

上級レベル聴解素材のWEB公開と項目分析による素材の評価

(<http://japanese.human.metro-u.ac.jp/mic-j/chokai.html>)

西郡仁朗 宮田剛章

1. はじめに

現在様々なWEBサイトで日本語学習に関する情報や教育素材が公開されているが、それらは主に初級の内容であり、中上級レベルの教育的配慮のあるものは限られている。特に中上級視聴覚系のファイルとなると数えるほどしかないのが実情である(西郡, 2002)。本稿ではこうした課題への対応として2002年度から公開を開始した上級者向け聴解素材について述べる。素材制作の経緯と留意点、また、これをテストとして用いた場合の試用結果を報告することで、素材を利用する学習者と日本語教育関係者への支援と情報提供を行うこととしたい。

2. 聴解素材制作の経緯と留意点

今回公開された聴解素材を制作する際にモデルとしたのは日本語能力試験聴解である。この試験は日本語学習者が出会う日常生活の中で解決を求められる聴解の課題を想定してそれを遂行する能力を測定することが標榜されており、「1.学習者が教室を含めた日常生活で出会う課題と同様の課題を解決する。2.情報の正確な聞き取りに留まらず、聞いて何かをすること、つまり広い意味での問題解決を求めるもの」であることが特徴となっている(『日本語能力試験出題基準』, 1994)

素材制作は日本語教育学を学ぶ大学院生・学部生のプロジェクトワークとして行われ⁽¹⁾、参加者が素案を持ち寄りグループによる検討、教師からのコメントを経て改良・修正されてスクリプトを作成した。その後、日本語教育関係者が録音しWEB上で公開されることとなった。

プロジェクトワークとして素材を制作する際、次の諸点に留意した。これら留意点は西郡(2002)と重複する部分もあるが、本稿ではより具体的に記述することとする。なお、以下はこれまでの聴解素材制作から経験的に得られた内容である。

A. 語彙・文型以上のコミュニケーション能力の要素を盛り込む。

今回の素材制作での使用語彙については日本語能力試験にある語彙表を基準とし、

上級の1級レベルとなるように統制を行った。また、文型については同「出題基準」や初級・中級の主な日本語教科書を参考にした⁽²⁾。ただ、語彙と文型をもとに情報を正確に聞き取る能力は「必要だ」がそれだけで「十分」ではなく、場合によっては、細部にこだわらず、重要な部分に焦点を当てて聞くスキニング的なストラテジーを駆使する能力や、スキミング的に重要情報を聞き取って整理し、あるいは処理することによって一定の判断や想像あるいは推理、仮説の設定、検証、修正などの問題解決のための知的な作業を行う能力が求められる内容となるよう指示した。

B. はじめにプリクエスチョンを設け、場面と課題を必ず提示する。

課題の事前提示は問題解決型の聴解素材とするために当然必要なことである。また、問題解決を現実的なものに近づけるためには場面の提示も欠かせない。認知心理学的に見れば場面の事前提示は問題を解決するために必要なスキーマを前もって活性化させ、トップダウン型の情報処理を促進することになる。現実の問題解決行動では、自分がどこにいて今何を解決しようとしているのかは通常すでに分かっていることであり、その上で言語を理解したり聴解行動をとったりする。テストあるいは聴解素材でこうした設定を行うのは究極的には困難な面もあるが、できるだけ現実に近い近づけるためには、分かりやすく場面と課題を示し、その上で聴解テキストを提示するというのもっとも自然な流れであろう。

C. 特定の知識や思想・観念・文化背景が必要な問題は避ける。

作成する聴解素材は世界の不特定多数の日本語学習者がWEB上で受信可能なものである。また、海外の教師による授業での利用も考えられる。その際、前提知識として要求するのは上記語彙表の意味的内容だけとし、特定分野の知識、特定の思想・政治観や宗教観、特定文化での知識や常識に依拠した内容は避けることとした。

D. 倫理的な問題に配慮する。

素材はWEBで公開され、様々な文化的宗教的背景の学習者が利用すると予想される。日本語能力試験ほど公的なものではないにせよ、こうした学習者が心理的な抵抗なく利用するためには一定の倫理的な配慮が必要であろう。差別語（性・人種・職業・身分・宗教・身体障害）や差別的内容を含まないのはもちろんのこと、飲酒場面の泥酔者の会話など特定の宗教を信仰する学習者にとってタブーとなる内容は避けることとした。また、当然ではあるが素材はオリジナルなものに限られている。

E. 聴解能力以外の能力に大きく依存した問題は避ける。

素材の中に固有名詞が多数出てきた場合、多くの学習者にとってこれらは無意味綴りにすぎず、綴りを記憶する必要が生じ大きな記憶負担となる。その上で課題を解決

するというのは聴解能力だけでなく記憶力に大きく依存することになる。また、素材の中に計算課題が含まれているような場合、計算能力が課題解決に影響することもありうる。今回の素材制作では、こうした聴解能力以外の能力が必要なものはできるだけ避け、聴解能力だけが必要なものとなるよう方向付けを行った。

F. 日本語の話し言葉の特徴を積極的に取り入れる。

文型積み上げ式の日本語教育では、一語一語が明確で文あるいは文章として完結性の高いものを聴解素材とすることが多い。しかし、本稿での聴解素材は現実生活での問題解決に近づけることを目標としており、より自然な会話に近い作例が求められる。近年の談話分析研究を参考に（ザトラウスキー, 1993; 堀口, 1997; 宇佐美, 1997 など）文が短くターンが多い、あいづち・倒置・繰り返し・言い換え・体言止めが多い、会話独特の文末表現、婉曲的な言い方、繰り返し、イントネーションによる発話意図の伝達等々の特徴が会話の中に自然に配置されるよう指示した。ただし、自然会話の特徴の一つである複数話者の音声のオーバーラップは下記の音声聴取上の問題があるので避けている。

G. 音声聴取上の問題を考慮する。

本素材集を利用するのは、制作者とは無関係の学習者・教師であり、赤の他人が制作したものを聴くことになる。こうした場合制作者には以下のような配慮が必要であろう。まず、声質であるが、聴取者が見ず知らずの複数の方の声を弁別する際には音声の物理的特徴が大きく異なっている必要がある。異性間の会話であれば声の基本周波数が男女で大きく異なるので同定・弁別しやすいが、男性同士または女性同士のダイアログは人の同定と弁別が困難となる。このため、同性間の会話では、録音する際に男性ならばテノール声とバリトン声、女性ならばソプラノ声とアルト声というように、声の物理的特徴が明確に異なる人たちに会話を録音してもらった。第二に、自然な会話では声のオーバーラップが頻繁に見られるが、本素材集ではこれを避けている。というのは、現実の会話の場合、声がオーバーラップしていても片方の声は会話の当事者である自分の声で「聴解」する必要はなく、相手の発話を理解するだけでよい。しかし、教材としての音声は、第三者である学習者が別の二人の音声を聞くことになる。オーバーラップした声は、音響的にもかなり聞き取りにくく現実の会話の理解よりはるかに大きな処理容量（resource）が必要になる。このため、オーバーラップは自然会話の大きな特徴ではあるが、今回の聴解素材では極力避けることとした。第三に、プリクエスチョンや会話が無声子音（特に摩擦・破擦系の子音）から始まらないように配慮している。無声子音から始まる発話は音声文脈がないため聞き取りに

くい。特にプリクエスチョンや会話の冒頭など、音声文脈も意味的文脈も乏しい場合にはネイティブスピーカーでもうまく聞き取れずに戸惑うことがある。このため無声子音が冒頭に来そうな場合には、他の表現に置き換えたり、語順を変えたり、「あのう」「ねえ」「すみません」を前置するなどして聞き取りやすいものにした。

H. その他

その他にも以下の諸点が配慮された。

- ・四肢択一課題とする。
- ・聴取者が登場人物の一人に自分を投影できるもの（原則上盗み聞きでないもの）
- ・学習者の母語、男女、年齢による理解の差がでないもの。
- ・日本語音声聞き分けだけのものは避ける。
- ・正解が曖昧になりやすいものは避ける。
- ・場面設定が長すぎるものは避ける。
- ・問題を聞かなくても常識で答えられるものではないこと。

3. 音声ファイル利用の技術的環境

現在、世界のどの国においてもインターネットのブロードバンド化が進行している。その目的は音声・動画等、大容量のファイルの転送を速め、遠隔地にあるサーバーとの双方向的な情報のやりとりをストレスなく行うことである。しかし、ブロードバンド回線上で、音声・動画ファイルを再生する際、ファイル形式と再生ソフトの標準化の問題がある。音声・動画ファイルの再生ソフトとして有力なのは、WindowsMediaPlayer、QuickTimePlayer と RealPlayer の3種であり、いずれも無償のものである。それぞれ独自のファイル形式もあるが、こと音声に関してはMP3 (MPEG Layer3)ファイルが共通して使用でき、ダウンロードした上での利用もストリーミングも可能である。MP3 は MiniDisc に音楽などを記録するファイル形式としても一般化しており、現在最も広範囲で使用されている音声のファイル形式と言えよう。圧縮率は再生ソフト独自のファイル形式 (VMA ファイルなど)ほど高くないが、利用者のパソコンのシステム環境 (Windows か Macintosh か Linux か) や再生ソフトを問わずに使用できる点が強みである。今後は Mpeg4 など移動体通信にも対応し、圧縮率の高い新しいファイル形式が主流になっていくであろうが、それまでの間は MP3 を使用することが適当であると考え、本素材集でも採用することとした。

今回制作されたのは 39 種の音声 MP3 ファイルであり、以下の URL でスクリプトの

テキストデータとともに公開されている。サイトがおかれている東京都立大学、言語文化研究所、中国復旦大学³⁾の LAN 内ではストレスを感じることなく、すぐにストーリーミング再生が始まるが、それ以外の場所からは数秒から数十秒の時間を要する。

東京都立大学 <http://japanese.human.metro-u.ac.jp/mic-j/chokai.html>

言語文化研究所 <http://www.naganuma-school.or.jp/mic-j/chokai.html>

中国復旦大学 <http://www.sirpa.fudan.edu.cn/jallearn/chokai.html>

4. 試行テストとその項目分析の結果

4-1. 試行テストの実施

今回制作した聴解素材の試行テストは2002年10月言語文化研究所東京日本語学校で行われた⁴⁾。被験者は84名で上級レベルを集中教育で学んでいる学習者である。ほぼ全員が同年末の日本語能力試験を受験を予定していて、その予行演習としてのテスト参加でありモチベーションは高い様子であった。国籍(地域)は、中国13名、韓国41名、台湾19名、タイ3名、インドネシア2名、日本・カンボジア・フランス・香港・ミャンマー・ベトナム各1名と多岐にわたっていた。

39のMP3ファイルをパーソナルコンピュータで次々に再生し、それをスピーカで拡声して教室内に流すという集団実験方式をとった。音声の聴取環境に問題はなかったと思われる。中には数問絵付きの問題があったが、これはプロジェクターで拡大して見せた。全員が確実に絵を見ることができるよう配慮し、絵付き問題については特に長く音声再生開始までのインターバルをとった。

なお、問題25はファイルの不具合により、分析からは外した。

4-2. 項目分析

上級日本語学習者に対して実施した聴解問題に関する分析を行う。39項目における得点の一致度を測るべく、Cronbach(1970)の係数を求めた。その結果、 $\alpha = 0.8062$ となり、項目の一致度が高いことが分かった。これは聴解問題全体の信頼性が高いことを意味する。

このように全体としては本稿で用いた聴解テストの信頼性が高く、調査対象者の聴解能力を反映していると考えられるが、良質の問題を開発するためには各問題についてさらに検討していく、つまり項目分析をする必要がある。各問題に対し項目分析

を行う際、重要とされる観点として以下の3点を挙げる。(参考:野口, 1985)

- ・通過率が30%以上70%未満の範囲内でなければならない。
- ・識別力が低すぎてはいけない。
- ・各項目における選択肢に解答の頻度が極端に低いもの(つまり、回答者の殆どが選択していない選択肢)があってはならない。

通過率とは本稿でいう各項目の平均値を意味する。満点を1点としているため、項目が良質の問題であるための条件として、 $0.3 < \bar{x} < 0.7$ となる。この範囲から逸脱する項目は調査対象者の聴解能力を十分に弁別していきなく、信頼性に欠くことになるため、改善の余地がある問題であると考えられる。

識別力は全項目の合計点と各項目との相関係数で示される⁽⁵⁾。これは各項目が合計点にどれほど貢献しているかを示す指標となる。本稿において、ある聴解問題の識別力が高ければ、それは聴解能力を測定するのに寄与度の高い重要な問題といえる。それは聴解能力に従って弁別できる問題とも言い換えられ、問題の信頼性の高さを示す。逆に識別力が低すぎる問題は聴解能力に反映せず、問題の改善が必要であることを示唆する。本稿では、 $r < 0.183$ の項目を改良の余地がある問題であると見なした。

項目分析ではある項目において解答頻度が極端に低い選択肢がないことが望ましい。多肢選択型テストの場合、どの選択肢ももっともらしい内容になっていることが理想的である。特に問題が分からない解答者にとって、誤答選択肢が解答を錯乱させるものでなければ、その問題の正答率を結果的に高めることになってしまい、従って弁別力を下げ、信頼性を低下させることに繋がる。本稿では特にゼロ頻度の選択肢が2つ以上含む問題に注目し、改善の余地がある項目と見なした。

4-3. 項目分析の結果と考察

以上の観点より、各聴解問題において信頼性を検討することでより良質な問題を開発できるよう思索していくことにする。信頼性は妥当性の前提条件である。信頼性が保たれなければ妥当性を考慮する余地がなくなり、テスト問題そのものの意義が問われかねない。そのことから、本章では聴解テスト問題そのものの信頼性を検討し、ひいては今後の課題としてどのように改良すべきかを考えていくことになる。分析結果を表1に各聴解問題における記述統計、識別力と不選択選択肢数を示した。

表1の透過率、識別力、不選択選択肢数において、信頼性に問題があると考えられる項目に網掛けを施した。

表1 各聴解問題における記述統計、および識別力・不選択選択肢数

	度数	合計	標準偏差	平均値 (通過率)	点双列相関係数 (識別力)	不選択選択肢数
問題1	84	43	0.503	0.51	0.41	0
問題2	84	44	0.502	0.52	0.40	0
問題3	84	70	0.375	0.83	0.32	0
問題4	84	66	0.413	0.79	0.25	0
問題5	84	78	0.259	0.93	0.08	0
問題6	84	70	0.375	0.83	0.39	0
問題7	84	27	0.470	0.32	0.28	0
問題8	84	62	0.442	0.74	0.45	0
問題9	84	36	0.498	0.43	0.29	0
問題10	84	69	0.385	0.82	0.34	0
問題11	84	84	0.000	1.00	0.00	3
問題12	84	60	0.454	0.71	0.55	2
問題13	84	50	0.494	0.60	0.21	0
問題14	84	70	0.375	0.83	0.15	0
問題15	84	50	0.494	0.60	0.61	0
問題16	84	70	0.375	0.83	0.34	0
問題17	84	71	0.364	0.85	0.34	0
問題18	84	36	0.498	0.43	0.44	0
問題19	84	78	0.259	0.93	0.47	1
問題20	84	82	0.153	0.98	0.14	2
問題21	84	78	0.259	0.93	0.30	0
問題22	84	73	0.339	0.87	0.50	0
問題23	84	71	0.364	0.85	0.48	1
問題24	84	17	0.404	0.20	0.10	0
問題26	84	52	0.489	0.62	0.28	0
問題27	84	73	0.339	0.87	0.39	1
問題28	84	56	0.474	0.67	0.62	0
問題29	84	73	0.339	0.87	0.36	1
問題30	84	76	0.295	0.90	0.26	0
問題31	84	84	0.000	1.00	0.00	3
問題32	84	80	0.214	0.95	0.15	0
問題33	84	40	0.502	0.48	0.40	0
問題34	84	80	0.214	0.95	0.25	2
問題35	84	44	0.502	0.52	0.50	0
問題36	84	75	0.311	0.89	0.16	0

問題37	84	76	0.295	0.90	0.37	0
問題38	84	56	0.474	0.67	0.55	0
問題39	84	52	0.489	0.62	0.35	0
問題40	84	69	0.385	0.82	0.35	2

次に、項目分析に取り上げた3つの観点のうち、2つ以上の観点から信頼性に問題があると見なされた項目について考察する。信頼性に問題があると思われる項目は、問題5、11、12、14、20、24、31、32、34、36、40の計11項目である。それぞれの項目についての問題点を表2にまとめる。また、これらの問題のスク립トを稿末に資料として掲げる。全体的にみて、解答を容易にさせている談話構造や難易度の低い語彙、パラフレーズされていない誤選択肢が、調査対象者の正答率を極度に高めてしまっているものと考えられる。

表2 信頼性が欠如していると考えられる聴解問題とその分析

項目	考えられる問題点
問題5	テキストは長い、語彙が容易である。また、構成も最後の便名だけを聞き取れば容易に解答できる形式となっている。
問題11	全体的にテキストの内容が容易過ぎる。解答の鍵となる語彙も数字や曜日等、初級レベルのものが用いられている。
問題12	解答の鍵となる「詐欺」は難易度の高い語であるが、始めにある「暇」と最後の「私と行きたくないのかなあと思っちゃったよ」という情報は、テキストの最初と最後ということで認知度が高く、しかも表現が容易であるため、選択肢の1と4が錯乱肢として機能しなかった。
問題14	テキストの最終部だけを聞き取れば、容易に解答できる形式となっている。選択肢と同じ「旅行」や「温泉」という旅行を連想させる語があり、さらには「結婚してから25年」という夫の発言や「うふふ」という妻の笑いがあるため、必然的に「夫婦の旅行」という解答が容易に導かれる。
問題20	正答選択肢がパラフレーズされていなく、解答となるテキストの言語形式がほぼ同一である。また、談話構成も最後だけを聞き取れば解答できる形式となっている。
問題24	解答の鍵となる「6時半」が始めに登場するが、テキストが長いために忘れてしまっている可能性がある。またテキスト最後に出現する「とりあえず、7時には起こしてくれ」という内容に多くの解答者が惑わされ、7時という選択肢を選んでいった。この「とりあえず」が雨天時なのか、晴天・雨天時両方に修飾されるのかの判断は語用論の問題であり、筆者を含めた日本語母語話者も誤答するという結果となった。

問題31	正答選択肢がパラフレーズされていなく、解答となるテキストの言語形式がほぼ同一である。テキスト上において解答の鍵となる箇所が感情的なイントネーションを伴って顕著になり、しかも感情的イントネーションは1回しか起こっていないため、解答が極めて容易となる。
問題32	テキストで使われている語句が、「ベル」・「時計」・「結果を郵送」・「お疲れ様」・「空いている席」・「机の上にあるもの」など場面を類推する要素が多分にある。
問題34	難しくてもせいぜい日本語能力試験2級レベルの「桁」など、語彙が容易である。また、構成も最後だけを聞き取れば解答できる形式となっている。
問題36	テキスト上において解答の鍵となる箇所が連呼されているため、顕著になり通過率も高くなったものと考えられる。下位者でも正答できただけでなく、上位者の中には正答できなかったものも何名かいたため、識別力も低い。これは他のテキストに比べ、発話速度が高く調子についていけなかったためと考えられる。
問題40	テキストで使われる語句が容易である。また、形式パターンも典型的な「提示 否定 提示・・・」型で、しかもこの項目では「提示 否定」で終わっているため、非常に分かりやすい談話構成となってしまう。

5. おわりに

これまで見てきたように、今回の聴解素材は全体としての信頼性は高いものの、個々の項目を分析すると39項目のうち11の項目で何らかの問題点が見出されている。被験者数を増やしたり、選択肢を工夫すれば解消されるであろう点も散見されたが、談話として自然で、かつテスト項目として好ましい統計的な結果を出すのは容易ではない。統計的な好ましさを優先すると自然さが損なわれるという二律背反的な要素も多分にある。

今後もこうした聴解素材の開発とデータベース化を継続していくが、たとえテスト項目としての信頼性に問題があっても、そのままの形でWEBでの公開を行っていく。その理由は、各素材はテストとしてではなく聴解テキストの部分だけでも、自習や授業などで十分に利用可能なものであるし、こうした素材を日本語教育を志す人々と共同作成することで教育メディア論を効果的に伝播できるからである。また、今後「作例」という行為の自己分析と他者による分析・評価を通じ、言語教育における教育的配慮や教材開発技法が高められると考える。

-
- (1) 東京都立大学人文学部「言語表現法」、早稲田大学大学院日本語教育研究科「マルチメディア教育論」の2002年度コースワークの一環として行われた。
- (2) 東京外国語大学『初級日本語』、『中級日本語』、筑波大学『日本語表現文型中級II』などを参考とした。
- (3) 西郡(2003)を参照されたい。
- (4) 同校の小島美智子氏と神部秀夫氏に協力していただいた。記して感謝する。
- (5) 識別力については、当初、積率相関係数や順位相関係数が考えられたが、本稿で扱うデータは2分名義変数が含まれるため、これらの採用により分析結果の妥当性を低下させてしまう恐れがあった。ブラウン(1999:199)によると、点双列相関係数とは2分名義変数からなる項目と間隔変数である全項目総数との相関関係を示すのに適しているとされ、これを採用した。点双列相関係数の臨界値はピアソンの積率相関係数の臨界値と対応するため、それをを用いた。5%水準の場合、自由度が80(つまり、標本数Nが82)の臨界値は0.183で自由度が90(N=92)の臨界値は0.173であるので、自由度が82(N=84)の臨界値はその間でやや0.183に近い値であると考えられる。そこで本稿ではそれよりも下回る相関係数を持つ項目は識別力が低いものと見なした。