になった方からの質問に回答しています。

なお、類似のご質問がたくさんありましたので、その場合はそのうちの一つの質問をお借りして回答しています。お手数ですが、ご自身のの質問に似た質問を探してください。ほぼすべてのご質問に回答したつもりです。

（松下 達彦、2024年2月4日）

Q. (1. 「語彙力はどれほど重要か」に関して) 「語彙力が読解力の分散をどの程度説明するか」の「分散」の意味と、その結果 ($r = \dots$) の読み解き方。

→分散はそれぞれの値の平均からの距離の二乗の平均。それを平方根にしたのが標準偏差。どちらもデータの散らばり具合を示す。語彙力が x で、読解力が y とした場合に直線に近い散らばり方をしているほど r (の絶対値) も r^2 も大きくなるはず。

→相関は双方向の関係を示すので、因果関係は説明できない。ある要因から別の要因を説明しようとするのが回帰分析。回帰分析では r^2 は決定係数と言い、2変数 x , y 間では x から y を予測する力を示す。数値が大きいほど予測力は大きい。

・相関係数 r は $-1 \sim 1$ に分布。決定係数 r^2 は二乗なので $0 \sim 1$ に分布。(決定係数を証明した人がいるらしい。数式を立てて証明していったところ、結果としてたまたま r^2 になったということらしい。)

参考サイト: 「相関係数 r 」や「決定係数 r^2 」について調べてみてください。

Q. 「3. 日本語学習者はどれだけの語を学習するのか」について、7:59のスライドの「初中級語彙を含めて」は理解できたのですが、次のスライドで中級と上級の数字が同じだったこともあり、その表す意味が理解できませんでした。

→この部分のレクチャービデオをぜひ聞きなおしてほしい。ポイントは、先に上級の定義を理解することが肝心だということ。初級と上級の定義を理解できれば、初級が終わって上級に達するまでが中級だということになる。中級の定義の最後が「～まで」で終わっており、上級の定義の最後が「～以上」となっていることに注意。

Q. 「4. 1年にどれだけの語を習得できるか」について→個人差も大きくよくわかっていないとのことですが、妥当な日本語教育を計画するために重要なポイントだと思います。私は日本語学校で高等教育機関への進学を志望する留学生に日本語を教えており、「大学等の教育機関で専門的な内容を学びながら身に付けていくレベル」（『「日本語教育の参照枠」の活用のための手引』p.7）の一つ下であるB2レベル辺りに到達目標を置くことが妥当なのだろうと思います。これは松下先生のご講義ではJLPT中級（N3, N2）6,000語、松下（2016）中級7,000語の手前（という言い方でいいでしょうか…？）に当たるレベルだと思います。例えば、日本語学校の2年の課程で、日本語未習者を受け入れて年3,000語超の学習を計画する（鬼のよう？）、2,000～3,000語を習得済みの初級既習者を受け入れて年2,000語程度の学習を計画する（これでも結構大変？）、…といったことをちゃんと考えて教育課程編成と入学者選考をしなければならないなと思いました。現状、こういった語彙の教育／学習の計画性についての共通認識が日本語教育関係者の間で十分に確立していないように感じているのですが、どうなのでしょう？

→確立はしていないと思うが、少しずつ広まっている感じはある。年間習得語数については、よくわからないが、非漢字圏の学習者について言えば、やはり高頻度漢字を含む高頻度語彙に早く習熟することが肝要。そこで「雪だるまの芯」ができれば中級で語彙習得が加速するように見える（[松下ほか, 2021](#)）。

→テスト結果を見ると、教科書でやったことしか身につけていない学習者と教室外でいろいろな語彙を拾っている学習者の違いを感じる。後者を意図的に促進するには、ストラテジー教育、多読の推進（→「[読み物いっぱい](#)」、Nation & Waring (Mitsue Tabata-Sandom ほか訳) 『[言語教育における多読](#)』くろしお出版）、テレビやネットを使った効果的学習方法の伝達、学習者コミュニティでのネットワーク化の促進などが有効ではないか。

Q. 1年の語彙習得数からその言語を把握する所要時間を推測できるといったが、それは個人差が大きいだろう。どのぐらいの単位の時間で測定するのが良いか、一週間の習得数か、あるいは1時間ほどの習得数で推測か。

→週か月ぐらいが良いのではないか。

→教室内のタスクで30分、1時間という単位で語彙習得を計算した研究もある (Jonathan Newton)

Q. 学校によって使用する教科書は様々ですので、「既知語の割合とテキスト理解の関係」について、もっと詳しい情報をいただきたいです。

Q. 語の頻度レベルと漢字の頻度レベルの対応のさせ方、同一のデータベースに基づいて、語彙頻度プロファイラーでテキストの語彙的難度を測定すること、および学習者のレベルを語彙テストで測定することの意味はよく理解できるのですが、具体的にはどのようなプロファイラー・テストを用いるのかを教えてください。

→まず、語彙テストでどの程度の語彙を習得しているのかをVSTRJ-15Kなどで測定して、どのレベルの語彙を何割ぐらい理解できるかを測定し、そのうえで、使っている教科書のテキストの語彙レベルをJ-LEXなどで測定すれば、どの程度の文章が、どの程度の語彙力の学習者グループに適合するかが、だんだんわかってくると思います。

VSTRJ-15K <http://www17408ui.sakura.ne.jp/tatsum/webtest.html#vstrj>
KCT <http://www17408ui.sakura.ne.jp/tatsum/webtest.html#KCT>

J-LEX <http://www17408ui.sakura.ne.jp/tatsum/webtools.html#jlex>

Q. 「7. 一定の既知語に達するのに、漢字はどのくらい必要か」について、「語のテキストカバー率」の意味自体が理解しにくいです。母数は何ですか。

→ 「テキストカバー率」は「ある語（あるいは語彙＝語の集まり）があるテキスト（文章）の延べ語数に占める割合」。母数は「あるテキスト（文章）の延べ語数」（＝テキスト長）ということになる。

「5. 既知語の割合とテキスト理解はどう関係しているか」で説明しているので、見直してみてください。

Q. 「9. 漢字学習と語彙学習をどう関連付けたらよいか」の中で、漢字学習と語構成の学習をどう関連付けていいのかについて、もう少し詳しく説明していただければと思います。そして、透明度という概念をどう使って漢字学習に役に立つのかも伺いたいです。

→ビデオの関連部分をもう一度聞きなおして考えてほしいが、語彙論の教科書などに出ている語構成のパターンを中級前半あたりで学習すれば、役立つかもしれない。「読書」（動詞＋目的語）、「白書」（形容詞＋被修飾名詞）、「夫妻」（並列）といったこと。ただし、例外や複数のルールが当てはまるケースもあるので、パターンを学習したうえで、学習者が新しい語に接したときにどのパターンなのか、確認しながら学習するような意識付けもあるとよいのではないか。例えば「魚肉」は「魚の肉」であり「魚と肉」ではない。このようなケースは難しい。

透明度については本多由美子（2022）『二字漢語の透明性と日本語教育への応用』（くろしお出版）が良い参考書。日本語教育への応用についての提案もある。ただ、上記の語構成の場合と同じで、透明度の高い（意味推測しやすい）語もあれば、そうでないものもあることには注意する必要あり。

いずれのケースについても言えることだが、優れた学習者は、ルールを使える場合はそれは採用して処理を効率化し、同じルールを使えない場合にはルールを無視する（脳内で抑制する）といったことができる（松下ほか，2004）。どのような場合にルールを使い、どのような場合に使わないかを考え、それをスイッチできるように仕向けることができればよい。

[松下達彦・Marcus Taft・玉岡賀津雄（2004）「中国語「単語」を知っていることは日本語漢字語の発音学習に役立つか？」](#) 記念論文集編集委員会編『平井勝利教授退官記念 中国学・日本語学論文集』白帝社、pp. 578-590

Q. 「9. 漢字学習と語彙学習をどう関連づけたらよいか」のついて、特に「漢字語の比喻を理解する」について、漢字が多義であることと、語の比喻との関係をもう少し詳しく教えていただけないでしょうか。例えば、「明白」の場合、「明」には「明らか」という訓もあります。

→ 「明らか」という語があり、そこに「明」という字があてられていること自体に比喻が関わっている。これはメタファーによる語義の拡張という現象で、このような語はたくさんある。つまり多義派生自体に比喻が関わっているということ。これについては「認知言語学」の分野で多くの研究があるので、調べてみてほしい。比喻に関する基本的な理解には、以下の本がわかりやすい。

瀬戸賢一（1995）『メタファー思考』講談社現代新書

多義派生については、糸山洋介の著作（認知言語学関連の入門書がわかりやすいであろう）。

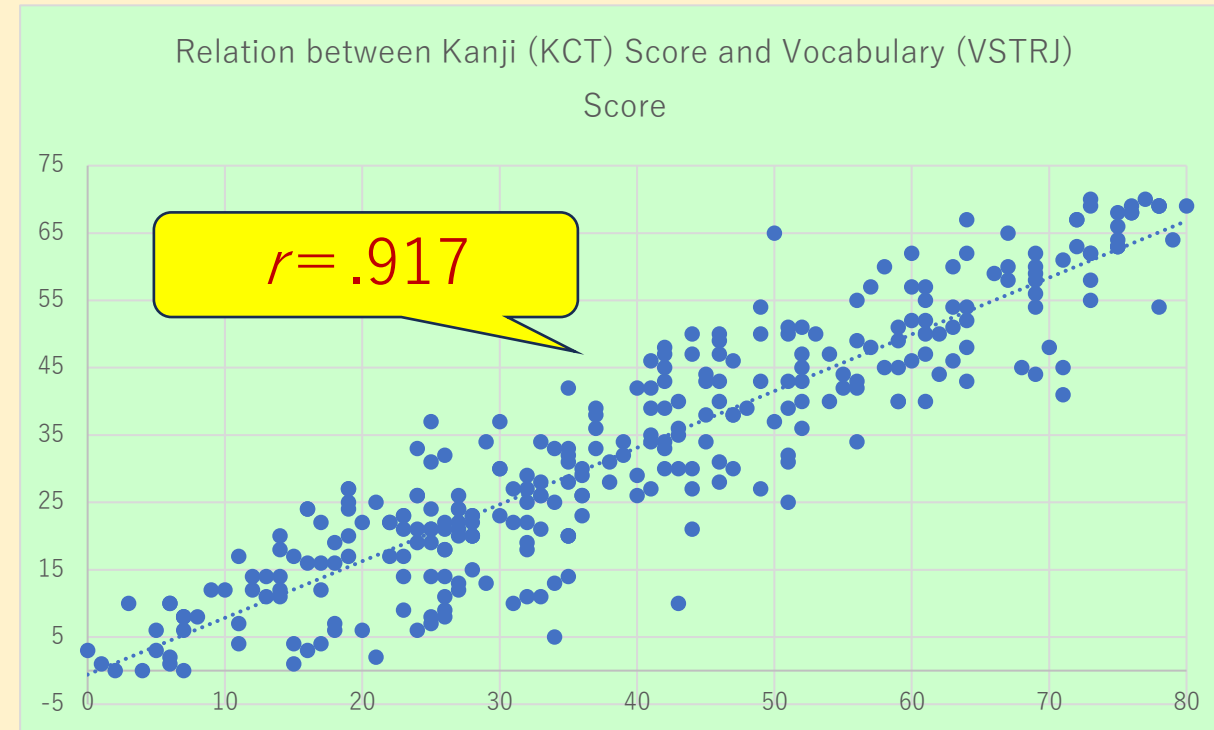
Q. 「9. 漢字学習と語彙学習をどう関連付けたらよいか」の中で、語の頻度レベルと漢字の頻度レベルを対応させておくといいという提案があったが、具体的にはどのようにすればよいか。複数のデータベースがあるが、それらをどう使えばよいか。

→KCTとVSTRJのテスト結果から単回帰分析を行って、漢字量（変換力）と語彙量の対応表を作ってみた（未発表の研究成果ですので、ここでは掲載を控えます）。直観と大きくは違わないが、問題はどうか。

ルビ振りなどの場合は少し「保守的に」、つまり少し学習者の知識を低めに見積もって考えるほうが良い。
（次ページ参照：『語彙ドン』などでのルビ振り基準）

コーパスで考えると

- ・ 漢字1,000字程度でテキストカバー率は95%になる
- ・ 語彙9,500語程度で95%になる
- ・ では、学習の実態で見れば、漢字1,000字を習得した時点で語彙数9,500語になり、テキストカバー率が95%になるかという、そうではない。
- ・ （非漢字圏学習者の場合）漢字数1,500ぐらいになって初めて1万語程度の語彙量になり、上級に近づくと考えられる。
- ・ 漢字が125字増えるごとに842.3語増加＝漢字1字増えるあいだに増える平均語数は6.7語



ルビ振り基準（案）（2019年3月21日）

- ＊すべての場合で、同一文章内では初出のみにルビ。
- ＊すべて「文字単位」でルビ
- ＊カタカナにはルビ不要
- ＊送り仮名にも不要

【別表】上記までを満たしたうえで、中級以上の語のうち、以下の漢字と語彙の基準のいずれか一つに当てはまる場合は、その語全体に文字単位でルビを振る。ただし、同一文章内では初出時のみ。

JEA教材では、以下のL4(1)に従うこととする

- ＊複合語の場合、単独の語構成成分が基準を満たせば、原則としてルビ不要（つけてもよい）

例：「取り入れる」「生き残る」

- ＊初級の語・文字やルビ不要のレベル基準を満たすケースであっても、低頻度の読み（例：「様子」＝ようす）が含まれている場合、熟字訓（例：「土産」＝みやげ）の場合はルビ。一つの漢字に複数の音や訓があるケースもルビを検討する）例：「発達」と「発足」など。

練習問題では以下の語句にはルビ不要。

「読解」「問い」「下線」「語句」「指す」「本文」「段落」

- ・語彙レベルが2000位以内であったとしても、漢字レベルが351位以上の字を含む場合は、語全体に文字単位でルビを振る。ただし、同一文章内では初出時のみ。
- ・そのほか教師的な視点で見て、難しいと思うもの（複数の読みのうち頻度の低いほうの読みが使われている場合など、例「様子」）があればルビを振る。「田舎」のような熟字訓にも語単位でルビを振る。

J-LEX (<http://www17408ui.sakura.ne.jp/index.html>) での基準順位

J-LEX説明<http://www17408ui.sakura.ne.jp/tatsum/webtools.html#jlex>

	漢字	語彙
L3前半	286～	1285～
L3後半	300～	1501～
L4(1)	351～	2001～
L4(1)(2)	401～	2501～
L4(2)	451～	3001～
L5前半	500～	3500～
L5後半	550～	4000～
L6	600～	4500～

『語彙ドン！』Vol.1, Vol.2（くろしお出版）もこの基準に従っており、Vol.1ではL.12以降、少しずつルビが減るという「しかけ」になっています。

現在作成中の中級後半の教材もこの基準に従っています。

Q. 「9. 漢字学習と語彙学習をどう関連づけたらよいか」について→松下先生の「いまのところの考え・・・語の順序を先に決める」が私もいいと思っています。「日本語教育の参照枠 報告」(p.68)では、基礎漢字122(読みの正確さや書き方ではなく、意味の理解を優先)→基礎漢字に加えて個々の学習者にとって必要な漢字を選定、という漢字学習のイメージが示されていますが、これは松下先生の「いまのところの考え」と矛盾なく成り立ちそうですか？

→基礎漢字122は見たが、評価が難しい。意図がよくわからない。目的や状況が不明確な中で漢字だけをポンと示すことにはあまり意味がないような気がする。やるなら語例を示すぐらいはしたほうがいいのではないか。

Q. J-LEXの使い方、日本語学術共通語彙リストの見方について少しご説明いただければありがたいです。

→J-LEXを利用したリライト（語彙レベルの調節）の方法については、「松下言語学習ラボ」の「口頭発表」のところから、松下（2015）を参照してほしい。

<http://www17408ui.sakura.ne.jp/tatsum/presentation.html>

Q. 「9. 漢字学習と語彙学習をどう関連づけたらよいか」について、「語彙プロファイラー」と「語彙テスト」を同じデータベースに基づいて測っておられるということですが、そのデータベースが公開されているのであれば、是非教えていただきたいです。

Q. 語の頻度レベルと漢字の頻度レベルの対応のさせ方、同一のデータベースに基づいて、語彙頻度プロファイラーでテキストの語彙的難度を測定すること、および学習者のレベルを語彙テストで測定することの意味はよく理解できるのですが、具体的にはどのようなプロファイラー・テストを用いるのかを教えてください。

→以下をご参照ください。

<http://www17408ui.sakura.ne.jp/tatsum/database.html>

語彙テストも一部公開しています。以下をご参照ください。

<http://www17408ui.sakura.ne.jp/tatsum/webtest.html>

ほかにも、国立国語研究所などでは様々な語彙データが公開されています。

<https://www.ninjal.ac.jp/resources/search/>

「検索パネルを開く」から検索してみてください。

Q. 非漢字圏の学習者には、漢字のイメージですぐ語彙が思い浮かべるのは非常に難しい。何かいい方法があれば、教えてほしい。

→この質問は「**非漢字圏の学習者はどう漢字語を学習したらよい**か」と同じ。いろいろあるが・・・

- ・L1の処理方略はL2にも転移する。音韻→意味のリンクは作ろうとしているはず。それに加えて、字形（→音韻）→意味のリンクを強化する必要がある。そのときに**字形→和語「明るい」、字形「明」→漢語「明白」「明瞭」「自明」「解明」**のようにしていくと、**和語（訓読み/akarui/）⇔漢語（音読み/mei/）**のリンクもできていくはず。そうすることで語彙学習負担が減るのでは。
- ・漢字による視覚的な連合が加わるとより強固で広がりのある語彙ネットワークになる。正しい知識への手掛かりは多いほうが良いはず。（玉岡先生いかがでしょうか。）

Q.（「8. どうすれば語彙学習負担を減らせるか」と「9. 漢字学習と語彙学習をどう関連付けたらよいか」に関して）8と9でご説明いただいた内容を最もよく体現した教材をお教えてください。

→8. については目的にあったジャンルの教材ということ。（大学進学など）アカデミックな目的の中級であれば、手前味噌だが、『語彙ドン！』はよいはず。専門に関係なく、（日常生活では頻度が高くないが）アカデミックな文章・談話で頻度の高い語彙を効率的に学習できるはず。

→9については、よくわからないが、案外少ないかもしれない。音訓の関係の整理は案外難しいのかも。なぜなら、類似項目による干渉の危険が同時に伴うから。一方が既習であることを保証する形で学習が進まないと、危険がある。

→既習項目に関連する新項目が出てきたときに既習項目を合わせて整理するという考え方がよさそう。（それを体系的にやっている教材があれば、それはよい教材かも。作ってほしい。）逆に類似項目を初出でまとめて学習しようとする（音と訓の初出同時学習、類義語のまとめ学習など）はよくない可能性が高い。

→折に触れて**学習者に学習方法の意識付けをする**という感じのほうがいいのではないか。

→その前提の上でいえば、案外、能力試験の対策本の中にもいいものがあるかもしれない。能力試験は語彙の異なる側面を異なる形式で測定するようになっているので、対策本もそうなのであれば、誤答に対するフィードバックに気をつければ、役に立つのかもしれない。また、漢検対策アプリみたいなのも無料で使える割にはおもしろいものがある。

Q. これは簡単に解答できることではないと思いますが、一般的に使われている日本語の教科書、特に初級、中級レベルのものが、どれだけ語彙学習負担を減らすことや漢字学習と語彙学習を関連づけることを念頭に作られているのか。どの教科書シリーズがそれに成功しているのかなども知りたいと思いました。

→教科書については、いろいろあるので、どれが成功しているかは勉強不足で良くわからない。ただ、複数の読みや意味がある漢字の場合、二つ以上を同時に新しく学ぶことは避けたほうがよいという原則があり、既に最低一つが既習かどうかを確認できないと、かえって混乱が起きやすくなる。そう考えると、そこまで考えて作られた教材がそれほどあるとは思えない。むしろ、似たものをまとめて教えるということを考えず、内容のおもしろい教科書で、重要な語は繰り返し使うようなタスクが多いものが良いと考える。音と訓の整理などは、本文を読み進めたり、タスクを行ったりする中で、教師が学習者のレベルを推し量りながら、学習者にとって難しそうな語が出てきたときに、訓読み語がより基礎的であれば「この語は訓読みでどう読みますか」のように尋ねながら語彙ネットワークの形成を促していくのが良いのではないか。（[本多2022](#)の「媒介語」は参考になる概念。）

Q. 「8. どうすれば語彙学習負担を減らせるか」と「9. 漢字学習と語彙学習をどう関連付けたらよいか」に関して、具体的な教室活動なども示していただきたいです。

Q. どうすれば語彙学習負担を減らせるかについて、「目的を定め、より少ない単語の数でテキストカバー率を達成する・・・」とありますがこれは教師側がすべきことですね。どうすれば学習者側が自律的に語彙学習に取り組めるかが依然課題だと思いました。何かヒントがありますでしょうか。「自分にとってより必要度の高いものから順に・・・」「学習する語彙の数を減らすこと」とありましたが、学習者側がこれらを念頭に置いて語彙学習をするようにするにはどのように導けばいいのでしょうか。

→ 8 については、教室活動以前に、まず素材の選択が学習者の目的に見合っているかどうかを検討する必要がある。学習者ニーズは多様だが、共通ニーズがあればそれに見合った素材を使用し、異なるニーズに対応するには、選択科目を設ける、あるいは1科目内で学習者の選択に任せて学習する部分を設けるなどの方法で対処するのがよいのではないか。自律学習については、以下の本も参考にしてほしい。

桜美林大学日本語プログラム「グループさくら」(2007) 『[自律を目指すことばの学習](#)』
(凡人社)

→ 9 については、レクチャービデオをもう一度聞きなおしてほしい。あとは他の質問への回答も参照してほしい。

Q. 「9. 漢字学習と語彙学習をどう関連付けたらよいか」に関して、語彙と文法の関係はありますでしょうか。

→関係はあるはずだが、難しい質問で、簡単には答えられない。以下の6を参照してほしい。

松下達彦 (2018) 「日本語語彙習得に関わる普遍性と個別性 –漢字をめぐる問題を中心に–」『第四回学習者コーパス・ワークショップ&シンポジウム –第二言語習得における語彙の役割–』発表資料、学習者コーパスに基づく第二言語としての日本語の習得研究, 国立国語研究所

Q. 職業に関する単語や国名に関する単語を一度にたくさん紹介する教科書がありますが、これは有効な教え方でしょうか。私自身の経験から申しますと、学生は一度しか出て来ない単語は覚えないと感じています。

→単なる参照資料であればよいが、それをまとめて学習するようなやり方には弊害もあり得る。Semantic clusterという語で研究されているが、意味や形の似ている複数の語を初出で同時に学習しないほうがよいと言われている。一方が既習である場合や、ある程度学習が進んだところで復習としてまとめるのは有効だと考えられる。

→「一度しか出て来ない単語は覚えない」は基本的にはその通りで、必要な語を（類似の異なる文脈で）繰り返し使用するようなタスクが語彙学習には有効。付随的語彙学習（incidental vocabulary learning）の研究では、多読など語彙学習を意識しない学習で未知語に何度ぐらい出会えば習得されるかという研究があるが、意識の向かない状態では一度や二度では習得されないのは明らか。

Q.（「10. 目的に応じた語彙学習とは」に関して）ドラマのことが上級に区分けされている理由は？ ドラマは小説とは異なり、ジャンルにかなり左右される印象があるが。

→確かにドラマのほうが小説よりは易しい（日常会話語彙に近い）と考えられる。画像があると理解が進むことや語彙習得も進むという研究はすでにある。

→小説では（低頻度語彙＝上級レベルの）文学的な描写がある場合でも、映画・ドラマになると映像に解消されて語彙的には易しくなることが多い。（例：デビッド・ゾペティ「いちげんさん」は小説のほうがおもしろい）

→ジャンルに左右されるのもその通りであろう。

Q. 「10. 目的に応じた語彙学習とは」に出てくるテキストカバー効率（TCE）の計算式とテキストカバー効率による語彙学習の効率性の評価を示した表の読み取り方をもう少し説明してほしい。

→ 計算式やその意味は、もう一度ビデオで復習してほしい。

要は語彙リスト中のある1語の100万語あたりの頻度ということ。

→ 必要な数値を自分で計算するには、形態素解析が必要。いまは「Web茶まめ」で誰でも簡単にできる。

結果をExcelで開いて、目的に応じて「語彙素」「書字形」「書字形基本形」などの列を取り出して、異なり語数や延べ語数を数えればよい。

（私は AntWordProfiler, AntConcなどを使うが、ここでは詳しい方法は省略。総研大（博士課程）、明治大（修士課程）、一橋大（修士課程）などの大学院に来るか、研究室のスタッフになってもらえれば教えられます。）

「10. 目的に応じた語彙学習とは—領域特徴語の抽出と利用—」について→「日本語教育の参照枠 報告」(p.70)では、今後の検討課題として「生活・就労・留学などの分野別漢字を示すためには、対象グループの言語活動調査を行い、各分野で必要となる漢字の抽出」が必要だとされていますが、松下先生の学術共通語彙のようなものが語彙調査としても漢字調査としても(留学だけでなく)「生活・就労」といった分野でも本当に必要だと思いますし、その先の上級ではより細分化された各領域の特徴語リストが必要だと思います。こういったことは各日本語教育機関や教師が個々に調査するのはとても大変で効率も悪いので、何か組織的な研究活動ができたらいいのにと思っています。そのような動きはもしかしてすでに始まっていたりしますでしょうか？

→様々な動きがあり、就労者を対象にした新しい日本語学習教材も出ているようだが、語彙に特化したものは管見の限り、見当たらない。必要性について、数人の話を聞いたところ、工業のほか農業、漁業などもあり、現場があまりに多様なので、難しそうだと感じている。また、技能実習や特定技能の職場について言えば、仕事で使う日本語にはパターンがあるようで、限られた語彙で済んでいるように見える。ただ、病院に行くときに困ることがあるなど、生活の日本語はもちろん重要で、それも含めた日本語力があると、職場の「リーダー」になれる(昇進できる)ということはあるようなので、まずは生活の日本語を上達させることが肝要で、そのうえで、仕事の日本語については不明点の聞き返しや確認、メモと復習といったコミュニケーションと学習のストラテジーを身につけることが大切ではないかと思う。

Q. 「10. 目的に応じた語彙学習とはー領域特徴語の抽出と利用ー」についてですが、『日本語を読むための語彙データベース』はBCCWJから各分野の領域特徴語を抽出してまとめたものでしょうか。

→いいえ、『日本語を読むための語彙データベース』はBCCWJの作成途上で公開されたBCCWJ 2009年モニター版（約3,300万語）に基づいています。書籍（約2,800万語）とYahoo知恵袋（約500万語）に基づいて作られており、完成版BCCWJ（約1億語）に含まれている新聞や雑誌のデータは含まれていません。これは欠点でもありますが、長所でもあります。欠点は時事用語・新語・固有名詞などがやや少ないということですが、長所はそれだけ語彙頻度の安定度が高いということです。

領域特徴語はこのデータベースに含まれる語の一部で、このBCCWJ2009年モニター版に基づいて抽出されています。詳しくは以下のサイトから、各語彙リストの説明を参照してください。

<http://www17408ui.sakura.ne.jp/tatsum/list.html>

Q. 「10. 目的に応じた語彙学習とは一領域特徴語の抽出と利用一」について。(1) 漢字の理解がほとんど問題にならない中国人日本語学習者にとって、語構成の理解などにおいて特に気を付けなければならないのは何でしょうか。語彙習得においては、漢字圏以外の学習者と違うところは何でしょうか。(2) 領域特徴語をどうやって抽出できるのでしょうか。

→ (1) 漢字語の語構成については、日中両語に共通の部分が多いので、使える知識は(自然に)使い、異なる部分については、意識的に「抑制する(inhibit)」ことを学習する必要がある。品詞性の異なる語(「参考」「関心」など)などにも注意する必要がある。学習時に連語などの語句レベル、あるいはさらに文脈を伴うレベルで学習することも有効。あとは中国語母語学習者の場合、テレビや映画を字幕付きで視聴するのはかなり有効な学習方法のように思われる。文字と音と意味を同時処理するようなタスクは有効だと予測される。

(2) 抽出の指標がいろいろあるが、対数尤度比(たいすうゆうどひ、Log-likelihood Ratio)を使うことが多い。詳しくは以下の文献を参照。抽出は、具体的には、形態素解析を行い、AntConcという無料のツールを使う。詳しく知りたい人は研究室スタッフになるか、大学院(明治大、一橋大)の授業に参加してほしい。

Chujo, K., & Utiyama, M. (2006). Selecting level-specific specialized vocabulary using statistical measures. *System*, 34, 255-269.

Q. 既知語の判定基準について伺いたいです。基本的に、教科書に載っている語彙の数を基準とされているでしょうか？

→教科書に出ていても習得されているとは限らないので、いちばんよいのはテストで測定して、おおよその既知レベルを推定することです。VSTRJ-15Kなどを使ってみてください。150問であれば10問ずつに区切って、10問中9問正解なら、そのレベルは概ねマスターされていると考えることができます。（半分の75問の場合は5問ずつに区切り、5問全問正解のレベルを見ていきます。）

VSTRJ-15K <http://www17408ui.sakura.ne.jp/tatsum/webtest.html#vstrj>

→ある特定のテキストについては、学習者に未知語に下線を引いてもらうといった活動もあり得ます。

→一般的な大雑把な予測としては、語彙は概ね頻度順に習得されるということがあります。ですのでJ-LEXでテキストの語彙がどのレベルに分布しているかを把握すれば、合わせてVSTRJ-15Kの結果を見れば、ある学習者グループのレベルと、テキストの語彙的難度が合っているかどうかを判定できるようになります。

Q. 松下先生ご自身の研究について非常に興味を持っています。講演会で研究方法や結果に関して詳しくご説明いただければ幸いです。どうぞよろしくお願いします。

→ありがとうございます。短時間ではとても説明しきれないので、まずは論文を読んでいただければと思います。あとは別の機会があれば、改めてお話ししたいと思います。