

ロシア語教育における基本語彙データベースの制作について

尾子 洋一郎

キーワード：ロシア語教育，基本語彙，データベース，e-Learning

0.0 本稿は，ロシア語基本語彙データベース（以下DB）の制作とその過程について述べるものである。また，MP3 ファイルの「ジャケット（絵）」を利用し，視覚的に語彙習得を目指す自習教材ならびに教室提示用ソフトウェアについて述べる。¹

第一部 データベース制作について

1.0 本 DB は，複数の教員が用い，任意の複数のロシア語教科書の語彙を包括的に含み，かつ音声情報をも含むことを目的として制作された。教員のコンピュータスキルに応じ，カスタマイズならびにメンテナンスが柔軟にでき，Windows でも MacOS でも（他の任意の OS でもよい）利用できるものを制作することを目的とする。

1.1 データには以下の情報を含む。

- a) 単語番号
- b) ロシア語つづり
- c) 教科書初出の課 c') 教科書に出てきたページごとの順番（位置）
- d) 音声
- e) 音声ファイルの位置

1.2 DB 制作にあたり，使用した教科書は以下の通り。

- a) 佐藤純一. 2001. 『NHK 新ロシア語入門』 日本放送出版協会
- b) 安徳二ナ, 木村英明, 柴田信子, 杉山秀子, 堤正典, 中本信幸, 2003. 『21 世紀のロシア語』 大学書林
- c) 北岡千夏, 横井幸子, 三浦由香利 2005 『ロシア語能力検定試験 合格への手引き—3 級・4 級対策問題集』 南雲フェニックス
- d) 米重文樹, ピョートル・トマルキン. 1996. 『話すロシア語入門』 白水社
- e) 諫早勇一, 服部文昭, 大平陽一, イリナ・メリニコワ. 2002. 『セメスターのロシア語読本』 白水社

¹ 本稿で触れるソフトウェアは特別に記載がなければ Windows 版である。

堤正典・小林潔編『ロシア語学とロシア語教育Ⅲ』神奈川大学ユーラシア研究センター，2011 年，pp. 35-45.
Masanori TSUTSUMI and Kiyoshi KOBAYASHI (eds.) *Russian Linguistics and Language Education*. III.
Yokohama: The Eurasia Research Centre Kanagawa University, 2011, pp. 35-45.

1.3 上記教科書から、統計的頻度ならびに経験的に基本語彙を抽出する（堤・小林による研究）。

1.4 抽出された語彙に、作業用の通し番号をひとつずつ割り当て、つづりならびに音声情報を付与する。この通し番号（仮に単語番号とする）は、表計算ソフトなどで、単語番号をもとに、DB のカスタマイズや並べ替えができるように割り当てたものであって、他意はない。データは、Unicode (UTF-8)を用い、ロシア語のつづりの他に、日本語の語義も付加できるようにした。また、汎用性を持たせるため、カンマ区切りのプレーンテキストファイル（CSV 形式）として、本 DB は制作された。（参照：堤ほか 2009）

1.5 CSV形式のファイルとしてデータを作成したことによって、任意のデータベースソフトウェアで利用ならびに加工が可能である。²

1.6 どの教科書で使われているかの情報も含んでいるので、単語から、教科書の逆引きが可能であり、語彙研究にも益すると考える。

A	B	C	D	E	F	G	H
通し番号	見出し語	МОЗАИКА	_21c	話すロシア語	検定	NHK	file
1	2 физика				109		file: トゥルビーンさん演
2	3 анкета		21				file: トゥルビーンさん演
3	4 аэробика		11				file: トゥルビーンさん演
4	5 варенье		16				file: トゥルビーンさん演
5	6 воспитательница		16				file: トゥルビーンさん演
6	7 влажный		18				file: トゥルビーンさん演
7	8 выпить	9					file: トゥルビーンさん演
8	9 гордиться		20				file: トゥルビーンさん演
9	10 документ		11				file: トゥルビーンさん演
10	11 дополнительный		21				file: トゥルビーンさん演
11	12 еда		17				file: トゥルビーンさん演
12	13 жара		18				file: トゥルビーンさん演
13	14 живопись	9					file: トゥルビーンさん演
14	15 журналистка		16				file: トゥルビーンさん演
15	16 завидовать	11					file: トゥルビーンさん演
16	17 зависть						file: トゥルビーンさん演
17	18 интернет		19				file: トゥルビーンさん演
18	19 июньский	10					file: トゥルビーンさん演
19	20 кардиограмма		17				file: トゥルビーンさん演
20	21 карта	8					file: トゥルビーンさん演
21	22 клен		18				file: トゥルビーンさん演
22	23 колоть		17				file: トゥルビーンさん演
23	24 исполнение	0					file: トゥルビーンさん演

図 1 作成した DB を Microsoft Excel 2007 で開いた例

2.0 問題点

2.1 本 DB 制作においての問題点は、音声情報をいかに取り込むかであった。結論から言えば、音声情報は、音声ファイル（実体ファイル）の位置（Path）を記述することで、本体の肥大化を防ぐとともに、テキストファイルとして DB を制作することができた。音声については

²² もちろんテキストエディタでの編集加工も可能である。

4.0 において後述する。

2.2 単語の初出情報ならびに、教科書での出現順をどのようにして再現するか。この問題は、教科書の初出情報の、該当する単語の課番号に、小数点以下の数値を付与することによって解決できた。

例) a という教科書の第 4 課の 3 番目に出る語であれば、語に与えられる番号は「4.003」となる（一課に 100 語現れると想定して、小数点以下の値が決められる）。

2.3 力点の問題。語に力点を打ち、それを表示する場合、力点の再現は環境依存になるので、本 DB では力点を全てオミットした。

3.0 教室提示用ソフトウェアならびに学習者利用ソフトウェアについて

3.1 本 DB は前述の通り、単なる Unicode のテキストファイルに過ぎないので、そのままでは教室提示や、自前の単語カード作成といった用途には適さず、ユーザインターフェースを持った、データベースソフトウェアないし表計算ソフトウェアが必要である。

3.2 データベースソフトウェアの選定にあたっては、複数の OS に対応していること、限りなく無償に近いこと（アカデミック版の有無）、カスタマイズの容易さ、Unicode に対応していることを考慮した。候補に挙げたのは、4D 社の「4D」、FileMaker 社の「FileMaker Pro」である。いずれもリレーショナル・データベースであり、安定性も高く、運用実績も多い。今回、本 DB のデータベースソフトウェアとして、「FileMaker Pro Advanced」を採用した。Windows と MacOS に対応し、本データベースソフトウェアを持っていなくても動作する、ランタイムアプリケーションを作成できるからである。また、簡易ウェブサーバ機能もあり、学習者に DB を開放することも可能である。同時に簡易なユーザインターフェースを自作することも可能である。



図 2 「FileMaker Pro」による教室提示用単語 DB. 単語左の「音声再生」ボタンを押すと、対応した音声別のソフトウェアによって再生される。当該語を含む一覧の Excel への書き出しも可能。

4.0 音声ならびに音声情報

4.1 3で述べたデータベースソフトウェアでは、音声そのものを格納することはできない。従って、音声ファイルの Path を、音声再生ソフトウェアに渡して、音声ファイルを再生させるようにした。

例) file:1223.mp3

4.2 音声ファイルは、MP3 形式で、32Khz のサンプリングレート、96Kbps のビットレートとした。(大須賀(2007: 32)は 64Kbps でも良いとしている。)

4.3 OS 側で、拡張子と、対応するソフトウェアの関連付けを行えば、任意の音声再生ソフトウェアで、音声ファイルを再生することが可能である。

4.4 付記すれば、この音声ファイルは、ロシア語ネイティブ(男性・モスクワ州出身)により、神奈川大学 20 号館のスタジオにて、2009 年 3 月 3 日に録音・音声波形編集ソフトの「Sound It!」で CD-DA と同じサンプリング周波数、ビットレートで録音されたものである。録音された音声ファイルは単一の wave ファイルであり、約 2500 語を一気呵成に読み上げた物である。

4.4.1 「Sound It!」で録音されたものは、マイクが悪かったのか、ソフトウェアの仕様なのか、語頭の[k]音の脱落が生じた。成蹊大学法学部の湯山トミ子教授に教えを請うた所、マイクが悪いのではないかと貴重な指摘をいただいた。

4.5 DB制作にあたり、この単一ファイルを語、ならびに成句ごとに切り分ける作業は人力で行った。はじめは、無音検出ソフトウェアによって、自動的に単語を切り分けたのだが、不精確で使用に耐えないため、一単語一単語、耳で聞いて、ファイルを区切ることにした。使用したソフトは「Adobe SoundBoothCS4」である。

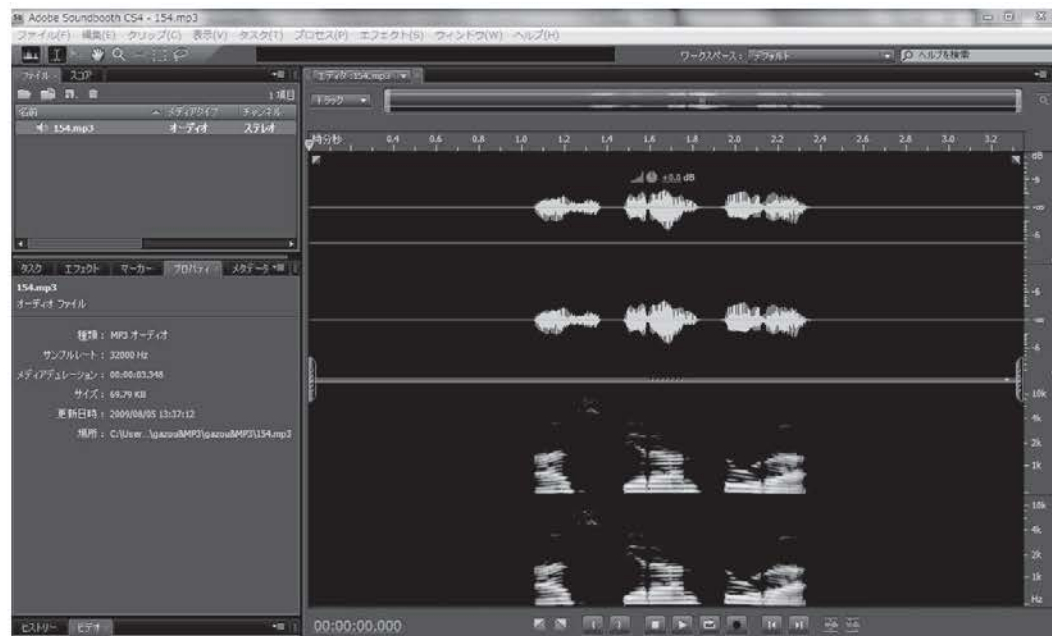


図 3 「Adobe SoundBoothCS4」で編集集中の音声

4.6 「Adobe SoundBoothCS4」は音響スペクトログラムも表示できる録音・音声波形編集ソフトであり、単語音声以外のノイズを視覚的に(手動で)除去できるのが特徴である。また、「Adobe Flash」での流用も視野に入れていたため、同ソフトウェアを選択した。4.5で述べたように、人力で一つずつファイルを区切り、ファイルを作成する際に、音声波形だけではわからないノイズを、スペクトログラムによって判断し除去し、単語／例文の前後に2秒ほどの無音を挿入し、ひとつひとつのファイルを作成した。

4.7 上記方法によって作成された音声ファイルを、任意の名称のフォルダに入れ、その Path を 4.1 のように、DB に記述した。この際、絶対 Path ではなく、相対 Path を用いた。

第二部 MP3 ファイルに画像を埋め込み活用する方法

5.0 音声ファイルの副次的利用法

5.1 MP3, AAC は ID3 タグを格納するデータ領域を持っているとともに, たとえばアルバムジャケットといった図(絵)をも格納するデータ領域を確保している. ISO/IEC 11172-3 MPEG-1 Audio の規定により, [ヘッダ | CRC | サイド情報 | メインデータ | 付加データ] の順にメインデータ(音声データ)そのもの以外の情報を「埋め込む」ことが可能である.

5.2 筆者は, この後者のデータ領域に着目し, 「ジャケット」表示は, 音声を再生させながら, 絵を表示させ, 視覚的に語彙を習得する手助けとなると考えた.

5.3 携帯音楽プレイヤー(SONY の Walkman, Apple 社の iPod など)が対応する音声形式で, 機種を問わず利用できるのが MP3 なので, AAC ではなく, MP3 を選ぶ.

6.0 アートワーク(アルバムジャケットアート)とカバーフロー~iTunes を用いて

6.1 Apple 社のマルチメディア再生・管理ソフトの「iTunes」(無償)を母艦として, iPod や iPad で表示される ID3 タグに例文や文法事項を書き込み, 音声とともに, 文字情報を表示させ, 語学学習の助けにするものは, ロシア語(水上 2006)に限らず, 英語などの他の言語でも用いられている手法である.

6.2 iTunes には, カバーフロー(Cover Flow)という画面表示機能があり, 再生する楽曲のアルバムアートを表示させることができる(図). 5.2 で言及した通り, ロシア語基本語彙に, キリル文字でつづりを表示させたり, 従来の紙ベースでの辞書では再現が困難であった色彩語彙においては, その色彩を表示させることをも可能である.(注: Windows Media Player でもアートワークの表示が可能である).

6.3 このアートワークを作成するに当たり使用したソフトウェアは「Paint.NET」(フリーウェア)である. なるだけ金銭をかけないようというコンセプトのもと, いわゆるペイントソフトを網羅的に検討したところ, Unicode 対応のフリーウェアが「Paint.NET」であった. 複数のレイヤーを駆使することのできる, 本格的なソフトウェアである.

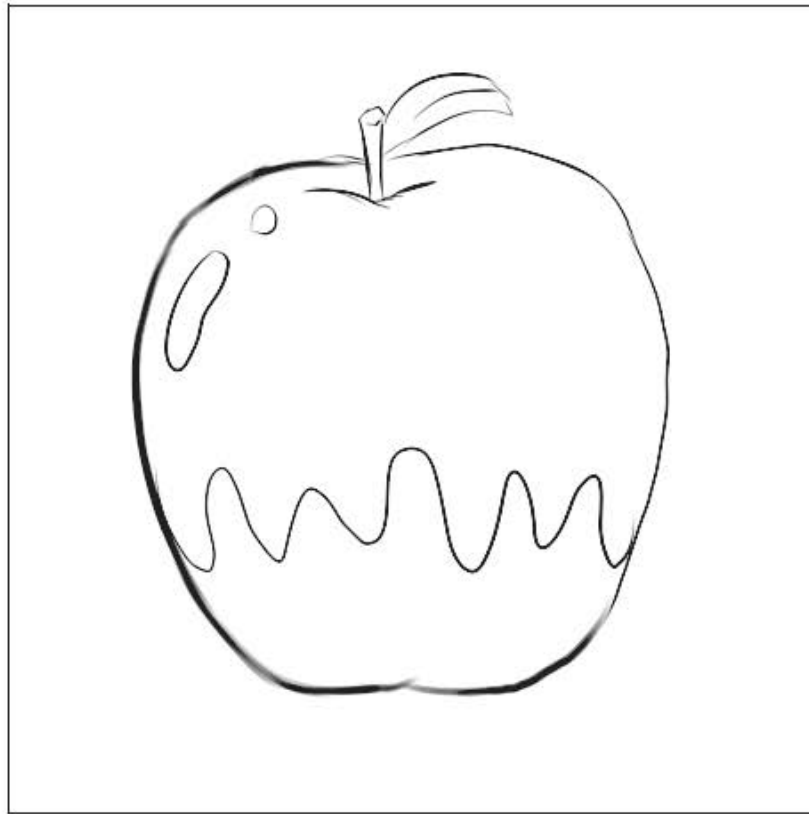


図 4 「Paint.NET」にてリンゴを描いた様子(井手征良氏作成)

6.4 アートワークの作成方法であるが、500 ピクセルの正方形を作り、その中にキリル文字なり写真なり、語彙に対応した絵を描けばよい（もっとも、「語に対応した絵」というものが難問中の難問なのであるが）。

6.5 MP3 ファイルへのアートワークの追加は、iTunes 上で行う。

6.6 アートワークをMP3 ファイルに格納するには、音声ファイルは最低限で 3 秒なければならない³。従って、「И」といった音声は 1 秒たらずなので、無音部分を前後に適宜付加し、全体で 3 秒を超えるように調節した。

6.6.1 図 4 にある左下の枠内に、作成したアートワークをドラッグ&ドロップすれば、自動的にアートワークが MP3 ファイルに格納される。

³ Apple社が仕様を公開していないので、筆者による経験則である。



図 5 iTunes で画像の埋め込みならびに表示をさせたところ。ID3 タグに記述した情報は画像の下に表示される。

6.7 特筆すべき点は、「ジャケット」の表示機能のあるソフトウェアであれば、iTunes でなくても、「ジャケット」の表示が可能な事である。すなわち、マイクロソフト社の「Windows Media Player」でも、SONY の「SonicStage」でも「表示」が可能である。これは、学習者が自習用に、任意の携帯音楽プレイヤーを使用して学習が可能であることを意味する。

7.0 ID3 タグの編集

7.0.1 6.1 で既に述べたように、ID3 タグに語意・語義を記述し、音声再生時に携帯音楽プレイヤーなどの画面にそれを表示させる事は行われており、英語、フランス語、韓国語、中国語教材などでは積極的に用いた物が書店店頭に並んでいる。⁴

7.1 ID3 タグの編集も iTunes を用いれば容易である。着目したのは、作曲者欄、ジャンル欄である。学習者の携帯音楽プレイヤーに楽曲が多い場合、あるいは、既にロシア語の学習教材がある場合、本 DB で作成した音声ファイルを検索するのは困難である。そこで、

⁴ 『iPodで中国語三昧プラス』（学習研究社 2009）など。この書籍ではiTunesの「歌詞」タブを用いている。筆者が着目した、アルバムアートや「情報」とは別である。

ジャンル欄に「ロシア語ロシア語」のように「ロシア語」という単語を重複させて書くことにより、授業で用いる単語の検索が簡単になる。

7.2 ID3 タグの作曲者欄は、iTunes 上でアートワークを表示させた場合、表示されない。しかしながら、同欄は検索時には検索対象となるので、便宜上定めた任意のキーワード+コロンと半角空白を記述することで、ヒットする。例えば、該当単語が、1.2 の a)のうち、第4課である事を示すには、「nhk:4」と書くことにより、a)を対象とした検索が可能になる。



図 6 iTunes 上で ID3 タグの編集を行っている図。「アートワーク」に先述のジャケットが格納される。「ファイル」メニューから「プロパティ」を選ぶ。

8.0 終わりに

8.1 本稿は以上に見てきたように、DB の制作について述べた。裨益するところは、MP3 ファイルそのものに絵を付け、視覚と聴覚から語彙習得、成句習得の一助になる事であろう。本稿執筆現在 (2011 年 3 月)、いまだに MP3 ファイルに絵をつけ語学学習に結び付けたものは、筆者の知る限り、ない。

8.2 問題点は MP3 ファイルの数だけ、「絵」を必要とする事である。多大な人力と時間を消費することである。如何に、短時間で要領良く、絵を用意できるかが今後の課題として残されよう。また、色彩を含む語彙の場合、ネイティブのインフォーマントによる色指定も必要であることも課題である。

参考文献

A)雑誌

大須賀史和. 2007. 「Webベースの授業用教材の作成方法と今後の課題」, 『スラヴィアーナ』(スラヴィアーナ編集委員会) 22, pp. 29-36.

B)ウェブサイト

水上則子. 2006. 『iPod をロシア語学習に使う』

http://www.nicol.ac.jp/~mizukami/study_Rus/iPod/iPod.html

C)電子データ

International Organization for Standardization 1993/1996 “Information technology - Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -Part 3:Audio”

<http://www.iso.org/>

D)発表

堤正典, 小林潔, 尾子洋一郎. 2009. 『ロシア語教育における基本語彙データベースの活用とその効果について』 私立大学情報教育協会平成 21 年度 教育改革 IT 戦略大会 9 月 3 日於私学会館.

小林潔, 尾子洋一郎, 堤正典. 2009. 『ロシア語初学者用語彙データベースの制作と運用』 日本ロシア文学会第 59 回定例総会・研究発表会 10 月 24, 25 日於筑波大学.

ロシア語教育における基本語彙データベースの制作について

尾子 洋一郎

5種類のロシア語教科書を元に必須単語 2500 語を抽出し、それを Unicode のテキストファイルにし、データベース化する方法について述べる。このデータベースは教科書ごとの単語検索も可能であるので、日本におけるロシア語教材についての研究にも様々に貢献するであろう。また、本稿が真に益するところのものは、MP3 音声ファイルとそれに格納したジャケット（絵）利用の提案である。これによって、視聴覚を用いた、単語や成句の自律学習を促進する礎となるとともに、教室提示用素材の用意の方法について言及する。本研究が既存の MP3 ファイルを用いた語学学習と異なる所は、ID3 タグを用いるのではなく、ジャケットを用いるという点であり、そこに真髄がある。