

2016 年度博士論文

日本語聴解テストにおけるテスト・アイテム分析と聴解行動
-アカデミック・リスニングの視点から-

桜美林大学大学院

足立 章子

目 次

1章 序論	1
1.1 本研究の背景	1
1.2 本研究の目的と意義	7
1.3 論文の構成	8
2章 言語テストの概観	11
2.1 言語能力の定義と言語テストの変遷	11
2.2 言語テストにおけるコミュニケーション能力	14
2.3 集団規準準拠テストと目標基準準拠テスト	15
2.3.1 集団規準準拠テスト	16
2.3.2 目標基準準拠テスト	17
2.4 日本語能力試験と日本留学試験	19
2.4.1 日本語能力試験	19
2.4.2 日本語能力試験 聴解	20
2.4.3 日本留学試験	30
2.4.4 日本留学試験 聴解	32
3章 聴解テストで測定される要素	36
3.1 アカデミック・リスニング	36
3.2 聴解能力と聴解テスト	38
3.3 聴解下位スキル	41
3.4 テストの入力ファセット	45
3.5 聴解の質的研究方法	50
4章 聴解テスト・アイテムの基礎的データ	52
4.1 基礎的データのための分析枠組み	52
4.2 結果と考察	57
4.2.1 理解の種類	57
4.2.2 場面	59
4.2.3 形態	63
4.2.4 トピック	64
4.2.5 人数	64
4.2.6 コミュニケーションの目的	65
4.2.7 機能	66
4.2.8 情報を得る理由	69
4.3 テキスト・バンクの提案	69
4.3.1 アイテム・バンク	69
4.3.2 聴解テスト・アイテムの作成のためのテキスト・バンク構築	71

4.4	改定前 2002 年版と 2010 年改訂版の比較・考察	76
4.5	結論と今後の課題	89
5章	聴解テスト・アイテムのスク립ト分析	92
5.1	研究の枠組み	92
5.1.1	文法的知識	96
5.1.2	テキスト知識	96
5.1.3	機能的知識	98
5.1.4	社会言語学的知識	98
5.1.5	方略的能力	99
5.2	分析対象と分析方法	100
5.3	結果の分析と考察	104
5.3.1	文法的知識	105
5.3.2	テキスト知識	106
5.3.3	機能的知識	109
5.3.4	社会言語学的知識	111
5.3.5	方略的能力	112
5.4	選択肢	113
5.5	結論と今後の課題	114
6章	テスト時の聴解行動（設問あり）	118
6.1	テスト時の聴解行動の研究目的	118
6.2	研究の枠組み	118
6.3	データ収集と分析の手順	120
6.3.1	対象者	120
6.3.2	対象アイテム	121
6.3.3	データ収集と分析の手順	122
6.4	留学生の聴解行動の分析結果と考察（設問あり）	123
6.4.1	全体	126
6.4.2	状況説明	128
6.4.3	プリタスク	129
6.4.4	本文	130
6.4.5	本文終了時	135
6.4.6	ポスト・クエスチョン	136
6.4.7	選択肢	136
6.5	日本人大学生の聴解行動の分析結果と考察（設問あり）	139
6.5.1	全体	140
6.5.2	状況説明	141
6.5.3	プリタスク	142
6.5.4	本文	144

6.5.5	本文終了時	148
6.5.6	ポスト・クエスチョン	149
6.5.7	選択肢	149
6.6	設問があるアイテムにおける2者の総括と考察	151
7章	テスト時の聴解行動（設問なし）	153
7.1	留学生の聴解行動の分析結果と考察（設問なし）	155
7.1.1	全体	156
7.1.2	本文	157
7.1.3	クエスチョン	161
7.1.4	選択肢	162
7.2	日本人大学生の聴解行動の分析結果と考察（設問なし）	166
7.2.1	全体	167
7.2.2	本文	168
7.2.3	クエスチョン	173
7.2.4	選択肢	174
7.3	設問がないアイテムにおける2者の総括と考察	177
7.4	提案と今後の課題	180
8章	結論	185
8.1	日留試験聴解テスト・アイテムの特徴	185
8.2	日留試験聴解テストのアカデミック・リスニング能力	186
8.3	留学生の聴解テスト時における聴解行動	187
8.4	今後の課題	188
参考資料		192
参考文献		203
謝辞		211

表 一 覧

表 1.1	研究の概要	10
表 2.1	集団規準準拠テストと目標基準準拠テストの違い	19
表 2.2	旧日能試の聴解課題	22
表 2.3	日留試と旧日能試のテスト課題の比較	31
表 2.4	聴解テスト・アイテムの比較：旧日能試、日留試、新日能試	33
表 2.5	日留試と TOEFL iBT の聴解テスト・アイテム比較	34
表 3.1	カンパセーショナル・スキルとアカデミック・スキル	42,93
表 3.2	聴解テスト・アイテムの設問、選択肢	48
表 4.1	理解の種類	57
表 4.2	場面の割合 (2002-2013)	60
表 4.3	アカデミック・ジャパニーズ・スキル・シラバス	61
表 4.4	アカデミック場面の下位場面 (2002-2013)	62
表 4.5	キャンパス場面の下位場面 (2002-2013)	62
表 4.6	ライフ場面の下位場面 (2002-2013)	63
表 4.7	形態 (2002-2013)	64
表 4.8	人数 (2002-2013)	64
表 4.9	コミュニケーションの目的 (2002-2013)	65
表 4.10	情報を授受する機能 (2002-2013)	66
表 4.11	聞き手に働きかける機能 (2002-2013)	67
表 4.12	相手の働きかけに応える機能 (2002-2013)	67
表 4.13	自己の感情を表す機能 (2002-2013)	68
表 4.14	個人・社会化に関する機能 (2002-2013)	69
表 4.15	情報を得る理由 (2002-2013)	69
表 4.16	アカデミック場面の形態 (2002-2013)	72
表 4.17	キャンパス場面の形態 (2002-2013)	72
表 4.18	ライフ場面の形態 (2002-2013)	72
表 4.19	理解の種類と比較 (改定前と改定後)	78
表 4.20	場面の比較 (改定前と改定後)	78
表 4.21	形態の比較 (改定前と改定後)	79
表 4.22	聴解テスト・アイテムのトピック一覧改定前 (2002-2009)	81
表 4.23	聴解テスト・アイテムのトピック一覧改定後 (2010-2013)	85
表 4.24	人数の比較 (改定前と改定後)	86
表 4.25	コミュニケーションの目的の比較 (改定前と改定後)	87
表 4.26	機能の比較 (改定前と改定後)	88
表 4.27	情報を得る理由の比較 (改定前と改定後)	88
表 5.1	Bachman & Palmer (1996)による言語能力の構成要素	93, 114
表 5.2	スクリプト分析におけるリスニング・スキルの枠組み①文法知識	96
表 5.3	スクリプト分析におけるリスニング・スキルの枠組み②テキスト知識	97

表 5.4	スクリプト分析におけるリスニング・スキルの枠組み③機能的知識	98
表 5.5	スクリプト分析におけるリスニング・スキルの枠組み④社会言語学的知識	99
表 5.6	スクリプト分析におけるリスニング・スキルの枠組み⑤方略的能力	100
表 5.7	下位スキル分析例	103
表 5.8	聴解テスト・アイテム本文のスクリプト・マーキング部分にみられた リスニング下位スキル (2002-2006)	104
表 5.8.1	聴解テスト・アイテム本文のスクリプト・マーキング部分にみられた リスニング下位スキル (2002-2006) 文法的知識	105
表 5.8.2	聴解テスト・アイテム本文のスクリプト・マーキング部分にみられた リスニング下位スキル (2002-2006) テキスト知識	106
表 5.8.3	聴解テスト・アイテム本文のスクリプト・マーキング部分にみられた リスニング下位スキル (2002-2006) 機能的知識	109
表 5.8.4	聴解テスト・アイテム本文のスクリプト・マーキング部分にみられた リスニング下位スキル (2002-2006) 社会言語学的知識	111
表 5.8.5	聴解テスト・アイテム本文のスクリプト・マーキング部分にみられた リスニング下位スキル (2002-2006) 方略的能力	112
表 5.9	本文から得た解答と選択肢の関係 (2002-2006)	113
表 6.1	アイテム全体の概念と聴解行動のヴァリエーション① (留学生設問あり)	126
表 6.2	状況説明時の概念と聴解行動のヴァリエーション② (留学生設問あり)	128
表 6.3	プリタスク時の概念と聴解行動のヴァリエーション③ (留学生設問あり)	129
表 6.4	本文時の概念と聴解行動のヴァリエーション④ (留学生設問あり)	131
表 6.5	本文終了時の概念と聴解行動のヴァリエーション⑤ (留学生設問あり)	135
表 6.6	ポスト・クエスチョン時の概念と聴解行動のヴァリエーション⑥ (留学生設問あり)	136, 161
表 6.7	選択肢時の概念と聴解行動のヴァリエーション⑦ (留学生設問あり)	137
表 6.8	アイテム全体の概念と聴解行動のヴァリエーション① (日本人大学生設問あり)	140
表 6.8.1	アイテム全体の地妖怪行動のヴァリエーション① (留学生と日本人大学伊勢設問ありの比較)	141
表 6.9	状況説明時の概念と聴解行動のヴァリエーション② (日本人大学生設問あり)	141
表 6.9.1	状況説明時の聴解行動のヴァリエーション② (留学生と日本人大学生設問ありの比較)	141
表 6.10	プリタスク時の概念と聴解行動のヴァリエーション③ (日本人大学生設問あり)	143
表 6.10.1	プリタスク時の聴解行動のヴァリエーション③ (留学生と日本人大学生設問ありの比較)	143
表 6.11	本文時の概念と聴解行動のヴァリエーション④ (日本人大学生設問あり)	144

表 6.11.1	本文時の聴解行動のヴァリエーション④ (留学生と日本人大学生設問ありの比較)「情報の受容」	145
表 6.11.2	本文時の聴解行動のヴァリエーション④ (留学生と日本人大学生設問ありの比較)「情報に対する態度」	146
表 6.11.3	本文時の聴解行動のヴァリエーション④ (留学生と日本人大学生設問ありの比較)「不安材料」	147
表 6.12	本文終了時の概念と聴解行動のヴァリエーション⑤ (日本人大学生設問あり)	148
表 6.12.1	本文終了時の聴解行動のヴァリエーション⑤ (留学生と日本人大学生設問ありの比較)	148
表 6.13	ポスト・クエスチョン時の概念と聴解行動のヴァリエーション⑥ (日本人大学生設問あり)	149, 173
表 6.14	選択肢時の概念と聴解行動のヴァリエーション⑦ (日本人大学生設問あり)	149
表 6.14.1	選択肢時の聴解行動のヴァリエーション⑦ (留学生と日本人大学生設問ありの比較)「確信・確定」と「確定」	149
表 6.14.2	選択肢時の聴解行動のヴァリエーション⑦ (留学生と日本人大学生設問ありの比較)「葛藤」と「不信」	150
表 7.1	アイテム全体の概念と聴解行動のヴァリエーション① (留学生設問なし)	156
表 7.1.1	アイテム全体の聴解行動のヴァリエーション① (留学生設問ありと設問なしの比較)	156
表 7.2	本文時の概念と聴解行動のヴァリエーション② (留学生設問なし)	157
表 7.2.1	本文時の聴解行動のヴァリエーション② (留学生設問ありと設問なしの比較)「情報の受容」	158
表 7.2.2	本文時の聴解行動のヴァリエーション② (留学生設問ありと設問なしの比較)「情報に対する態度」	159
表 7.2.3	本文時の聴解行動のヴァリエーション② (留学生設問ありと設問なしの比較)「不安材料」	161
表 7.3	クエスチョン時の概念と聴解行動のヴァリエーション③ (留学生設問なし)	162
表 7.4	選択肢時の概念と聴解行動のヴァリエーション④ (留学生設問なし)	163
表 7.4.1	選択肢時の聴解行動のヴァリエーション④ (留学生設問ありと設問なしの比較)「確信・確定」	163
表 7.4.2	選択肢時の聴解行動のヴァリエーション④ (留学生設問ありと設問なしの比較)「葛藤」と「葛藤・吟味」	164
表 7.5	アイテム全体の概念と聴解行動のヴァリエーション① (日本人大学生設問なし)	167
表 7.5.1	アイテム全体の聴解行動のヴァリエーション① (日本人大学生設問なしと設問ありの比較)	167
表 7.5.2	アイテム全体の概念と聴解行動のヴァリエーション① (留学生と日本人大学生設問なしの比較)	167

表 7.6	本文時の概念と聴解行動のヴァリエーション②（日本人大学生設問なし）	168
表 7.6.1	本文時の聴解行動のヴァリエーション② （日本人大学生設問ありと設問なしの比較）「情報の受容」	169
表 7.6.2	本文時の聴解行動のヴァリエーション② （留学生と日本人大学生設問なしの比較）「情報の受容」	169
表 7.6.3	本文時の聴解行動のヴァリエーション② （日本人大学生設問ありと設問なしの比較）「情報に対する態度」	170
表 7.6.4	本文時の聴解行動のヴァリエーション② （留学生と日本人大学生設問なしの比較）「情報に対する態度」	171
表 7.6.5	本文時の聴解行動のヴァリエーション② （日本人大学生設問ありと設問なしの比較）「不安材料」	172
表 7.6.6	本文時の聴解行動のヴァリエーション② （留学生と日本人大学生設問なしの比較）「不安材料」	172
表 7.7	クエスチョン時の概念と聴解行動のヴァリエーション③ （日本人大学生設問なし）	173
表 7.7.1	クエスチョン時の聴解行動のヴァリエーション③ （留学生と日本人大学生設問なしの比較）「確認」	174
表 7.7.2	クエスチョン時の聴解行動のヴァリエーション③ （留学生と日本人大学生設問なしの比較）「照合」	174
表 7.8	選択肢時の概念と聴解行動のヴァリエーション④（日本人大学生設問なし）	174
表 7.8.1	選択肢時の聴解行動のヴァリエーション④ （日本人大学生設問ありと設問なしの比較）「確定」と「確信・確定」	175
表 7.8.2	選択肢時の聴解行動のヴァリエーション④ （留学生と日本人大学生設問なしの比較）「確信・確定」	175
表 7.8.3	選択肢時の聴解行動のヴァリエーション④ （日本人大学生設問ありと設問なしの比較）「不信」と「吟味・不安」	175
表 7.8.4	選択肢時の聴解行動のヴァリエーション④ （留学生と日本人大学生設問なしの比較）「葛藤・吟味」と「吟味・不安」	176
表 7.9	留学生、日本人大学生の個人的特性と話題の知識の違い	178

図 一 覧

図 1.1	日留試の測定対象能力	5, 31
図 2.1	意思伝達言語能力の構成要素	15
図 2.2	標準得点の正規分布	16
図 2.3	旧日能試 聴解テストで測定する能力	21
図 2.4	課題遂行のための言語コミュニケーション能力測定（新日能試）	24
図 2.5	日留試の測定対象能力	29
図 3.1	ACT-Model	39
図 3.2	言語テスト得点の変動要因	45

図 4.1	日留試聴解テスト・アイテムの構成	52, 100
図 4.2	留学生のコミュニケーション能力	54
図 4.3	日留試聴解テスト・アイテムの分析例	57
図 4.4	場面の割合 (2002-2013)	60
図 4.5	場面の割合の推移 (2002-2013)	60
図 4.6	聴解テスト・アイテムのテキスト・バンク (2002-2013)	73
図 4.7	聴解テスト・アイテムのテキスト・バンク キャンパス場面・人数 (2002-2013)	74
図 4.8	聴解テスト・アイテムのテキスト・バンク 私的会話・キャンパス場面・人数・機能 (2002-2013)	75
図 4.9	聴解テスト・アイテムのテキスト・バンク 私的会話・キャンパス場面・人数・機能 鳥瞰図 (2002-2013)	75
図 5.1	スクリプト・マーキング例	101
図 6.1	SCQRM の構造モデル	119
図 6.2	分析の手順	123
図 6.3	留学生の聴解行動 (設問ありのアイテムに対する行動)	125
図 6.4	日本人大学生の聴解行動 (設問ありのアイテムに対する行動)	139
図 7.1	留学生の聴解行動 (設問なしのアイテムに対する行動)	155
図 7.2	日本人大学生の聴解行動 (設問なしのアイテムに対する行動)	166

1章 序論

1.1 本研究の背景

本節では、本研究の背景として、(1) 日本留学試験テスト・アイテム¹の基礎的研究が少ない (2) アイテム・ライター制度の認知度が低く、アイテム・バンクの研究が少ない (3) 日本留学試験はアカデミック・ジャパニーズに特化したテストであるが、アカデミック・ジャパニーズに関する定義が曖昧である (4) 日本留学試験における聴解領域の研究が少ない の4点にまとめて記述する。

(1) 日本留学試験テスト・アイテムの基礎的研究が少ない

日本語を母語としない人々の日本語能力を測定する大規模な日本語能力テストは、1984年に始まった「日本語能力試験」(Japanese Language Proficiency Test ;以下「日能試」)があげられる。日能試とは、一般的な日本語能力を測定・認定する一般言語能力テスト (test for general language abilities) である。しかし、学習者の多様化に伴い、ある特定の目的のための日本語能力 (Japanese for specific purposes) を測定するテストが必要になり、1996年に「ビジネス日本語能力テスト」(BJT: Business Japanese Proficiency Test) ²が始まった。そして、日本の大学に留学するための日本語能力テストとして 2002年、「日本留学試験」(Examination for Japanese University Admission for International Students ;以下「日留試」)が実施され始め、その後、日能試が2010年に大幅に改定され、同年、日留試も小改定された。現在、日本語の大規模テストとしては、日能試と日留試が代表的である。本研究では、2009年までの日能試を「旧日能試」、2010年以降の日能試を「新日能試」とし、双方のテストを指すときは「日能試」と呼称する。

日留試ができるまでは、日本の高等教育機関は旧日能試の結果の提出を留学条件に挙げていたが、2002年以降は積極的に日留試に変える大学が増えてきた³。日留試ができたことで、日本語教育に関わる人々や学習者は、テストで何を測定するか、テストの結果が何を示しているかについて、意識をするようになった。例えば、日能試では「文字・語彙・文法」という言語知識の量を測定しているが、日留試は測

¹ カタカナ語は元来の意味と日本語のみで使用される意味の両方が存在する場合があります、混乱を避けるために、本研究では主要語彙の「テスト・試験」と「アイテム・問題」について以下のように定義する。

①テスト (test) ・試験

一般的に英語の「test」の訳は「試験」と「テスト」の2種類が使用されている。前者の例は「Japanese Language Proficiency Test : 日本語能力試験」、後者の例は「Language test : 言語テスト」である。評価の研究分野においては、英語では「test/testing」が使用されており、英語の専門書の和訳はほとんどが「テスト/テストング」である。よって、本研究では固有名詞として既に定着している「日本語能力試験」のような語彙はそのまま「試験」を使用し、それ以外はテストを使用することにする。例えば、「Japanese Language test : 日本語テスト」「listening test : 聴解テスト」のような使い方である。

②アイテム (item) ・問題

日本留学試験ではテスト問題をテスト・アイテムと呼ぶ。本研究では日本留学試験以外のテストについて話すときも、混乱を避けるため、問題という言葉は使わずアイテムを使用する。

また、本研究では多くの参考文献が英語であったため、邦訳のある参考文献はなるべくそのままの日本語訳を使用した。ただし、訳者によって同じ語彙の訳が異なる場合は、その際には日本言語テスト学会が編集した『日本言語テスト学会言語テストング用語集 日本言語テスト学会テストングの実施規範』(2006年)の用語集を参考に、統一した。なお、稿者が訳したものもあるため、日本語訳で誤解が生じないよう、なるべく英語を()内に並記するようにした。

² BJTは公益財団法人日本漢字能力検定協会が実施している。独立行政法人国際交流基金が実施する日能試、独立行政法人日本学生支援機構が実施する日留試と比較すると実施機関としての公的性は低いが、特定の日本語能力を測定するテストとしての例としてあげている。

³ 日留試は主要アジア諸国とウラジオストク、香港(中国本土は除く)で受験することができるがその他の地域では実施されていない。そのため、2015年現在でも、日留試が行われない地域の留学希望者は、日能試あるいは他の日本語テストの結果を留学時の日本語能力の資料として用いている高等教育機関も多い。

定しない。旧日能試は配点に重みづけ⁴があり、素点を用いて結果としていた⁵。一方、日留試と新日能試は統計処理をした点数⁶を出しており、時代の流れとともに、テストの結果にはより信頼性のある統計処理が使用されるようになった。日能試は決められたゴール（級・レベル）の合格点を超過すれば合格するテストであるが、日留試は決められた目標到達点があるわけではなく、受験者全員の中での位置を表すテストである⁷。日留試も新日能試も点数を統計処理するが、この2つのテストは目的が異なり、テストの設計や尺度の解釈も違う。しかし、旧日能試に慣れてきた日本語教師や学習者、テストの結果を利用する人々のあいだには、日留試が、対象受験者を留学生に限定している以外は日能試と同じテストではないかという誤解もあり、2つのテストの違いは一般にはわかりにくいようだ。

日留試が始まった当初、アカデミック・ジャパニーズ・グループ⁸が中心になって、日留試のテスト・アイテムを対象にした研究を活発に行っていたが（門倉 2003b、2005）、その後、日留試のテスト・アイテムを対象にした研究は減少した。また、旧日能試は毎年、テストの結果やテスト・アイテムの内容分析を行い、『分析評価に関する報告書』（国際交流基金・日本国際教育支援協会）を公開してきたが、日留試は毎回実施されたテスト問題冊子のなかの1フォームのテスト⁹を公開しているのみである。このような背景から、日本語の大規模テストの代表の1つと言える日留試は依然として認知度が低く、これに関連する研究もほとんど進んでいない。日留試は2010年に一度小改定が行われており、今後大きな改定が行われる前に基礎的研究を進めることが急務である。

（2）アイテム・ライター制度の認知度が低く、アイテム・バンクの研究が少ない

日留試が行なった新しい試みとして、アイテム・ライター（item writer）制およびアイテム・バンク（item bank）の構築とその利用がある。アイテム・ライター制を導入したことで、多くの日本語教師や学校教師が、大規模テストのテスト・アイテムを作成する機会が得られたことは、日本語教育の発展において望ましいと言える。多くの日本語教師は、自らが携わる日本語教育現場で、プレースメント・テストや到達度テストを作成あるいは使用した経験はあるが、日留試のような大規模テストは、特定のカリキュラムの教育効果を見るためには適していないため、教師が触れる機会は少ない。しかし、アイテム・ライターとして、今まで経験できなかった大規模テストの作成に携わることで、テスト作成の方法やスキルを体系的に体感できるようになった。

アイテム・ライター制は多くの作成者が関わることで、広範囲なトピックや多種多様な視点からアイテムが作られるという利点もあるが、すべてのライターがテストの目的やテストで測定する能力につい

⁴ テスト・アイテムによって点数を変えることを重みづけという。例えば、1番は1点、2番は3点というような配点である。日能試は2010年に改定され、新日能試からは重みづけのあるアイテムはなくなった。

⁵ 素点とはテストの配点を合計した数値であり、換算をしていない点数である。例えば、1アイテム1点で50アイテム出題されたテストで50点をとった「50点」は素点である。同じ能力の受験者が受験した場合、素点では難しいテストでは低く、易しいテストでは高く点数が出てしまう。

⁶ 統計処理の点数については2章の標準得点の正規分布（p.16）で詳述する。

⁷ 日能試と日留試の違いについての詳細は2章（p.19）で記述する。

⁸ 日本語教育学会が、日本語教育に関連ある領域のテーマについての研究活動を推進するために、「日本語教育学会テーマ領域別研究会活動に関する規程」を設けた際、テーマ領域別研究会として、「アカデミック・ジャパニーズ・グループ」（AJG）が結成された。主に、高等教育機関における日本語教育の研究をすすめているグループである。

⁹ 日留試は、同時に内容の違うテスト問題冊子を複数作成している。それは日留試が世界各地で実施されるため、テスト実施時間に時差が生じ、テスト内容の漏洩の問題や、セキュリティ対策のためである。その1冊1冊をフォームとここでは呼ぶ。日留試は異なるフォームの得点比較を可能にする得点等化（等化についての詳細は第2章 24.p.25の脚注33を参照）を行っているため、テストアイテムの公開をしないのが原則であるが、受験者や日留試結果を利用する機関などの便宜のため、1フォームのみ公開している。

て統一した見解を持っていなければならないという難点もある。質のよいテスト・アイテムを作成するためにはテストで何をどう測定するのか、そのためのアイテムの構成はどうあるべきかなどの議論が十分に行われ、そこでの合意内容がアイテム・ライター全員に浸透していなければならない。Hughes (2003 邦訳 pp.64-65)は、アイテム・ライターは自分で作成したアイテムを客観的に評価し、批判に対しては柔軟に対応すること、そして、母語話者¹⁰に近い言語能力、知性¹¹、想像力（アイテムの文脈を作り出す能力、誤解される危険性を予測する能力など）が必要だと述べている。特に聴解テスト・アイテム作成には、知性および想像力が不可欠である。なぜなら、日留試の聴解テスト・アイテムは自然談話をそのまま使用してはいないため、アイテム・ライターは談話そのものを作り出さなければならないからである¹²。また、聴解テスト・アイテム作成の難しい点は他にもある。それは作り手（アイテム・ライター、テスト実施機関）の視点と受け手（受験者）の視点が異なっていることである。作り手は、聴解の談話スクリプトを文字で書き、それを視覚情報で捉える。作成途中および最終的には聴覚で確認する作業が行われるはずだが、最初は視覚に頼って作成する。この受け手との視点の違いがテストそのものに及ぼす影響があるのではないかという疑問から、本研究では視覚情報と聴覚情報の双方から聴解テスト・アイテムの分析を試みることにした。

多大な時間と労力を使って作成されたテスト・アイテムは、1度使用された後、問題がなければアイテム・バンクに登録される。テスト・アイテムを再利用することで、アイテム作成の労力を大幅に減らすことができ、テストの質があがり、信頼性の確保が確かなものとなる (Hughes, 2003 邦訳 p.250)。アイテム・バンクは既に TOEFL や IELTS などの大規模な言語テストで使用されているが、日本語テストでは日留試が初めての導入であり、これまで研究の蓄積はほとんどない。日留試は1年に2回実施され、複数のフォームを使用していることから、膨大な数のテスト・アイテムが存在する。その管理のためには、テスト・アイテムのラベル化が必要である (伊東 2008b, p.5)。ラベルとは個々のアイテムを特徴づけるもの (属性 properties) のことである (Downing & Haladyna, 2006 邦訳 p.284)。ラベルはあるアイテムを探そうとしている人にとって論理的であり、その分類は特定の課題のために構想されなければならない、すべての分類システムはその時の判断や知識の変化に伴って常に変化する (同掲 pp.289-290)。例えば、日留試の聴解テスト・アイテムでは、テスト・アイテムの内容 (設問、本文、選択肢の文字情報と音声

¹⁰ 本研究では、「ネイティブ、ノンネイティブ」を使用し、「母語話者、非母語話者」を避ける。大平 (2009, pp.85-110) は、「ネイティブ」の定義を、①時間説：当該言語との接触開始時が誕生あるいは幼児期であること、②能力説：当該言語の運用力が有能であること、③理想説：完全な能力を有する理想的な話者であること、の3つが現在までの説であり、それが混同されていると述べている。大平は、Hymes, D. (1972)の「もしもある言語のネイティブスピーカーという概念を認めれば、言語に固有の、また唯一の正しい規則があることを認める事になってしまう」と、Bachman (1990)の「ネイティブスピーカーは言語使用能力の全ての側面で誰もが同じようにうまく振る舞えるわけでも、また誰もが同じようにノンネイティブよりよく振る舞えるわけでもないことが徐々に明らかになりつつある」という言葉を引用し、「人はネイティブ・ノンネイティブ『である』だけでなく、ネイティブ・ノンネイティブ『になる』のである」とまとめている。本研究では①時間説にのっとり、日本人大学生を「ネイティブ」とした。日本人大学生は、誕生時もその後も日本語社会の成員であるネイティブであるが、それは彼らの日本語が完璧であり理想的であると保証するものではない。なお、参考にした文献で「母語話者、非母語話者」を使用している場合は、この定義とは限らないため、そのまま「母語話者、非母語話者」とした。

¹¹ ここでいう「知性」とは「世界知識」を指していると考ええる。

¹² 中級以上のレベルの場合、読解テスト・アイテムの問題本文 (読解文、テキスト) は既に印刷・出版されたものを使用することが多いため、アイテム・ライターは質問文および選択肢などは作成するが、本文は作らないことが多い。しかし、聴解はすべてのレベルにおいて、本文談話を作成している。

情報)、正答、解答の理論的根拠、困難度¹³、識別力¹⁴、使用履歴などが分類項目と推測される。本研究では、日留試の聴解テスト・アイテムの分析結果から、さらにどのような分類の可能性が考えられるかについて考察する。

(3) 日本留学試験はアカデミック・ジャパニーズに特化したテストであるが、アカデミック・ジャパニーズの定義が曖昧である

日留試の日本語シラバス (日本国際教育協会 2003, p.244) には、このテストの目的を「外国人留学生として日本の高等教育機関、特に大学学部で留学を希望する者が、日本の大学での勉強に対応できる日本語力 (アカデミック・ジャパニーズ) をどの程度習得しているかをシングルスケールで測定すること」と記している。「アカデミック・ジャパニーズ」の定義についての議論は、日留試がこの用語を使用したことから始まり、研究者間でも狭義な解釈と広義な解釈があり、言語のみに関わるのかどうかなどに見解の相違が見られる (嶋田 2005, p.82、森 2005, p.121、門倉・筒井・三宅 2006) ¹⁵。英語教育では「academic listening、academic writing」などが本の題名にもなっており、既に「academic」の用語は定義され、広く浸透している。ここで使用されている「academic」という用語は、高等教育機関における活動に限定して用いられ、言語活動の面では講義 (lectures) 活動に焦点を絞って使用される場合が多い (Richards 1985、Flowerdew (ed.) 1994)。これは狭義の「アカデミック」の定義と重なる。それに対して、日留試は「日本での留学生活を送る上で、日本語によるコミュニケーション能力があるかどうか、また、自国での初等・中等教育修了までに習得した知識を前提としながら、日本の大学で学習・研究活動を行うための日本語能力があるかどうかを測定する言語テストであり、かつ標準テストである」(「日本留学のための新たな試験」調査研究協力者会議 2000, p.12) と位置づけられており、ここで言及されているアカデミック・ジャパニーズは、狭義の「アカデミック」が意味する講義活動の範囲における日本語だけを指すものではない。図 1.1 (次頁) は日留試の日本語シラバスの中にある測定対象能力について説明する図であり (日本国際教育協会 2003, p.245)、円の中の網掛けの部分測定することを目的としている。これを見ると、大学外の生活場面に必要な言語能力、社会適応能力や事務的処理能力、社会知識や専門知識など、言語以外の能力についても言及されている。日留試は日本の大学の留学生が遭遇するあらゆる場面が必要となる日本語能力、スキルおよび知識の測定を目的としていることがわかる。これは広義のアカデミックの定義である。しかし、日留試が目指すアカデミック・ジャパニーズが、どのような日本語運用能力かについての具体的な記載はない。

¹³ テスト関連で頻繁に使用する「困難度 (difficulty)」について本研究での扱いを説明する。「difficulty」の邦訳、あるいは日本語で書かれた文献には「難易度」と「困難度」が使用されていた。本研究では『日本言語テスト学会言語テスト用語集』(2006年) にならぬ「困難度」を使用するが、参考とした文献で使用している用語を尊重し、「難易度」を使用している文献を引用する場合には難易度の後ろに (困難度) をつけることにする。

¹⁴ テスト関連で頻繁に使用する「識別力 (discriminability, discrimination)」について本研究での扱いを説明する。「discriminability, discrimination」の邦訳、あるいは日本語で書かれた文献には「弁別力」「弁別度」「識別力」「識別度」が使用されていた。『日本言語テスト学会言語テスト用語集』(2006年) で、discriminability は「弁別力」、discrimination は「弁別」となっているが、『テスト作成ハンドブック』(2008年) では、discrimination は「識別力」と訳されている。参考にした文献および邦訳では新しい年代になるにつれ、「識別力」が多く使用されていた。よって、本研究では「識別力」を使用する。ただし、参考にした文献で使用している用語を尊重し、「弁別力」「弁別度」「識別度」を使用している文献を引用する場合には、各用語の後ろに (識別力) をつけることにする。

¹⁵ 「アカデミック」の定義についての詳細は3章「3.1 アカデミック・リスニング」(p.36) で述べる。

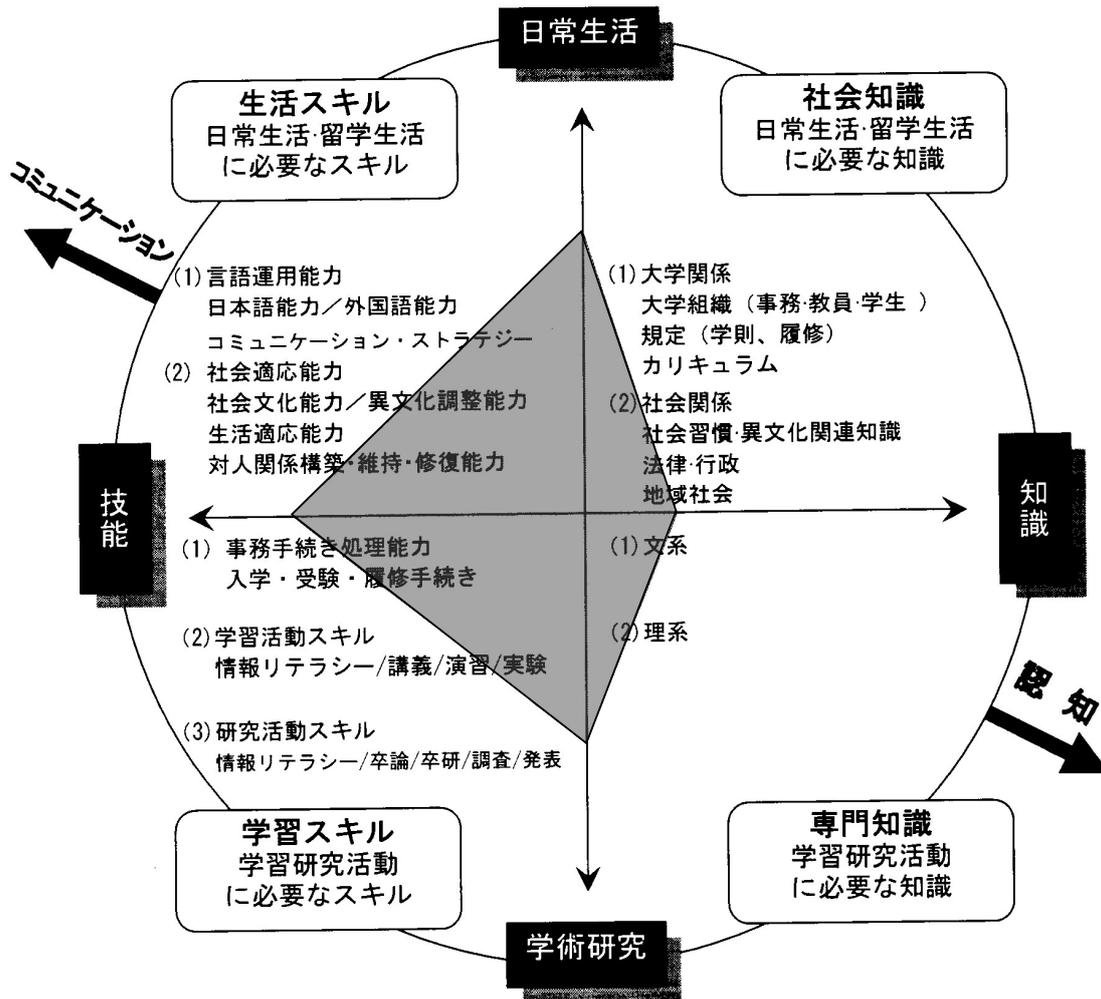


図 1.1 日留試の測定対象能力 (日本国際教育協会 2003, p. 245)

日留試は 2002 年から実施され、1 年に 2 回、6 月と 11 月に実施されてきた。初回から 16 回の実施を終えたあと、2010 年の 6 月からはシラバスの記載方法と領域構成が改定された (新シラバス)。以下に改定前と改定後の「試験の目的」を記す (日本学生支援機構¹⁶2010, p.290)。

¹⁶ 日留試の事業は日本国際教育協会が行っていたが、2004 年 4 月 1 日に留学生交流事業、留学生に対する奨学金関連、生活調査などの事業を整理・統合し、学生支援事業を総合的に実施する日本学生支援機構が誕生した。本研究では、引用元が以前の日本国際教育協会の場合はその名称を使用し、日留試の製作・実施団体について述べるときは日本学生支援機構とする。

「試験の目的」

2002年（改定前）

この試験は、外国人留学生として日本の高等教育機関、特に大学学部留学を希望する者が、日本の大学での勉学に対応できる日本語力（アカデミック・ジャパニーズ）をどの程度習得しているかをシングルスケールで測定する事を目的とする。

2010年（改定後）

この試験は、日本の高等教育機関（特に大学学部）に、外国人留学生として入学を希望する者が、大学等での勉学・生活において必要となる言語活動に、日本語を用いて参加していくための能力をどの程度身につけているか、測定することを目的とする。

新シラバスでは、日留試のシラバスの発表当初から議論されてきた「アカデミック・ジャパニーズ」という用語は使用されなくなり、「日本の大学での勉学に対応できる日本語力」が「大学等での勉学・生活において必要となる言語活動に参加するための能力」に置き換えられ、「生活」が付加されている。2002年当初の文言には「生活」がなかったが、日留試の測定対象能力（図 1.1 前頁）には、「日常生活」が含まれており、改定後の目的の文言のほうが、日留試の測定している範囲を正確に表している。文言が変更されてはいるが、測定対象能力は図 1.1 の網掛けの範囲で示された内容であり、テストの目的は基本的に変わっていないと考えてよいだろう。日留試のシラバスから「アカデミック・ジャパニーズ」という用語はなくなったが、既に日本語教育界では使用が定着しているため、本研究でもこの用語を使用する。

また、新シラバスの各領域の概要¹⁷（日本学生支援機構 2010, p.291）では、読解・聴解・聴読解のテストで測定する能力が3つ挙げられている。

1. 直接的理解能力： 例えば、内容を正確に把握する、主題主旨を的確に捉える
2. 関係理解能力： 例えば、重要部分を抽出する、情報の関係を理解する、異なる媒体の情報を比較対照する
3. 情報活用能力： 例えば、内容から結果や帰結を導き出す、具体的事例を一般化する、一般論を具体事例と照合する、異なる媒体の情報を相補的に組み合わせる

これらの能力の大学生レベルがどの程度であるのか、留学生に期待される能力とは具体的にどのような能力であるのかは、日留試のテスト・アイテムを分析することから見てくると考える。

（4）日本留学試験における聴解領域の研究が少ない

日留試で測定される言語技能は、「①聞く ②話す ③読む ④書く ⑤翻訳する」の5つであるが、「②話す ⑤翻訳する」は部分的ないし間接的測定とされている。テストは「聴解・聴読解・読解・記述」の4つの領域¹⁸に分かれている。2002年当初から2009年までは「聴解」「聴読解」「読解」それぞれが20アイテムずつ、同一の重み付けで出題され、点数配分は「聴解」120点、「聴読解」120点、「読解」160点の、計400点であった。そして、「記述」はこの得点とは別に、文法的能力と論理的能力それぞれ3点ずつ、計6点で示された。2010年の改定後、「読解」200点、「聴解」と「聴読解」合わせて200点と読解の比重が増え、記述は50点満点で、得点では9段階、レベルでは6段階で示されるようになった

¹⁷ 詳細は参考資料1「日留試日本語シラバス各領域の概要」（p.192）を参照。

¹⁸ 日留試では「聴解・聴読解・読解・記述」を「領域」としている。

¹⁹。比較材料として旧日能試を挙げる。旧日能試は「文字・語彙・文法・聴解・読解」の5つの領域で、点数配分は「文字・語彙」100点、「聴解」100点、「文法・読解」200点で計400点であった。日留試と旧日能試の点数配分から、日留試は旧日能試に比べ聴解能力に重点をおいていると言える。これは、日留試が測定しようとしている「生活上の日本語によるコミュニケーション能力」また「大学で学習・研究活動を行う日本語能力」では、聴解能力が重視されていることを示している。當作（1988, p.59）は従来の外国語教授法は話す能力に重点を置き、聞く能力を付随的なものとして位置づけてきたが、最近では聞く能力が話す能力につながるとする教授法も出現したと述べ、聴解能力の重要性を示唆している。また、聴解スキルがコミュニケーション活動の中で最も高い割合を持つこと（竹蓋 1984, p.5、Benjamin 1986 邦訳 p.3）、かつ聴解テストが他のスキルのテストと非常に高い相関があること（武井 2002, p.59、庄司他 2004, p.50、日本語能力試験実施委員会 2004, p.78）から、聴解能力が言語能力のなかで重要であることは明確である。アカデミック・ジャパニーズ・グループの研究報告（門倉 2003b、2005）では、21件の研究報告がなされているが、スキルや領域に関する研究は、作文が2件、読解が2件、語彙が1件であり、聴解のみに関しての研究報告は皆無である。そして、アカデミック・ジャパニーズ・ジャーナルが2009年から6回発行されているが、2015年現在、聴解のみに言及している論文・報告は1件（足立 2014、筆者自身の投稿）しかない。これらのことから、聴解研究の必要性が高いことは明白であり、本研究では「聴解」領域に焦点を当てることにした。

1.2 本研究の目的と意義

本研究では日留試の聴解テスト・アイテム分析を行う。この研究から日本の高等教育機関が留学生に期待している日本語聴解能力が見えてくるのではないかと考えたからである。この研究結果によって日本への留学を考えている学習者、既に来日している就学生、日本語聴解教育に関わる人々に、具体的な学習目標を提示することができる。加えて、聴解テスト・アイテムの特徴を分析することで、アイテム・ライターがテスト・アイテムを作成する際に参考になるであろうテキスト・バンク²⁰構築の1つの具体例を提案したい。大規模テストにおけるテスト・アイテム作成は大量に、継続的に作成しなければならない。アイテム・ライターの力量だけに頼っていても限界はある。そのためにテキスト・バンクの活用を推奨したい。本研究ではテスト・アイテムを多角的に分析することによって、テスト・アイテムの持っている特徴を立体的に示すことを目指す。

聴解テスト・アイテムは、普段、日常の中の聴解活動とは著しく異なっている。例えば、聞き返しができる双方向コミュニケーションや、話者の非言語情報、談話がされている環境などの視覚情報がない。また、談話が起きた背景、話者との関係などの情報も希薄である。そして、テストということで課題が出され、その答えを選択肢から選ばなければならない。聞かされる音（声）は、滑舌がよく、スピード・コントロールされた、ノイズのない談話である。では、日本語学習者が授業で聴解練習をする場合はどうだろうか。課題が与えられ、CDなどからコントロールされた談話が流れる。解答方法は選択肢の場合もあれば、自由解答の場合もあって一律ではないが、いずれにせよ、答えなければならないというプレッシャーはかかる。つまり、日本語学習者がよく行う聴解はテスト聴解に近いのである。このことより本研究においてテスト聴解の過程を分析することが、聴解テスト・アイテム作成や、日本語教育現場の聴解練習、聴解学習にも貢献できると考える。

¹⁹ 改定前と改定後の詳細については日本学生支援機構のHP、<http://www.jasso.go.jp/eju/kaitai.html>（2015年3月）を参照。

²⁰ テキスト・バンクはアイテム・ライターがアイテムを作成する際に、トピック、場面、談話展開などの談話構成の参考にするためのものである。テキスト・バンクについては4章にて詳述する。

以上のことから、次の3つを本研究の目的とする。

- 1 日本留学試験聴解テスト・アイテムの特徴を明らかにする。
- 2 日本留学試験聴解テストが留学生に期待している聴解能力を、アカデミック・リスニングの視点から明らかにする。
- 3 聴解テスト時、受験者はどのような聴解行動²¹をとっているのかを明らかにする。

これらを研究目的とした背景には、日留試が留学生にとって大学への入り口テストであるという位置づけがある。「留学生 30 万人計画」や「アジア人材資金構想」では、留学生が卒業・修了後に日本で就職することを構想している。西原 (2011, pp.7-10) は日本への入り口テストとして日留試を挙げ、出口テストとしての日能試と BJT ビジネス日本語能力テストにつないでいく必要を述べている。留学したい学習者は日留試を、日本語能力を試したい学習者は日能試を、日本の企業で働きたい人は BJT を別々に受験するというように、それぞれが独立した目的のために存在しているテストではなく、日留試を受けて大学のカリキュラムを修了したのちに日能試または BJT へとつながっていくようなテストの位置づけがなされているのである。菅井 (2004, p.40) は日留試が測定している能力は受験時だけではなく、その後半年から1年後のことも含めていることから、テストが測定しているのは将来の能力であると述べている。将来とは、日本の大学に入学したときから卒業までのことを指している。アメリカやカナダの大学に入学する際、留学希望者は TOEFL などの英語テストの他に、GRE (Graduate Record Examination) の受験が義務づけられていることがある。GRE は言語 (大学院教育レベルの読解力・論理的思考能力・単語力、学術的な文章を英語で書く能力)、数学、記述 (Analytical Writing) の3つのセクションに分かれており、英語ネイティブも受験するテストである。日留試も日本語のテスト以外に「理科、数学、総合科目」があり、GRE と機能が似ている。日留試と GRE に共通しているのは、大学生活に必要な知識やスキルを測定することであり、それは「テストで出題される知識やスキルを準備してきなさい」という大学からの発信なのである。日留試は、留学生にその知識とスキルの準備をさせることで、充実した豊かな大学生活を送れることを約束する。本研究では、3つの目的を明らかにすることで、日留試が日本の大学生活を送る際に必要な日本語能力を測定しているかどうかについて検討したい。

本研究では公開されている日留試の聴解テスト・アイテムを使用するが、テストの信頼性、テスト・アイテムの識別力パラメーター、困難度パラメーターの数値は公開されていない。本研究では、テストの信頼性、妥当性などの検証を目的としてはおらず、基礎的研究として公開されているアイテムを分析し考察するにとどめている。

1.3 論文の構成

本稿の1章から3章で、本研究に関する背景およびテストの概観をまとめる。1章では、研究の背景と意義について述べ、そこから本研究の目的を提示した。2章では、言語テストの定義およびテストの変遷をまとめ、現在コミュニケーション能力が注目されていることと、それをテストでどのように測定しようとしているのかについて述べる。また、日本語の大規模テストとして代表的な日能試と日留試の違いを、特に聴解テストにおいて比較することにより、日留試が定義しているアカデミック・リスニングの骨格を明確にする。3章では、「アカデミック」の概念について詳述する。その後、聴解テスト研究

²¹ 日常的に行われる聴解活動と聴解テスト時に行われる聴解を区別するため、この本研究では、「聴解行動」とした。詳細は6章「6.1 テスト時の聴解行動」(p.118)を参照。

の先行研究を概観し、本研究で使用する分析方法として聴解下位スキルが有効であることを示す。また、聴解テスト形式がテスト得点に影響を与えることについて、本研究では課題の有無に注目してその理由を述べる。

4章から7章で、前述した3つの研究目的について論ずる。

4章では、日本留学試験聴解テスト・アイテムの特徴を明らかにする。日留試の聴解テスト・アイテムの分類および分析を行い、基礎的なデータをまとめる。その結果から、聴解テストのテキスト・バンク構築のための1モデルを提案する。5章では、日本留学試験聴解テストが測定している聴解能力をアカデミック・リスニングの視点から明らかにする。日留試の聴解テスト・アイテムを分析して、日留試が測定しようとしている「アカデミック・リスニング」について明確にし、留学生にとって必要性が高い聴解テストとは何かを考察する。6章、7章では、聴解テスト時、受験者はどのような聴解行動を行っているのかを明らかにする。そのために、留学生の聴解行動だけではなく、既にアカデミック・リスニングのレベルを保持していると想定されている日本人大学生²²との比較・分析をする。そして、聴解テストの得点に影響を与える要因として、課題の有無に注目し、課題のあるアイテムと課題がないアイテムの2つの聴解行動を比較する。

最終章で、3つの研究の結果を総合的に考察し、結果を記述する。この研究の研究意義と限界について触れ、最後に今後の課題をまとめる。

表1.1(次頁)は4章、5章、6章、7章の研究概要をまとめたものである。4章と5章は聴解テスト・アイテムのスク립ト分析であり、6章と7章は聴解行動分析である。

²² 日本語母語話者の大学生。

表 1.1 研究の概要

研究の目的	章	データ収集法	対象	分析方法
日本留学試験聴解テスト・アイテムの特徴を明らかにする。	4章	2002年度から2013年度までに公開された、聴解テスト・アイテムのスク립トを使用。	聴解テスト・アイテム440項目	テスト・アイテムを以下の8つの項目で分類 ①理解の種類 ②場面 ③形態 ④トピック ⑤人数 ⑥コミュニケーションの目的 ⑦機能 ⑧情報を得る理由
日本留学試験聴解テストが留学生に期待している聴解能力を、アカデミック・リスニングの視点から明らかにする。	5章	2002年度から2006年度までに公開された、聴解テスト・アイテムのスク립トを使用。	聴解テスト・アイテム200項目	リスニング下位スキルを以下の言語能力の5つの構成要素で分類 ①文法的知識 ②テキスト ²³⁾ についての知識 ③機能的知識 ④社会言語学的知識 ⑤方略的能力 その後、スク립ト・マーキングによるリスニング・スキル分析と選択肢の分析
聴解テスト時、受験者はどのような聴解行動をとっているのかを明らかにする。 (設問 ²⁴⁾ の有無による受験者の聴解行動に注目)	6章	公開された中から3アイテムを選出(設問あり)。 半構造化インタビューによる発話思考法と発問法の併用。	留学生(大学生、大学院生)21名、 日本人大学生9名	構造構成的質的研究法を使用し、概念モデルを以下の4つ作成 ①留学生設問あり ②日本人大学生設問あり ③留学生設問なし ④日本人大学生設問なし
	7章	公開された中から3アイテムを選出(設問なし)。 半構造化インタビューによる発話思考法と発問法の併用。		

²³ 「text」の訳として、「テキスト」と「テクスト」がある。本研究では『Longman Dictionary of Applied Linguistics』(1985)と『英語教育用語辞典』の訳である「テキスト」を使用する。ただし、引用の場合、訳者を尊重し「テクスト」を使用している場合はそのまま引用する。

²⁴ 設問とは日留試の聴解テスト・アイテムの本文の前にある場面・状況説明とブリタスクのことである。詳細は4章「4.1 基礎的データのための分析枠組み」p.52を参照。

2章 言語テストの概観

2.1 言語能力の定義と言語テストの変遷

言語テストは言語能力を測定するため、言語能力の定義によりテストの目的や方法が違ってくる。また、教授法や指導法と関係が深い。それは、どのような言語能力を、どのように教えたかによってテストで何を測定するのかが決まるからである。言語能力と言語テストの推移については、研究者により使用する用語や概念などで多少違う分類がされている。Brown (1996 邦訳 pp.28-31) は、前科学時代、心理測定的・構造言語学的時代、統合的社会言語的時代の3つに分け¹、教授法との関わりから分類した。Brown は、3分類した後、今後は学習者にタスク²を課し、実生活での予測不可能なやりとりに近づいていくだろうと推測している。McNamara (2000 邦訳 pp.24-29) と菅井 (2006b, pp.3-6) は、テスト内容により個別的要素テスト、統合的語用論的テスト、コミュニカティブ・テストの3つに分けている。Heaton (1988 邦訳 pp.16-29) は、テスト方法によって論文・翻訳による方法、構造主義の方法、統合的な方法、コミュニカティブな方法の4つに分け、Bachman (1990 邦訳 pp.4-7) は、言語能力を中心に技能と個別要素、語用論的予測能力、意思伝達言語能力を測定するテストと、3つに分けている。これらを参考に、言語テスト研究の経緯から時期的、内容の特徴などで以下のように4分類にまとめた。

(1) 前科学時代 論文・翻訳による方法 翻訳テスト

前科学的時代は文法訳読法³と結びついており、現場の教師が文法の知識や翻訳能力を測定するために、訳読や自由作文などを自ら作成していた。テスト作成のための特別な技術や専門知識は必要なく、テストの多くは論文、翻訳、文法問題から構成されていた。また、テスト内容は文学的、文化的側面に偏っていた。このテストの問題点は、テストを作成した教師が主観的に採点をしていたことである。しかし、教師の主観的判断による評価という問題点があるにもかかわらず、今なおこの指導法とテストは行われている。

(2) 心理測定的・構造言語学的時代 構造主義の方法 個別的要素テスト (Discrete point tests)

心理測定的・構造言語学的時代になると、客観性、信頼性、妥当性に関心が起こり、言語テストは急速に科学的になった。1960年代にLado (1961) や Carroll (1961, pp.30-40) が提案した技能と構成要素(参考資料2「個別的要素テストの評価項目」p.194参照)を中心とした個別的要素テストの時代は、言語学の構造主義と心理学の行動主義の研究に基づいている。言語は、音素、文法項目、語彙で細かく分解されるが、それらはあるルールにそった構造を持っているという考えからテスト開発がされた。テストは、音素、文法項目、語彙の測定単位がテキストの中で、あるいはテキストから切り離されて問われる。多肢選択法が使用され、科学的測定の対象となり、統計的処理が行われるようになった。また、技能は、「話す・聞く・読む・書く」というマクロ言語技能として厳密に分けて実施された。オーディオ・リング法³の全盛期である。このテストへの信頼は高く、今現在もインタビューなどの評価者間信頼性が問題視

¹ Brown (1996 邦訳 p.28) は言語テストの動向を言語教授法と結びつけ、Spolsky (1978, Introduction: Linguistics and language testers, Spolsky (ed.) *Advances in language testing series: 2*, Arlington, VA: Center for Applied Linguistics) と Hinofotis (1981, Perspectives on language testing: Past, present and future. *Nagoya Gakuin Daigaku Gaikokugo Kyoiku Kiyo*, 4, pp.51-59) が分けた3つの時代を用いて説明している。

² 「タスク」と「課題」は同義語であるが、参考とした文献での使用語彙に従った。

³ 文法訳読法は19世紀に現代語の教授法としてヨーロッパで広く普及した。(Richards 他 1985 『ロングマン応用言語学用語辞典』邦訳 p.161)

されるものよりは安易に作成でき、評価も単純であることから多く使用されている。

(3) 統合的・社会言語学的時代 統合的な方法 統合的・語用論的テスト (Integrative and pragmatic tests)

統合的・社会言語学的時代には、コミュニケーション能力は言語の文法を操るだけでなく、様々な状況に応じて言語を使用できる能力であると考えられた。このときのテストとしては、クローズ・テスト (cloze test)⁴とディクテーションがある。この方法のよい点は、構造言語学的時代には異を唱えられた心理測定的測定法を支持したことである。これは後に、語用論が関わる統合的測定テストの意義に発展した。

1970年代、コミュニケーションな教授法への移行や外国人留学生の増加に伴い、実践的かつ統合的なパフォーマンスをテストする必要が出てきた。その際、文脈の理解と言語の体系的な知識(発音、文法、語彙など)を測定するためのテストとして、面接やある程度の長さをもった作文テスト、読解テストなどが行われた。これが、統合的テスト⁵である。しかし、実施に時間がかかり、採点が難しいなどの問題が挙げられた。これを解消したのが Oller (1979) である。Oller は、細かく分解された言語知識の集合が言語能力であるという考えに反論し、語用論的予測文法 (pragmatic expectancy grammar) の概念を加えた語用論的テストを開発した。Oller は、言語能力は部分部分に区切る事ができない統合的なものであると考え、個別的要素テストでは無視されていた文脈を重視した。そして、経済的かつ効率的なクローズ・テストが、産出的なテストと同じ技能を測定していると考えた。しかし、後にクローズ・テストは読解であることから、表出的なパフォーマンスの測定はできていないこと、また、結局は文法や語彙という個別的要素を測定しているということがわかった⁷。

(4) 現在 コミュニカティブな方法 コミュニカティブ・テストイング (Communicative testing)

1970年代初頭に、Hymes(1972)がコミュニケーション能力 (communicative competence) の理論を展開し、言語テストに影響を与えた。そして、1970年代末には、実際のコミュニケーションの場でどのようなパフォーマンスができるかを測定すべきであるという考えに至った。これは、コミュニケーションな教育法と呼応している。コミュニケーションな言語テストの特徴を、McNamara (2000 邦訳 p.28)は以下の2つにまとめている。

1. コミュニカティブな言語テストは、パフォーマンス・テストであり、学習者や受験者が受容的、産出的あるいはその両方のある程度の長さのコミュニケーション活動を行っている際にアセスメントを行うことが必要である。

⁴ クローズ・テストとはある程度の長さを持った文章をある一定数(5~7)の単語ごとに空欄を入れたテストである。その空欄を埋めるためには文法・語彙・文脈・語用論的知識が統合的に必要である。(Oller, 1979 邦訳 p.28)

⁵ Oller (1979 邦訳 pp.25-26)は統合的テストと語用論的テストは区別されるものであると説明する。個別的要素テストが言語規則や技能を分解し、部分的、あるいは一部をテストすることに対して、統合的テストは言語規則や技能の2つ以上の側面を同時にテストするものであるという点で対立している。語用論的テストは目標言語の通常文脈的、時間的な制約の中で言語を扱うものでなければならず、また言語外の意味のある文脈がなければならない。語用論的テストは必ず統合的テストとなるが、統合的テストは語用論的自然さの文脈の中でなければ語用論的テストとはならない。

⁶ 通常の言語使用では言葉と身振りを含んだ言語的文脈 (linguistic context) と人、状況、環境や人間関係、個性などの言語外文脈 (extra-linguistic context) がある。(Oller, 1979 邦訳 p.18)

⁷ Oller (1983 ed., *Issues in Language Testing Research*. Rowley, Mass.: Newbury House)自身も自らの主張(言語能力は部分部分に区切る事ができない統合的なものである)を、言語能力はいくつかの独立した分割可能な能力から構成されている可能性が高いと修正した。(武井 2002, p.42)

2. コミュニカティブな言語テストは、受験者が実際に体験しそうな現実社会を想定し、その中での社会的役割に注目し、そのような役割に何が必要かを詳細に示す。

2番目の特徴が統一的・語用論的テストと決定的に違う点である。

Bachman は、Oller が付け加えた語用論的知見のほかに、言語を適切に使用するためには社会言語学的場面を含む文脈が重要であることを、意思伝達言語能力 (communicative language ability) という用語で説明している。意思伝達言語能力を測定するテスト開発では、文法、談話、社会心理学、方略的能力を含んだ広範な言語能力に加え、テスト・アイテムに含まれる言語的情報と機能的意味を相互作用させて処理できる真正性 (authenticity) の高い内容を目指した。もう1つの言語能力に関する見解は、ACTFL (American Council on the Teaching of Foreign Languages) に代表されるもので、言語能力の熟達度⁸は、ACTFL が提示したガイドラインにそって広範囲な機能と話題を持つ一般的言語能力 (general language ability) の到達度⁹で示されるものである。双方とも言語能力とは、ある場面で伝達する際に、言語知識、技能、方略、社会言語学的要素などを駆使して処理することとしているが、ACTFL の場合は、決められた言語能力の一本の尺度計のどのレベルに達しているかに注目している。尺度計に関して、最近のテスト研究者たちに一貫して見られるのは、特定の言語、特定の場面や談話とは関係がない「共通尺度」を作ることを目標としていることである。(Bachman, 1990 邦訳 pp.4-6)

静他 (2002, pp.54-55) は、コミュニケーション・テストとは情報を伝達する様式 (productive mode) と情報を受容し理解する受容的様式 (receptive mode) のどちらか一方、あるいは両方を含むパフォーマンス・テストであり、テストでは、実際のコミュニケーション上で受験者が担うであろう役割とタスクが重要視されるものであると述べている。

菅井 (2006b, p.6) は、Morrow (1979, pp.151-156)¹⁰のコミュニケーション・テストの概念を以下のように邦訳し、これらの概念がその後の課題をも提示していると述べている。

1. パフォーマンス・テストであるが、それを可能ならしめる前提としてコミュニケーション能力がある。細かく寸断された能力を足し合わせても、総合的な能力にはならず、欠けている部分がある。全体を測定するには、実際にありえるコミュニケーションの状況の中に受験者が入って、あるタスクに基づきパフォーマンスをすることで測定できる。
2. タスクは、無限にある可能性があるが、例えば、タスクを遂行するのに必要な受験者の行動には共通する部分も多く、それらの行動における共通した部分をスキルと呼び、細分化されたスキルの総合としてタスクを一般化することができる。スキルとは、コンピタンスを細分化したものではなく、特定のタスクを一般化するためにそれに必要な行動を細分化したものである。
3. 実際のコミュニケーションの状況を注意深くサンプリングする必要がある。しかし、サンプルとして

⁸ 熟達度 (proficiency) は、言語運用の総合的な能力で、ある状況、ある目的のためにどれだけ話す・聞く・読む・書く・理解することができるかという能力の度合いを表す。それまでの学習期間、方法、指導とは必ずしも関係せず、学習者がその時点で持っている、あるいは将来の状況での言語運用能力をさす。(Richards 他 1985 『ロングマン応用言語学用語辞典』 邦訳 p.205、白畑他 1999 『英語教育用語辞典』 p.245、McNamara, 2000 邦訳 p.17)

⁹ 到達度 (achievement) は、ある一定の期間、指導され学習した結果、特定の学習目標にどれだけ到達したのかを表す。(Richards 他 1985 『ロングマン応用言語学用語辞典』 邦訳 p.199、白畑他 1999 『英語教育用語辞典』 p.7、McNamara, 2000 邦訳 p.17)

¹⁰ 菅井は Morrow (1981, pp.18-23) Communicative language testing; revolution or evolution. In the British Council. Issues in language testing. *ELT Documents*, Vol.111 を参考に邦訳している。

抽出された状況やタスクの中で観察されるパフォーマンスが、どの程度サンプル以外の状況やタスクでも観察することができるのか、という問題は今後の課題である。

4. 個別的要素テストや統合的・語用論的テストは有用性がないわけではなく、診断テストや到達度テストに於いては、細かい文法項目等のどの部分を学習者が不得意としているのかを見るのに必要である。しかし、将来特定の社会的文脈の中でどのようなことができるのかを知りたいければコミュニケーション・テストが最適であり、大学入学に必要なテストや、特定の職につくための習熟度¹¹テストに適している。

武井 (2002, p.42) は現在、統合的・社会言語学的時代の考え方である分割可能性仮説¹² (divisibility hypothesis) より、言語能力は独立した様々な構成要素で構成されているが、かなりの部分が重なっている部分的分割可能性仮説¹³ (partial divisibility hypothesis) が有力視されているとしている。しかし、統計手法の限界があるため、言語能力とは何かという確固たる解答はないだろうと結んでいる。

2.2 言語テストにおけるコミュニケーション能力

2.1 で記述したように、近年はコミュニケーション・テストが主流となっている。コミュニケーション能力という用語は、Hymes (1972) が初めて使用した。その後多くの研究者が様々な用語を使用し、コミュニケーション能力を定義している¹⁴。

1970年代、Chomskyの生成文法 (Generative Grammar) に反し、Hymes (1972, pp.290-291) は抽象的な知識体系ではなく、人間関係や状況に影響される現実での使用が重要であると主張した。そして、言語知識のほかに、会話ルールや言語の機能、会話行動領域などの社会言語学的要素が必要だと述べた。Munby (1978, pp.22-27) は、第二言語のコミュニケーション能力として、「社会文化的方向付け (sociocultural orientation) (コミュニケーションをとるコミュニティにおいて、場面に適切でありコミュニケーションのニーズに合致している)」、「言語学知識の社会意味論的基礎 (sociosemantic basis of linguistic knowledge) (言葉の使用は社会構造に由来しており、使用者は適切なアプローチを知っている)」、「操作の談話レベル (discourse level of operation) (発話の表面的な意味、発話の意図、発話参加者の役割、場面などによって変わる)」が必要であると述べている。1980年代には、Canale & Swain (1980) が、コミュニケーション能力を、①文法的能力 (grammatical competence) : 文法・語彙・音韻など ②談話的能力 (discourse competence) : 場面、文脈などを正確に把握でき、的確に伝達できる ③社会言語学的能力 (sociolinguistic competence) : 言語使用の社会的規則や約束事を知っている ④方略的能力 (strategic competence) : コミュニケーションを円滑に行うために言語あるいは言語資質の不十分な部分を補う の4つに下位分類し、言語教育と言語テストの理論的根拠を示した。Savignon (1997 邦訳 pp.18-19)は、コミュニケーション能力は、①動的なものであり対人関係に基づいている ②書き言葉・話し言葉・記号などにもあてはまる ③

¹¹ 習熟度 (mastery) は、言語テストにおいて、その言語テストに該当する言語能力を最低限持っていることを表す。これは目標基準テストにおける「合否分割点 (cut-off score) : あらかじめ決められた最低限能力 (minimum competence) を持っているか否か」に相当する。(Bachman 1999 邦訳 p.241)

¹² 筆者訳

¹³ 筆者訳

¹⁴ 多くの研究者がコミュニケーション能力について記述している。本論ではコミュニケーションの定義や流れなどについて、Bachman (1990 邦訳 pp.93-125)、Richards 他『ロングマン応用言語学用語辞典』(1985 邦訳 pp.63-64)、長澤 (1988, pp.48-69)、白畑他『英語教育用語辞典』(1999, pp.64-65)、McNamara (2000 邦訳 p.30)、Buck (2001, p.83)、静他 (2002, p.54)、菅井 (2006b, pp.6-9) を参考にした。

コンテキストに特定された能力であり、状況と参加者により適切な言語使用域 (register) と様式 (style) を選ぶ ④言語能力 (competence) と言語運用 (performance) は違うが、我々は運用のみ観察し、学び評価する ⑤相対的であり参加者によって適切なレベル (level) を選ぶ と説明している。Bachman (1990 邦訳 pp.96-123)は、意思伝達言語能力が上記に挙げられたコミュニケーション能力の要素と関連しているとし、言語能力 (language competence) とストラテジー能力 (strategic competence)、更に言語使用に含まれるものとして心理生理的機能 (psychophysiological mechanisms) を付け加え、それぞれの下位要素を以下のように分類した (図 2.1)。

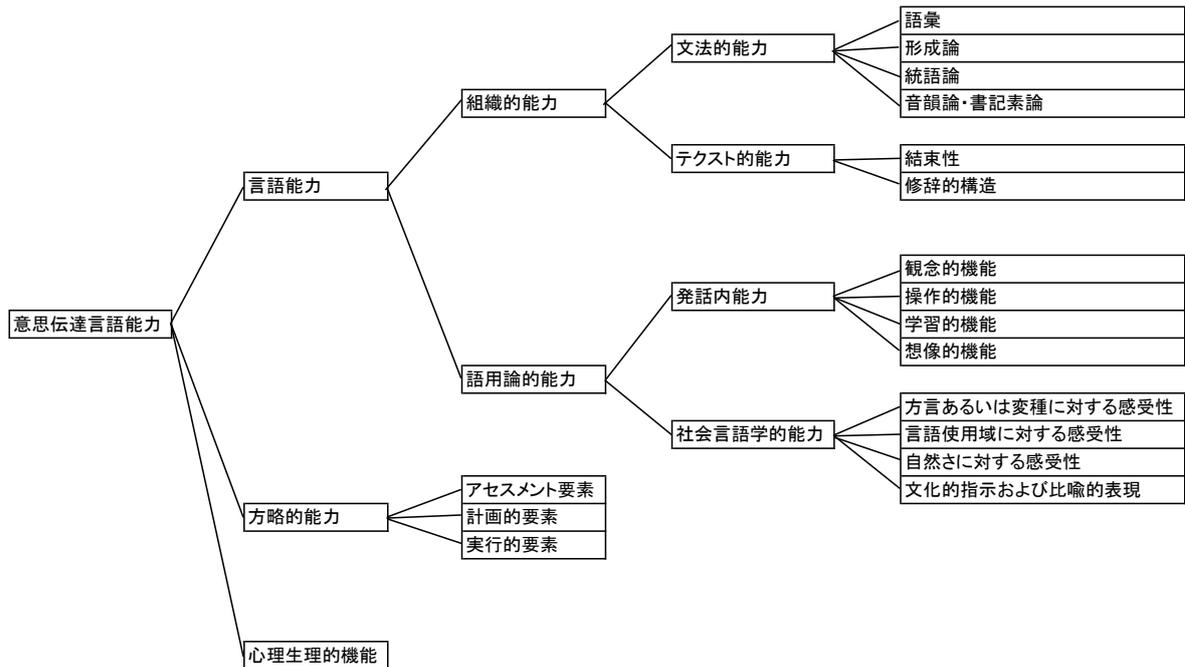


図 2.1 意思伝達言語能力の構成要素 (Bachman, 1990 邦訳 pp. 96-123)

Bachman は、心理生理的機能は、視覚的伝達経路と聴覚的伝達経路、産出モードと受容的モードを区別できると説明している。受容的言語使用時には、視覚と聴覚的技能を使用する。産出的言語使用時には、神経筋肉的技能 (有節音や区切りなど) を用いる。例えば、写真を説明する際、言語能力を正しく使い、視覚的技能で情報収集をし、調音的技能 (イントネーションや正しい発音) を使用して説明する。(Bachman, 1990 邦訳 p.123)

上記の研究から、コミュニケーション能力とは、言語知識のみでなく、コミュニケーションが行われる際の様々な状況や場面に適した言語を選択し使う能力と言える。

2.3 集団規準準拠テストと目標基準準拠テスト

テストは、設計と構造と開発方法の違い、そして、尺度とその解釈方法の違いから、集団規準準拠テスト (norm-referenced test) と目標基準準拠テスト (criterion-referenced test) の2つに分けられる (Bachman, 1990 邦訳 p.87)。

2.3.1 集団規準準拠テスト

集団規準準拠テストは、他の受験者との相対的な優劣により評価されるテストである (Hughes, 2003 邦訳 p.21、Bachman, 1990 邦訳 p.84、石田 1992, p.26、静他 2002, p.118)。集団規準準拠テストの言語テストとしては、日留試、英語テストの TOEFL (Test of English as a Foreign Language)¹⁵、TOEIC (Test of English for International Communication) などがある。

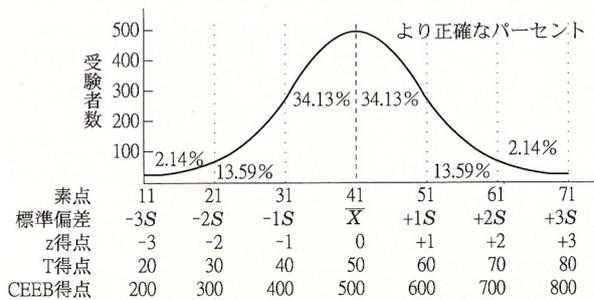


図 2.2 標準得点の正規分布

Brown (1996 邦訳 p. 153)

よく設計された集団規準準拠テストでは、得点分布が正規分布をして、釣り鐘型のカーブ (bell-curve) を描き、平均値からどのぐらい標準偏差 (standard deviation:SD)¹⁶が離れているかという

ことで、集団の中の順位を知ることができる。Brown (1996) は、正規分布の特徴を以下のように説明している。図 2.2 (同掲 p.153) は、100 点満点のテストを受験した結果が正規分布になっていることを示している。 \bar{X} は平均値、S は標準偏差を表す。図 2.2 の場合、最低点が 11、最高点

が 71、平均値が 41、標準偏差が 10 である。平均値 41 から 10 多い 51 までに受験者の 34.13%が存在しており、51 から 61 までには 13.59%、61 から 71 までには 2.14%存在している。平均値より低い場合も同様である。例えば、素点 61 点をとった受験者は、上からほぼ 16 番目 (2.14%+13.59%) あたりの成績と考えてよい。図 2.2 の「標準得点」とは、得点が標準化されていることを示す。得点を標準化する利点は、異なる内容の複数のテストを異なる集団が受けた際の結果を比較することができる点である¹⁷。よく知られたものに、z 得点、T 得点、CEEB 得点がある¹⁸。(Brown, 1996 邦訳 pp.147-158)

McNamara (2000 邦訳 pp.78-79) は、使用が適切であれば複数の異なるレベルを設定でき、個々の受験者間の差の程度を明らかにできる、また、テストの信頼性と妥当性を調べる方法も確立されていると、集団規準準拠テストの利点について触れている。しかし、教育的視野からみると、集団の中での相互比較の順位付けにより、必要以上の競争心をあおり、平均値に近い位置にいる学習者のやる気をそぐと警鐘もならしている。尾崎 (2008, p.54) も、どんなに努力しても、他の人と比較されることによって自分の努力した結果が点数として反映されないこともあり、学習者の学習動機をそぐ可能性があるとして述べている。

Brown & Hudson (2002)は、集団規準準拠テストの問題点として、「①教育とテストが一致しない ②教育指導との関連性が低い ③カリキュラムとの関係が薄い ④正規分布には制限がある ⑤弁別できる

¹⁵ ある特定の基準にしたがって学生の能力を測定するテストは目標基準準拠テストである。TOEFL は安定した出題基準に基づいて作成されているが、その得点や順位が実際の言語使用の程度を予測はできないため、集団規準準拠テストに分類されている (Hughes, 2003 邦訳 p.21、静他 2002, p.119)。集団規準準拠テストの得点や順位が実際の言語使用にどのような意味を持つかは不明であり、TOEFL のように大学の勉学に必要とされる言語能力があるか否かを判断するには換算表を調べる必要がある。その表は過去において類似の得点を得た学生がどうであったかを参考にした経験によるものでなければならない (Hughes, 2003 邦訳 p.23)。

¹⁶ 標準偏差はすべての得点を平均値から引いた差を平均したものであり (Brown 1996 邦訳 p.124)、得点が平均値からどのぐらい離れているのかを測定できる (Richard 他 1985 『ロングマン応用言語学用語辞典』 邦訳 p.352)。

¹⁷ 例えば、ある受験者が国語と数学のテストを受けた場合、どちらのほうが集団の中で高い位置にいるかを知ることができる。あるいは 3 年間に何度か受けた英語テストの得点を比較することで、その受験者の変化を知ることができる。

¹⁸ z 得点、T 得点、CEEB 得点の詳細は参考資料 3 「Z 得点、T 得点、CEEB 得点の算出方法」 (p.194) を参照。

(識別できる) アイテムであることに制限がある」の5つを挙げている。「①教育とテストが一致しない」という理由は、大規模テストであるがゆえに巨大な尺度のもと標準化されたテストを用いているため、各教育機関で行なっている綿密な言語教育に沿っていないことがあるからである。「②教育指導との関連性が低い」とは、ある特定の目的や目標をもった言語教育機関の学習者の言語能力を適切に測定できず、またそのような言語プログラムの欠陥などを診断できないことである。「③カリキュラムとの関係が薄い」は①と②の理由からもわかるように、特定の教育機関のカリキュラムに基づいているわけではないので、カリキュラムの改善やそのカリキュラムで学習した学習者の評価には適していない。「④正規分布には制限がある」というのは以下の理由である。集団規準準拠テストは、長年統計的にテスト結果が歪みのない正規分布を描くように設計されてきた。しかし、受験者集団がかなり高いレベルを持っており、テストが簡単過ぎた場合は、その分布には歪みが生じる。この場合、他のテストの結果との比較などには不適切となる。「⑤弁別できる(識別できる) アイテムであることに制限がある」とは、集団規準準拠テストのテスト・アイテムは受験者の半数ができないことをよしとする前提で作成されていることにある。これは、高い能力を持った受験者と低い能力の受験者を明確に識別するためである。そのため、テスト設計者はカリキュラムや何を学んできたのかより、いかに識別することができるかを重視してテストを設計してしまう原因となる。(Brown & Hudson, 2002, pp.6-8)

以上のように、集団規準準拠テストは、各国で違う内容やレベルで学習してきた受験者たちを1つのテスト結果の中で位置づけることに適している。しかしその結果、受験者は自分の言語能力に注目することより要求された得点をとることに集中してしまいがちである。特に大学入学時の際に使用されるTOEFLや日留試など、受験者にとってテスト結果がその後の人生に深く関わってくる場合は、1点でも高い得点を得たいと考えるだろう。内容理解が不十分でも正解できるスキルがあるとすれば、コミュニケーション能力を伸ばそうという曖昧な目標より受験者にとっては魅力的である。言語学習がテスト対策になってしまうことを避けるために、目標言語を使用して何ができるか(can-do)を明確に細かくテスト使用者に対して説明することが得点以外への注目を促すという点で有効である。Can-do リストは、言語教育に関わる機関でカリキュラムやシラバスをたてる際に有用となり、テスト結果の数値の解釈にも役立つ。

2.3.2 目標基準準拠テスト

目標基準準拠テストは、明示された知識や技能のレベルに受験者が達しているかいないか、あるいは複数のレベルのうちどのレベルに達しているかを測るテストであり、あくまでも設定された目標に個人が達したか否かをみる絶対的な評価である(Hughes, 2003 邦訳 pp.22, Bachman, 1990 邦訳 p.86, 石田 1992, p.26, 静他 2002, p.63)。このテストは、テスト内容が適切に内容領域の代表的なサンプルであること、能力水準の度合いを明確に示す事ができるかどうかということが重要である(Bachman, 1990 邦訳 p.86)。Bachman はそのために、テストの内容領域を細目化(specified)する必要があると述べている(同頁)。目標基準準拠テストの言語テストは、日能試や実用英語技能検定、ACTFL (American Council on the Teaching of Foreign Language)¹⁹、IELTS (International English Language Testing Systems)などである。診断テスト(diagnostic test)や到達度テスト(achievement test)も、目標基準準拠テストである。集団規準準拠テストの結果が正規分布になるのに対し、目標基準準拠テストの結果は正規分布する必要はない。受験者集団がテスト項目をよく学んでいれば高い得点に集中し、そうでなければ低い得点に集中する。評価

¹⁹ Rost (1990, p.187)はACTFLのガイドラインを集団規準準拠テストの基準として挙げている。その基準は、母語話者のための基準であり、受験者はその基準と相対的な比較をしているため、集団規準の例としたのである。しかし、本研究では、特定の基準があり、その基準に合格するという意味でACTFLは目標基準準拠テストの例とした。

の出し方には合格・不合格、あるいは到達度により A・B・C・D などで段階づけたものもある。

Brown (1996) は、目標基準準拠テストの妥当性と困難度について以下のように述べている。目標基準準拠テストは、診断テストや到達度テストとして使用されるため、ある言語プログラムの内容や目的に沿ったものが望ましい²⁰。このテストは、テスト・アイテムの品質分析をすることで妥当性を見る。集団規準準拠テストは、一般的な言語能力を測定し、受験者間の弁別度（識別力）が高いアイテムが必要なため、難易度（困難度）と弁別度（識別力）分析の数値にある程度依存してよい。しかし、目標基準準拠テストは、テスト・アイテムが言語プログラムの目標と、そのための知識や能力が測定できるようになっていなければならない。テスト・アイテムの内容分析においては、そのテスト・アイテムの細目一覧が明確にまた具体的に記述されており、その細目一覧に従いテスト・アイテムの整合性や適切性の判断ができる専門家が必要である。テスト・アイテムの難易度分析（困難度分析）には2つの方法があり、1つは介入方式（intervention strategy）で、事後テストから事前テストの各テスト・アイテムの正答率を差し引いたものである。もう1つは、差異集団方式（different groups strategy）である。これは、既にテストで測定される能力を身につけている集団と身につけていない集団、この2つの集団の各テスト・アイテムの項目難易度指数（困難度指数）を比較する。どちらの方法でも、2つのテストの結果の差異は差異指数（difference index）とよばれ、テスト・アイテムの能力を持っている受験者とそうでない受験者を区別する度合いを-1から+1の間で示している²¹。差異指数を使用する場合、2回テストを実施しなければならない。この問題点を解決したのがB-指数（B-index）である。B-指数はテスト・アイテムの統計値で、テストの合格者と不合格者の項目難易度（困難度）を比較する。合格か不合格かを定めるためには、分割点が必要である。分割点以上の受験者の各テスト・アイテムの正答率から、分割点以下の受験者の正答率を差し引くとB-指数が求められる。差異指数もB-指数も-1から+1の間の数値をとるが、-1から0の間の指数を持つテスト・アイテムは、受験者の能力測定には不適切であり、差異指数の場合は、言語指導に何らかの問題があったのではないかとすることも考えられる。差異指数とB-指数の解釈は似ているが、言語プログラムの効果や受験者の個々の問題点などをみるには差異指数がよく、大規模試験で一般的な言語能力の到達度をみる場合はB-指数が適していると言えるだろう。（Brown, 1996 邦訳 pp.87-96）

目標基準準拠テストは、能力水準、シラバス、テスト内容の細目化が事前に明示されているため、学習者は学習動機が高まる（Hughes, 2003 邦訳 p.23、尾崎 2008, p.53）。しかし、静他（2002, p.64）は、言語使用に関しての目標基準が恣意的であると問題視している。特殊なあるいは狭い範囲を目標とした言語プログラムは、恣意的である可能性が高い。日能試や実用英語技能検定などの大規模テストの場合、かなり広い範囲を網羅した基準を設定していると思われるが、これらの基準が一般的かどうか、また複数のレベル設定が適切かどうかの判断は難しいと言わざるをえない。Oller（1979 邦訳 p.52）は、目標基準準拠テストは教師に目標やプログラムの適切さ、教授方法などを再考させる教育的効果があるとしているが、Brown（1996 邦訳 p.10）は、教師がテストで高い点数をとらせるためにばかり時間を使い、ほかの大切なことに時間を使わなくなると指摘している²²。

目標基準準拠テストの特徴として、Brown & Hudson（2002, p.9）は、以下の5つを挙げている「①教育と

²⁰ ただし、特定のプログラムの内容や目的にそってなくても、明確な水準が提示できる場合は、この限りではない。

²¹ -1は事前テスト・未習者集団で全員正答し、事後テスト・既習者集団で全員不正答の場合。+1はこの反対である。+1となるのが理想的である。

²² Brown はテストのための教育を否定しつつも、「良いテストのための教育であれば適切なものとなる（Brown, 1999, p.10）」と述べており、テストのための教育・学習が必ずしもすべて良くないと言っているのではない。

テストが一致している ②言語指導の精度に注目している ③カリキュラムが妥当である ④正規分布は不必要である ⑤アイテムの弁別度（識別力）は不必要である」。この5つは、前述した（p.15） 集団規準準拠テストの問題点と相反している。しかし、これは目標基準準拠テストが集団規準準拠テストに勝っているということではない。あくまでも、目標基準準拠テストは学習した内容をどの程度習得しているのかを測定するテストであるため、言語教育の現場から見たときに使いやすいテストであるという考え方である。

Brown (1996 邦訳 p.6、Brown & Hudson, 2002, pp.9-11)は、この2つのテストの違いを比較している（表 2.1）。テスト使用者および利用者は、それぞれのテストの特徴と、どのような学習者のどのような能力を測定しているのかを理解し、テストの目的に鑑みて適切なテストを選択することが望ましい。

表 2.1 集団規準準拠テストと目標基準準拠テストの違い (Brown, 1996 邦訳 p. 6)

特徴	集団規準準拠テスト	目標基準準拠テスト
1. 解釈の仕方	相対的（ひとりの受験者の成績がパーセンタイル値 ²³ によってほかの受験者の成績と比較される）	絶対的（ひとりの受験者の成績は習得した言語材料、すなわちパーセンタイルで対比される）
2. 測定の種類	一般的な言語能力や言語熟達度を測定する	目標に定められた特定の言語項目を測定する
3. テストの目的	受験者を一般的言語能力や言語熟達度に基づいて直線上に並べる	個々の受験者が習得した言語材料の量を測定する
4. 得点の分布	平均値を中心にした正規分布をする	普通は正規分布しない（例えば、学習事項を全部知っている受験者は全員100%の得点をとるはずである）
5. テストの構造	様々な内容からなる少数の比較的長い下位テスト問題	同種類の内容からなる明確な目標のある下位テスト群
6. 問題の予想	受験者は問題内容をほとんど予想できない	受験者はテスト問題の内容を正確に予想できる

2.4. 日本語能力試験と日本留学試験

2章 2.3 では、テストの目的や尺度、解釈の違いでテストの種類が決まることを述べた。本節では、日能試と日留試を比較し、その違いを詳述する。テスト全体の記述後、本研究は聴解テスト研究であるため、聴解テストに焦点を当てて論ずる。

2.4.1 日本語能力試験

日能試は、1984年から2009年まで実施された旧日能試と、2010年から始まった新日能試があるので、その2つのテストについて述べる。

旧日能試は、日本語を外国語として学習した人達の日本語能力を測定する試験として、日本で初めて開発・実施された。旧日能試は、4段階（1級～4級）のレベルで測定をしていたが、新日能試より5

²³ パーセンタイル (percentile) は統計学的に、受験者の順位を示す。たとえば、70パーセンタイルの得点をとった受験者は受験者100人中70人の受験者より成績がよかったという意味である。(Richards 他 1985 『ロングマン応用言語学用語辞典』 邦訳 p.274、Brown, 1996 邦訳 p.7)

段階 (N1～N5)²⁴となった。また、旧日能試では、「文字・語彙、読解・文法、聴解」という分類であったが、新日能試では「言語知識 (文字・語彙・文法)、読解、聴解」に変更された。大きな違いは文法と読解が明確に分かれたという点である。ただし、下位レベル (N4、N5) においては、科目として言語知識と読解は分かれているが、採点上は分けていない。旧日能試は『日本語能力試験出題基準』(1994年)に、各レベルで提出される「文字、語彙、文法」リストを提示していた。読解においては、テスト・アイテムのテキストの語彙逸脱率や漢字含有率なども示され、文字・語彙の知識が読解においても必要であることを明確にしている。出題基準には、読解、聴解に関して、どのような能力をどのような形式で測定するのかの指針が書かれている。このことからわかるように、旧日能試は言語知識偏重であったことに対し、新日能試は公開された基準はなく、文字・語彙の言語知識と読解・聴解のスキルを明確に分けたテストであると言える。新日能試は「課題遂行のための言語コミュニケーション能力」を測定しているが、そのために必要な語彙、文法が何であるかは明示していない。課題遂行はどのような方法であれ、課題達成ができればよく、そのために必要な言語知識も多種多様であると考え、「何が」必要であるかの規定はしにくい。もし、新日能試で設定したレベルに必要なであると規定した語彙、表現、文法などが公表された場合は、新日能試の受験者および利用者にとって、どのレベルがどの程度の日本語能力であるかの理解がしやすい。また、具体的なリストや詳細な能力記述は、日本語教育のシラバスを考える際の見本となり、日能試のスタンダードとして日本語教育関係者、利用者に広く認知され、より統一されたレベル観を育成する。このような利点を考えると、特に言語知識に関しては具体的な基準を示すことが望まれる。また、読解と聴解においても、旧日能試は詳しい説明を公開しているが、新日能試は各大問²⁵のねらいのみである。今後、新日能試の具体的な目標の公開の有無に注目したい。

2.4.2 日本語能力試験 聴解

・旧日本語能力試験 聴解

聴解に関しては、旧日能試も新日能試もどの程度の言語知識が必要かという提示はない。旧日能試の聴解テストでは、構文や語彙の知識も必要だが、談話テキストが理解できるためには、談話が一本の流れではなく脇道にそれたりした際、それについていく力も必要である。そして、当該レベルの構文や語彙の知識だけでは十分ではなく、必ず未知のものが含まれているとしている²⁶ (『日本語能力試験出題基準』1994, p.191)。このように、構文や語彙、情報を正確に聞き取る能力も含め、それ以上のものを「聴解行動能力」とよび、以下のように説明している (同掲 pp.193-194)。

- ① 不必要な細部にこだわらず、重要な部分に焦点を当てて聞くなどの聞きのためのストラテジーを駆使する能力
- ② 情報を整理し、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、さらには仮説の設定、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

²⁴ 旧日能試の1級から4級、新日能試のN1～N5についての詳しい説明は参考資料4「旧日能試と新日能試のレベル」(p.196)を参照。

²⁵ 大問とは測ろうとしている能力の名称である。例えば、N1の文字・語彙の大問は「漢字読み、文脈規定、言い換え類義、用法」であり、漢字の読み方を問う、文脈で規定された語彙が何であるか問う、意味的に近い語や表現を問う、文の中でどのように使われるかを問うというねらいが記述されている。(国際交流基金・日本国際教育支援協会 2009『新しい「日本語能力試験」ガイドブック』)

²⁶ 読解テスト・アイテムにも未知なものは含まれているが、表意文字の漢字、また何度も繰り返し見ることができるという時間的な余裕と視覚情報があることから、聴解に比べより言語知識に頼る割合が多いと考えられる。

聴解テスト・アイテムは、日常生活で生じる課題を解決する過程を前提としており、課題解決の知的行動に基づきなされるのが、コミュニケーション行動である。聴解テストで測定する聴解行動能力は、コミュニケーション行動の一部をなしている（図 2.3）（同掲 pp.193-195）。

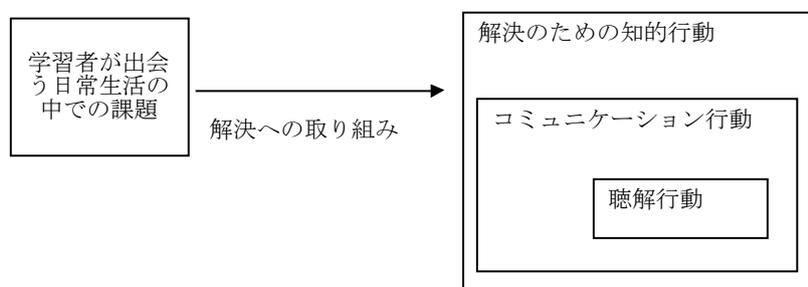


図 2.3 旧日能試 聴解テストで測定する能力（『日本語能力試験出題基準』 p. 195）

学習者が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、テスト・アイテムは作成される。測定する聴解能力は、以下の4つのテスト・アイテム・タイプに分けられる（同掲 pp.185-201）。

テスト・アイテムの「指示（プリタスク）」を聞き、「判断」をすることを前提とし、「本文」からは

- ① 判断に必要な情報と不必要な情報があるので、何が大切かを聞き分ける。
- ② 一般的な規則や傾向を具体的なデータに基づいて検証できるかどうか、検証されないならばどのように修正されるべきかの判断が求められる。
- ③ 情報を比較して、共通点を見つけ、解決策を捜し出す。
- ④ 断片的な情報からその情報の全体像をつくりあげ、さらに違う断片的な情報から検証し、修正し、全体像を強化していく。

そして、テスト課題としては以下の7つを提示している（表 2.2）（同掲 p.201）。

表 2.2 旧日能試の聴解課題（『日本語能力試験出題基準』 p. 201）

課題名		1級	2級	3級	4級
情報完成	指示による情報完成				
	情報完成による論理的判断				
規則・法則・傾向の適用／検証	規則・法則・傾向の具体例への適用				
	規則・法則・傾向の検証／修正				
	具体例に基づく規則・法則・傾向の判断				
情報総合	情報の比較（相違点）				
	情報の比較（共通点）				
	情報の総合				
意図の汲み取り	はい／いいえの意図の判断				
	意向の判断				
照合	4つの事物・人物の照合				
	1つの事物・人物の照合				
順序の再構成	時間的順序の再構成				
	手順・手続きの再構成				
事物・人物の同定					

*各級の色濃度は出題される傾向が強いほうが濃い

ここで、旧日能試の聴解テスト1級のテスト・アイテム構成を見てみる。1級を分析する理由は、日留試が大学入学のために利用されることから、旧日能試の1級に相当するレベルと考えられ、内容やレベルを比較しやすいためである²⁷。

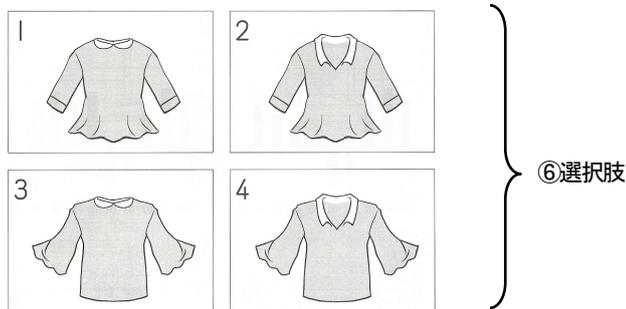
²⁷ 日本の大学に留学する場合、旧試験では1級あるいは2級以上を求められていた。ここでは1級のテスト・アイテムを例として挙げる。1級～4級の構成は同じである。

例1 旧日能試 2009年2回目 1級 聴解問題I1番

お母さんと娘が電話で話しています。・・・①状況説明
 お母さんはどのシャツを買えばいいですか。・・・②プリタスク } ③設問

F1:もしもし、エリちゃん、買ってきてって言っていたシャツ、どんなんだっけ。
 F2:襟つきのブルーのやつだよ。襟が小さくて、すそがちよっと広がってて、あっ、襟んところが
 広く開いてるのじゃないから、間違えないでね。
 F1:ああ、わかったわ。 } ④本文

お母さんはどのシャツを買えばいいですか。・・・⑤ポスト・クエスチョン



例1は「⑥選択肢」4つが視覚情報として与えられている。聞くのは「①状況説明」「②プリタスク」が「③設問」として最初にあり、その後「④本文」を聞き、「⑤ポスト・クエスチョン」を聞く²⁸。

例2 旧日能試 2009年 1級 聴解問題II1番

交通事故についてのニュースです。・・・①状況説明
 どんな事故でしたか。・・・②プリタスク } ③設問

F:今日の午前11時頃、南区のスーパーの駐車場で車が店の壁に衝突するという事故がありました。
 事故当時、駐車場や店内には客が大勢いましたが、幸い負傷者はいませんでした。車を運転して
 いたのは近くに住む30歳の男性で、駐車場に車を止めようとした際、ブレーキとアクセルを踏み
 間違えてしまったと話しています。 } ④本文

どんな事故でしたか。・・・⑤ポスト・クエスチョン

1. ブレーキの故障で事故が起こり、けが人が出た。
2. ブレーキの故障で事故が起こったが、けが人はなかった。
3. 車の操作ミスで事故が起こり、けが人が出た。
4. 車の操作ミスで事故が起こったが、けが人はなかった。

} ⑥選択肢

²⁸ テスト・アイテムの各部分の詳しい説明については5章図5.1「日留聴解テスト・アイテムの構成」p.100を参照。

例2は選択肢が視覚情報として与えられていないこと以外は、例1と同様の形式である。旧日能試の聴解テスト・アイテムの出題形式は、この2種類である。

旧日能試の報告書（日本語能力試験実施委員会 日本語能力試験企画小委員会 2004, pp.24-25）では、各レベルの信頼性および類間の相関を出している。例えば、2002年度（平成14年度）実施の1級は、総点との相関が「聴解」0.744、「文字・語彙」0.761、「読解・文法」0.942である²⁹。3類間の相関は、全級とも「文字・語彙」と「読解・文法」が1番高く、次に「読解・文法」と「聴解」、最後に「文字・語彙」と「聴解」の順である。特に1、2級では、「聴解」は運用力に関わる能力を測定しており、「文字・語彙」のような知識とは関連があまりないと推察している³⁰。青山他（2003, p.27）は、旧日能試の因子分析を行っており、級が低い場合は各類は一因子性を示し、級が上がると因子数が増したと報告している。これは、級があがるほどそれぞれの技能に必要な能力が多様になっていることを示している。これらの結果からわかることは、高いレベルの日能試の聴解は運用力を測定しており、聴解スキルは言語知識（文字、語彙、文法など）のように部分的に分解できる能力の測定ではないと言ってよいだろう。

・新日本語能力試験 聴解

次に、新日能試について、旧日能試と比較しながらまとめる。新日能試は、文字・語彙・文法をどの程度知っているかに加え、その言語知識を利用してコミュニケーション上の課題を遂行できるかどうかを重視している（図2.4³¹）。



図2.4 課題遂行のための言語コミュニケーション能力測定（新日能試）

旧日能試と新日能試の違い³²は、「レベル」「認定の目安」「試験科目と試験時間・配点」の3点である。「レベル」設定は、4段階から5段階になった。旧日能試1級は、新日能試N1に相当する。「認定の目

²⁹ ただし、「読解・文法」が200点で他の2類が100点であるため、「読解・文法」の配点が高くなっている。

³⁰ この傾向は1984年の開始時より続いており、ここで記述した2002年度以降（～2009）の報告書でも同じ分析となっている。

³¹ 2015年3月 日本語能力試験ホームページより <http://www.jlpt.jp/about/points.html>

³² 参考資料4「旧日能試と新日能試のレベル」（p.196）を参照。

安」は旧日能試の1級では「高度の文法・漢字（2000字程度）・語彙（10000語程度）を習得し、社会生活をする上で必要な、総合的な日本語能力（日本語を900時間程度学習したレベル）」であり、出題基準が公開され、具体的に習得すべき文法・漢字・語彙が示されている。一方、N1は「幅広い場面で使われる日本語を理解することができる」とし、「文字・語彙・文法」などの言語知識を基礎とした「読む」と「聞く」のスキルを以下のように説明している。

読む：幅広い話題について書かれた新聞の論説、評論など、論理的にやや複雑な文章や抽象度の高い文章などを読んで、文章の構成や内容を理解することができる。さまざまな話題の内容に深みのある読み物を読んで、話しの流れや詳細な表現意図を理解することができる。

聞く：幅広い場面において自然なスピードの、まとまりのある会話やニュース、講義を聞いて、話しの流れや内容、登場人物の関係や内容の論理構成などを詳細に理解したり、要旨を把握したりすることができる。

新日能試には、旧日能試のような出題基準は公開されておらず、文字・語彙や文法は、「読む・聞く」スキルを使用して課題遂行するために必要な言語知識という位置づけである。「試験時間と配点」は、旧日能試1級が「文字・語彙45分100点」「聴解45分100点」「読解・文法90分200点」合計180分400点に対し、新日能試N1は「言語知識（文字・語彙・文法60点）と読解60分で110分」「聴解60分60点」合計170分180点である。得点は旧日能試は素点だが、新日能試は等化による尺度得点である³³。旧日能試との比較からわかるように、聴解の時間配分および配点が全体の1/4から1/3に増え、聴解がコミュニケーションをとる上でより重視されてきたことがうかがえる。

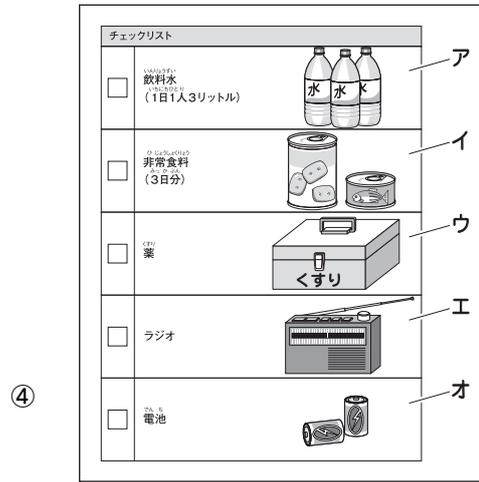
N1聴解には5つの大問があり、それぞれにねらいが示されている。それぞれのテスト・アイテム例（国際交流基金・日本国際教育支援協会2009『新しい「日本語能力試験」問題例集』より）を以下に挙げる。テスト・アイテム例にある数字は、①：状況説明、②：プリタスク、③：設問、④：本文、⑤：ポスト・クエスチョン、⑥：選択肢である。

³³ 素点はテストで得られた生の点数のことである。素点はそのテストのみに対して有効であり、違うテストとは比較できない。例えば、学習者Aが日本語テストAを受験し素点100点を取り、その半年後に日本語テストBで素点50点をとった場合、学習者Aの日本語能力が低くなったかどうかは判断できない。テストBの困難度が非常に高かった場合には、このような結果になる可能性があるからである。等化はあるテストの2つのフォームの結果を比較可能にする。例えば、同じ細目に基づき、同じ形式で作成されたテストの場合、困難度の小さな違いによる得点の違いはなくなる。（『テスト作成ハンドブック』2008,p.58、p.487）新日能試は素点をより適切な解釈をするために尺度得点に換算したものであり、テストの困難度に左右されないため、より正確な結果を出すことができる。

例3 課題理解：まとまりのあるテキストを聞いて内容が理解できるかどうかを問う（具体的な課題解決に必要な情報を聞き取り、次に何をするのが適当か理解できるかを問う）

女の人と男の人が、地震に備えて表を見ながら話しています。 ①
 男の人は、何を買わなければなりませんか。 ② } ③

F：何が足りないかな…と、水、1日1人3リットル。
 M：今は6リットルあるよ。
 F：あ、そう。でもうちは3人家族だから…。
 M：そうだね、じゃ追加だね。
 F：そうね、それから…。非常用食料、これは3日分。
 M：3日分、はい。あ、でもこれはもう期限が切れてるよ。
 F：あら、そう。じゃ、新しいのがあるね。
 M：うん。
 F：あと…薬は…それからラジオはあったわよね？
 M：両方ともあるよ。でも、電池は、新しいのを用意しよう。
 F：そうね、それがいいわ。それじゃ、買い物、お願いね。



⑥視覚情報

男の人は、何を買わなければなりませんか。 ⑤

- 1 ア イ ウ
 - 2 ア イ オ
 - 3 ア エ オ
 - 4 ウ エ オ
- } ⑥視覚情報

「⑥選択肢（図あるいは文字）」を見ながら、「③設問」である「①状況説明」「②プリタスク」を聞き、「④本文」、再度「⑤ポスト・クエスチョン」を聞く。

例4 ポイント理解：まとまりのあるテキストを聞いて内容が理解できるかどうかを問う（事前に示されている聞くべきことをふまえ、ポイントを絞って聞くことができるかを問う）

大学で男の人と女の人が話しています。 ①
この男の人はどうして先生が怒ったと言っていますか。 ② } ③

M：あー、先生を怒らせちゃったみたいなんだよねー。困ったなあ。

F：え、どうしたの？

M：うーん、いやそれがね、先生に頼まれた資料、昨日までに渡さなきゃいけなかったんだけど、いろいろあって渡せなくて…。

F：えー、それで怒られちゃったの？

M：うん、いや、それで怒られたっていうより、おととい、授業の後、飲み会があつてね。で、ついそれを持ってっちゃったんだけど、飲みすぎて、寝ちゃって、忘れてきちゃったんだよ。

F：え？ じゃ、なくしちゃったわけ？

M：いや、出たはきたんだけどね、うん。先生が、なんでそんな大事な資料を飲み会なんかを持って行くんだって。

F：ま、そりゃそうよね。

この男の人はどうして先生が怒ったと言っていますか。 ⑤

- 1 昨日までに資料を渡さなかったから
 - 2 飲み会で飲み過ぎて寝てしまったから
 - 3 飲み会に資料を持って行ったから
 - 4 資料をなくしてしまったから
- } ⑥視覚情報

例4は例3と同様の構成だが、「③設問」のあと、「⑥選択肢」を読む時間が20秒程度与えられる。

例5 概要理解：まとまりのあるテキストを聞いて内容が理解できるかどうかを問う（テキスト全体から話者の意図や主張などが理解できるかを問う）

大学の先生が話しています。 ①

M：今日は最初の授業なので、授業内容について簡単に説明します。えー、犬の祖先は、今の犬とは、外見だけではなく、習性もずいぶん違っていました。ちょっと例をあげてみますと、進化の結果、犬は、よくほえるようになりましたが、犬の祖先はめったにほえませんでした。これはですね、人間の都合によって、ほえる犬が選択されたためです。それから、進化の過程で形を変えた動物もいます。ある鳥は、細長い花の蜜をすうために、くちばしが異常に長くなりました。あと、住む環境にあわせて、形を変化させたものもいますね。えー、この授業では、こういう現象をみていきたいと思います。

} ④

この授業でとりあげる内容はどのようなことですか。 ⑤

- 1 動物の種類
- 2 動物のすみ環境
- 3 動物の進化
- 4 動物と人間の関係

} ⑥

例5では、「①状況説明」のあと、すぐに「④本文」となり「⑤ポスト・クエスチョン」「⑥選択肢」を聞く。概要理解と銘されているので、本文の概要を聞き取ることが求められていることは事前に分かっているため、プリタスクがなくとも、全体理解を問われると予測できる。

例6 即時応答：質問などの短い会話を聞いて、適切な応答が選択できるかを問う。

M：あーあ、今日はお客さんからの苦情が多くて、仕事にならなかったよ。・・・⑦語りかけ

- F：1. いい仕事できてよかったね。
2. 仕事なくて、大変だったね。
3. お疲れ様、ゆっくり休んで。

} ⑧応答選択肢

例6は、例1から例5までの構成と大きく違い、「⑦語りかけ」を聞いた時点で状況判断をし「⑧応答選択肢」を聞いて選ぶ。今までの聴解テスト・アイテムは本文の全体理解や部分理解であったが、「即時応答」は応答力を問うことから、「話す」につながる聴解テスト・アイテムと考えられる。

例7 統合理解：長めのテキストを聞いて、複数の情報を比較・統合しながら内容が理解できるかを問う

お店の人がジュースの説明をしています。①

M1：えー、こちらをご覧ください。当店ではおいしさだけでなく栄養のバランスを考えた健康ジュースをご用意いたしました。黄色、紫、緑、赤の4種類。それぞれ効果が異なりますので、皆さまの体調や目的に合わせてお選びいただけます。まず男性サラリーマンの方に人気なのがこの黄色でして、こちらは疲労回復効果があります。えー、次に女性の方にお勧めなのが、この紫で、美容に大変いい要素が豊富に含まれております。緑には、パソコンなどによる目の疲れを取る働きがございます。またちょっと高いんですが、若い男性の方々からは、黄色にさらに美容効果を加えたこの赤が、大変好評でございます。

} ④

F：へえ、よさそう。やっぱりきれいになるっていうのは魅力的ね。

M2：でも、仕事で一日中パソコン使ってるんだよね？

F：そうなのよね。目の疲れってのもつらいのよね。そっちのほうがいいかな。

M2：じゃあ、飲んでみたら？僕は最近体がだるいから…。

F：赤ならお肌もきれいになりそうね。ちょっと高めみたいだけど。

M2：いやあ、僕は美容効果の方はいいよ。

} ④

質問1 この女の人にはどのジュースが最も効果的ですか。⑤

- 1 黄色 2 むらさき 3 緑 4 赤 ⑥視覚情報

質問2 この男の人にはどのジュースを飲もうと考えていますか。⑤

- 1 黄色 2 むらさき 3 緑 4 赤 ⑥視覚情報

例7は「①状況説明」から「④本文モノログ ④'本文ダイアログ」が続く。その後「⑤ポスト・クエスチョン」が2種類ある。「⑥選択肢」は文字により視覚情報として与えられている。ただし、この大問は3項目あり、うち2項目は質問が1つで、選択肢は聞いて答える形式である。

新日能試の聴解テスト形式の特徴をまとめる。例3と例4の課題理解とポイント理解は選択肢が視覚情報として与えられており、「設問(状況説明・プリタスク)、本文、ポスト・クエスチョン」を聞くと同じタイプである。例5の概要理解は、視覚情報はなく「状況説明、本文、ポスト・クエスチョン、選択肢」で構成されている。例7の統合理解は、質問が2つある場合は選択肢が視覚情報として与えられているが、そうでない場合は与えられない。選択肢が視覚情報か否かではなく、統合理解は「状況説明、本文、ポスト・クエスチョン、選択肢」の構成であるため、テスト・アイテムの形式としては例5の概要理解と同じタイプと考える。例6の即時応答は、プリタスクがなく選択肢が視覚情報ではないという点では概要理解、統合理解と同様であるが、1文の語りかけに答え、選択肢も4つではなく3つであることから、1つの独立した形式とする³⁴。旧日能試はすべてに設問(状況説明・プリタスク)が

³⁴ 新日能試のN3～N5には発話表現という大問がある。1枚の絵とその状況説明を聞き、指示された人物の発話文を選択肢(3つ)から選ぶものである。

あるが、新日能試では設問（状況説明・プリタスク）があるタイプは選択肢が視覚情報として与えられており、状況説明はあるがプリタスクがないタイプは選択肢が視覚情報として与えられていない。プリタスクがないアイテムは、本文を聞きながら大切な情報を抽出する能力測定を意図していると考えられる。新日能試で特筆すべき点は、選択肢が書いてあり、受験者は選択肢を読むという変更である。「読む」といっても、読解のような「読み」能力は必要なく、短い文章レベルである。この変更は、4つの選択肢を記憶保持する必要がないことを意味する。新日能試以前の聴解テストでは、聴解以外のスキルを測定しないという強いピリーフが存在していた。旧日能試・日留試を経て、初めて大規模日本語テストの聴解で文字が提示された。聴解テストは聴力テストでもなく、記憶力テストでもないということである。設問（状況説明・プリタスク）および選択肢も含め、全てのテキストを聞かせることが聴解能力の測定であったとした旧日能試、日留試といった大規模日本語テストが行われてきた日本語テストにおいて、これは画期的なことである。そして、選択肢を見せる形式は、今後設問を見せるということに進む可能性も含め、聴解テストの形式を自由にしたと言える。

2.4.3 日本留学試験³⁵

本節では、日留試について説明する。日留試には日能試のようにスキルで分けた記述がないため、全体像についてまとめる。日留試の日本語シラバスでは、「日本での留学生活を送る上で、日本語によるコミュニケーション能力があるかどうか、また、自国での初等・中等教育修了までに習得した知識を前提としながら、日本の大学で学習・研究活動を行うための日本語能力があるかどうかを測定する言語テストであり、かつ、標準テストである」とうたっている（「日本留学のための新たな試験」調査研究協力者会議 2000, p.12）。これは、「言語知識だけではなく、社会的、文化的慣習に即した適切さや、コミュニケーションを効果的に進める能力などを含めて、外国語でのコミュニケーション能力を測ろうとするテスト」であるコミュニカティブ・テストの定義と（白畑他 1999 『英語教育用語辞典』 p.66）一致する。同辞典ではコミュニカティブ・テストの測定方法として、インタビュー・テストやロールプレイ・テストに限定はしているものの、「発音、文法、語彙、流暢さ、理解度などの細かな項目を設定して評価を行い、言語的な正確さと意志伝達の成功度の両方を評価する方法が一般的によく用いられている」と述べている（同掲 p.67）。日能試では、語彙、文法などの言語コードを測定する項目があるが、日留試には言語コードを測定する項目は見られない。日留試では、知識の有無、量は測定しないという前提にたっているためである。では、日留試は、どのような聴解能力をどのように測定するのだろうか。その答えを得るためには、テスト・アイテムの構成内容と、テスト作成時の構成概念の推測をするのが有効である（Bachman & Palmer, 1996 邦訳 p.89、McNamara 2000 邦訳 p.24）。

³⁵ この節では2002年度から2009年度まで、改定前の日留試のシラバスを扱う。

日留試の日本語シラバス（日本国際教育協会 2003, pp.244-249）を参考に、日留試で測定する対象能力（図 1.1）（網掛け部分がテストで測定する部分）³⁶、言語下位技能、試験の類型を以下に挙げる。

・言語下位技能

1. 情報の全体の流れをとらえる
2. 情報の全体を、ある判断や評価をしながらとらえる
3. 特定の情報を抽出してとらえる
4. 推測しながら情報をとらえる
5. 予測しながら情報をとらえる
6. その他

・課題の類型

1. 指示を実行する
2. 事物を特定する
3. 事物を描写する
4. 事物を比較・配列する
5. 物事の推移・展開を予測する
6. 物事の背景や意図を把握する
7. 物事を構造化し、法則性を発見する
8. その他

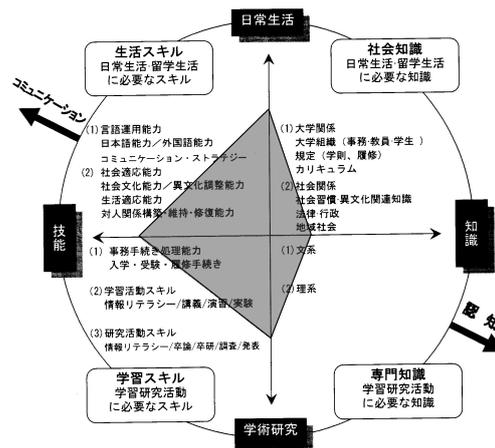


図 1.1 日留試の測定対象能力（日本国際教育協会 2003, p245）

日留試の「課題の類型」と旧日能試の聴解テスト課題（『日本語能力試験出題基準』p.201）を比較すると、課題はほぼ同じであることがわかる（表 2.3）。旧日能試の聴解は、表 2.3 の「指示」を聞き、「判断することが前提である」という文言から、「指示」が「課題（プリタスク）」で示され、本文を聞いて「判断（解答）する」ことが求められる。日留試は、「課題」を「日本に留学するに当たって、あるいは、日本の大学に入学後直面する現実的な諸課題の総称である」（日本国際教育協会 2003, pp.246）と説明しており、その課題がプリタスクで示され、本文を聞いて解決することが求められている。旧日能試と日留試は、課題を与えられ、本文を聞いて解決するというテスト構成の点では同じである。

表 2.3 日留試と旧日能試のテスト課題の比較

日留試（全体）	旧日能試（聴解）
1.指示を実行する	「指示」を聞き、「判断」をすることが前提である
2.事物を特定する	1.情報完成 5.照合 7.事物・人物の同定
3.事物を描写する	7.事物・人物の同定
4.事物を比較・配列する	3.情報総合 6.順序の再構成
5.物事の推移・展開を予測する	4.意図の汲み取り
6.物事の背景や意図を把握する	4.意図の汲み取り
7.物事を構造化し、法則性を発見する	2.規則・法則・傾向の適用／検証

³⁶ 図 1.1 (p.5) と同じ図。

2.4.4 日本留学試験 聴解

この節では、日留試と日能試、また TOEFL などの日本語以外の聴解テストの比較をしながら論ずる。

日留試と旧日能試の相違点は、聴解テスト・アイテムの場面が、旧日能試は生活や仕事、学生生活など広く場面を設定しているのに対し、日留試は留学生が遭遇するであろう生活場面と大学場面に限っていることである。もう1点、日留試は技能に学習スキルをうたっている。一般的な学習スキルか大学における学習スキルかは明示されていないが、日留試の目的から考えると、大学における学習スキルであろう。旧日能試は、日常生活の中に教室内を含んでいるが、特に学習スキルを測るという言及はない。この点から、日留試は大学入学後に、つまりテストを受ける時点の能力だけではなく、将来大学で学習する際に必要となる能力（熟達度）も含めて測ろうとしており、旧日能試はテスト受験時まで学習してきた成果（熟達の到達度）を測ろうとしている³⁷と思われる。

日留試の聴解テスト・アイテムは、例8に見られるように「①状況説明」「②プリタスク」による「③設問」があり、その後「④本文」を聞き、「⑤ポスト・クエスチョン」「⑥選択肢」を聞き答えるという単形式である。旧日能試では視覚情報として選択肢が与えられているタイプもあったが、そのほとんどが絵や図であった。文字情報もあったが、単語レベルにとどまり読むスキルは排除されていた。聞くスキル以外を排除していることから、日留試の聴解テスト・アイテムは旧日能試の流れを汲んでいると言える。

例8 日留試 2008年2回目聴解1番

先生が、授業で、発表の準備について話しています。 ①
この先生に連絡する必要があるのは、どのような場合ですか。 ② } ③

この授業では、グループでテーマを決めて、順番に発表をしていってもらいます。
発表の時に使う資料ですが、コピーをするなら、事前に事務室に申し込む必要があります。前日までに申し込まないとコピーができないので、余裕をもって準備してください。また、発表の準備にデジタルカメラやビデオカメラを使おうと思っている人は、事務室で貸し出ししますので、予約をして借りてください。
それから、資料室にある本や研究論文を見たいときは、事前にメールで私に連絡してください。あの部屋は普段鍵がかかっていますので、当日、急に行っても開いていませんからね。 } ④

この先生に連絡する必要があるのは、どのような場合ですか。 ⑤

- 1. 発表用の資料をコピーするとき
 - 2. デジタルカメラを使用するとき
 - 3. 資料室の資料を見るとき
 - 4. 事務室に鍵がかかっているとき
- } ⑥

³⁷ 表2.1 (p.19) の集団規準準拠テストと目標基準準拠テストの特徴にあてはまる。

この3つの言語テスト（旧日能試、新日能試、日留試）の実施開始年を古い順に並べると、1984年（旧日能試）、2002年（日留試）、2010年（新日能試）となっており、各テストには少しずつ変化が見られる。表2.4（次頁）は、3つの言語テストの変化をまとめたものである。日留試には聴解と聴読解があり、聴読解には聴解能力が関わっているため、表に加えた。

表 2.4 聴解テスト・アイテムの比較：旧日能試、日留試、新日能試

	旧日能試（1984年開始）	日留試（2002年開始）	新日能試（2010年開始）
視覚情報（文字情報も含む）	問題II：無し	聴解テスト：無し	プリタスクがないアイテム：無し
	問題I：選択肢が絵か図、あるいは単語レベル	聴読解テスト：問題文と選択肢（図、絵、文字、グラフ、表、ちらし、広告など）と文字（単語あるいは文章）による4選択肢	プリタスクがあるアイテム：文字による4つの選択肢。課題理解では図なども用いられている。
本文前のプリタスク ³⁸ の有無	有り	有り	課題理解、ポイント理解：有り 概要理解、即時応答、統合理解：無し
1アイテムに対する解答の数	1アイテムに答えは1つ（多肢選択法）	1アイテムに答えは1つ（多肢選択法）	1アイテムに答えは1つ。統合理解の最後のアイテムのみ答えは2つ（多肢選択法）
アイテムの長さ	1～2分	1～2分 （聴読解も1～2分）	1～2分

表2.4を見ると、テスト・アイテムの形式上、設問（状況説明・プリタスク）と選択肢が事前に文字情報として与えられるようになったのは、日留試の聴読解が初めてである。これ以前は、聴解テストは聴解以外のスキルを測定すべきではないという考えに基づいており、文章レベルの文字情報を読ませることはなかった。日留試の聴読解の「読解」部分は、「読み」の時間を聴解時と別にとっているのではなく聴解をしながら見るため、非常に短く、種類も限られている（ポスター、案内、メモなど）。グラフや文字を伴った図も含まれており、日留試の聴読解の「読解」は「文字を含んだ視覚情報がある」という定義ではないかと思われる。一方、日留試の聴解は、聞くスキルのみで測定するという考えで構成されている。聴解テストで選択肢を文字で示し、視覚情報としたのは新日能試からである。多肢選択法で基本的に1アイテムに答えは1つ、聞くアイテムの長さは1～2分程度という点で変化はない。日留試の聴読解の文字情報は、旧日能試の問題Iの絵や図で4選択肢が示されるものと同様のタイプと、1枚のグラフや広告など全体が一度に視覚情報として見られるものの2つのタイプがある。後者は、コミュニケーション・テストの特徴である実際のコミュニケーションの視覚情報に近づいたと言える。新日能試

³⁸ 本研究では状況説明とプリタスクの両方をさして設問と定義しており、旧日能試と日留試、新日能試の課題理解とポイント理解には設問がある。新日能試の即時応答には設問がないが、概要理解と統合理解には状況説明がありプリタスクがない。

の課題理解で出題される図も、文章レベルの文字情報はないが、日留試の聴読解に近いタイプとなっている。日留試の聴解テスト・アイテムの形式は旧日能試の問題Ⅱと同タイプである点で、旧日能試からの変化はあまり見られないが、聴読解という文字情報を含んだ聴解が新たに加わった。そして、新日能試では、選択肢が文字で提示されるようになったことから、過去30年の間に、聴解テストに徐々に文字情報が取り入れられてきたことがわかる。

ここで、日留試と同じ目的のテストである英語テストの TOEFLiBT (TOEFL 公式 Web サイト : 国際教育交流協会 HP³⁹) と日留試の聴解テスト・アイテムの比較をしてみる (表 2.5)。

表 2.5 日留試と TOEFLiBT の聴解テスト・アイテム比較

	日留試	TOEFLiBT
視覚情報	無し	文字による (単語～文レベル) 4 選択肢
本文前のプリタスクの有無	有り	無し
1 アイテムに対する解答の数	1 アイテムに答えは1つ	1 アイテムに答えが複数 (4 選択肢から 2 択する場合もある)
アイテムの長さ	1～2分	会話 3分、 講義 3分～5分
アイテム数	20 (改定後 15)	会話 2～3 アイテム各 5 問、講義 4～6 アイテム各 6 問
聴解テストの時間	35 分程度 (改定後 23 分程度)	60～90 分

表 2.5 からわかるように、日留試と TOEFLiBT の聴解テスト形式はかなり異なっている。まず、日留試の 1 アイテムの長さがかかなり短いことがわかる。聴解テストの構成および長さも違うため、聴解の 1 アイテムの長さを同等に論じることはできないが、TOEFLiBT のように本文前のプリタスクがなく選択肢が書いてある場合は、1 アイテムを聞く時間の短縮ができる。日留試から得られるデータは 35 分で 20、TOEFLiBT は 60～90 分で 34～51 であり、データ数としては、さほど変わらない。日留試のようなタイプのテストの利点は、多くの話題や場面を出題できるため、受験者の個人差 (既有知識や興味の対象) の影響を拡散させることができる。TOEFLiBT の長所は、大学生の聴解場面に即しており、真正性が高く、必要とされる聴解能力の測定がしやすい。

次に、TOEFLiBT のように本文前にプリタスクがないアイテムの場合、課題予測能力・課題発見能力が期待されており、これは大学生の学習スキルに相当すると考えられる。日本の大学では、教師から常に課題が与えられるわけではないため、自ら課題を発見する能力が必要である。また、ある程度の情報量をもった談話の中から課題が何であるかを予測し、重要と判断した複数の事柄の整理、全体の流れの再構築など、大学生に必要なスキルと言えよう。

最後に、日留試では聴解スキルという単一のスキルのみの測定が絶対条件となっていることである。視覚情報がないアイテムだけが出題されているが、日常生活における聴解、特に大学生が受ける授業で視覚情報がないほうが稀である。広い視野で捉えれば、日留試には聴読解という分野もあるが、聴解テストにおいて、選択肢などの視覚情報を一切排除することが必要かどうかは疑問である。

³⁹ 2015 年 3 月 <http://www.cieej.or.jp/toefl/toefl/format.html>

日留試の特徴の分析と他のテストとの比較から、日留試の聴解テストには、プリタスクの有無、1アイテムの長さ、選択肢の視覚化など、テスト・アイテム形式の見直しの必要性があるのではないかと結論づける。なお、テスト形式については、第3章「3.4 テストの入力ファセット」(p.45)で詳しく述べる。

3章 聴解テストで測定される要素

3.1 アカデミック・リスニング

1章でも述べたが、英語教育での「アカデミック」という用語は、高等教育で必要とされるスキルに限定し、使用されている。しかし、日本語教育での「アカデミック」の定義は、まだ明確にされていない。日本語テストにおいて「アカデミック」という用語が注目され始めたのは、日留試のシラバスの中で「アカデミック・ジャパニーズ」が使用されてからである。しかし、日留試の日本語シラバスにはアカデミック・ジャパニーズとは何かという具体的な記載はない。そのため、現在でも様々な意見や解釈が存在している。

日留試の日本語シラバスでは、アカデミック・ジャパニーズは日本の大学で勉学に対応できる日本語能力と定めているが、それはまだ抽象的な理念にとどまっているとし、アカデミック・ジャパニーズを研究テーマとした「アカデミック・ジャパニーズ・グループ」という研究会が2001年に発足した。この研究グループでは、既に学部入学をした留学生と日留試の分析により、大学で必要とされる日本語能力を研究している。そして、「アカデミック・ジャパニーズ」に対する議論を深め、日本語学校、大学・大学院、海外の日本語教育機関との連携を図っている（堀井2009）。この研究会の発足を機会に、それまであまり連携が見られなかった日本語学校と高等教育機関の日本語教育が、海外の日本語教育機関も含め、連携を深めることとなった。さらに、日本語教育関係者以外の他分野の教師たちへも影響を与えている。このように日留試が使用した「アカデミック・ジャパニーズ」という用語が、大学で必要とされている日本語とは何か、留学生の日本語教育とは何かを再考させるきっかけとなった。以下に、複数の研究者が述べている「アカデミック・ジャパニーズ」、「アカデミック・スキル」についての定義をまとめる。

嶋田（2003, pp.18-20）は、留学生が大学入学後に求められる活動を、入学後からの時系列場面とともに列挙している。活動課題を、①人間関係の構築 ②情報の受信・収集 ③情報の発信 ④情報の交換 ⑤働きかけ の5つに分類し、課題達成のためには知識と言語運用能力（言語技能と技能選択・統合能力）が必要だと述べている。知識には社会文化知識、専門知識、言語・非言語的知識、コンピュータ・リテラシーを挙げ、言語技能には学業日本語運用能力、自立能力、学園情報獲得能力、そして技能選択・統合能力には能力操作能力を挙げている。

村上（2003, p.58）は、留学生の1年生と2年生にアンケートとインタビューを行い、学部留学生が習得しておく学習の困難が軽減する事柄として、①人間関係を形成する能力 ②講義でよく使われる語彙・表現 ③文章を書く力 ④カタカナ語 ⑤批判的思考能力 の5つをとりあげた。

因（2003, p.72）は、学部留学生に意識調査を行い、留学生に必要な技術として未知の項目への対処技能と複線的行動を行う技能（複数の情報源を同時に処理する）の2つが必要であると述べている。また、情報源が1つの場合、例えば音声のみの場合でも、「音が取れる」「表面的な意味が掴める」だけではなく、聞いた内容の解釈などを他の知識と関連づける、次の段階の準備をしつつ聞く、情報解析をしながら知識体系の中に位置づけるなどの作業が、講義や演習参加に必須であると説明している。

堀井（2003, pp.115-116）は、大学学部入学後の留学生の日本語の授業で使用されている教材を分析し、アカデミック・ジャパニーズとは「学習を行うための講義理解、レポートを書く、口頭発表をするなどの日本語能力（スキル）」、「学習を行うためのネットワーク作りのためのコミュニケーション力」であると述べている。そして、アカデミック・ジャパニーズを使用する際には、基礎知識と問題発見解決能力

とスキルも必要だとした¹。

門倉 (2003a, pp.127-128) は、大学の勉学に必要な日本語能力に焦点をおいて作成された教材を分析し、「講義を聞く技術、口頭発表・討論の技術、速読力、論理的な文章を書く力」を、アカデミック・ジャパニーズにとって欠かすことができない能力であると提言している。また、「勉学のためのストラテジー」だけではなく、「先生や学生との協力関係を作るためのストラテジー」の重要性にも触れている。

ネウストブニー (2003, pp.140-141) は、大学での言語活動を、相互行動のプロセスと捉え、アカデミック・ジャパニーズより広い枠組みとして「アカデミック・インターアクション」と呼ぶことを提唱している。アカデミック・インターアクションは ①文法能力 ②文法外コミュニケーション能力 ③社会文化的な問題 の3つの層からなり、日本語教育でこれらをどのように取り扱っていくべきかと疑問をなげかけている。

森 (2005, p.121) は、アカデミック・ジャパニーズとは「大学において授業（講義、演習、ゼミ、卒業研究等）をこなし、単位を取得することができる日本語能力である」と定義している。

アカデミック・ジャパニーズを測定するとうたった日留試に対し、倉持 (2004, p.283) は、日留試はすべてを測定する必要はなく、留学生の大学生活と学習活動上必要な基礎学力を測ればよいと述べている。

上記に挙げた研究者たちの見解から、アカデミック・ジャパニーズは、専門用語などを含んだ講義場面とキャンパス内でのコミュニケーション場面で使用される日本語であり、日留試が測定範囲としている留学生の生活場面での日本語は含まないと言える（図1.1「日留試の測定対象能力」p.5を参照）。アカデミック・ジャパニーズでは言語の構造的知識以外に、与えられた情報を複合的に利用し既存の知識や情報と関連づけて理解する能力、言語以外の社会文化的な理解、課題を自ら発見し解決する能力などが求められていると考えてよいだろう。

次に、本研究のテーマである聴解能力について、特に「アカデミック・リスニング」とは何かについて、どのような考えがあるかをまとめる。

日能試では、聴解テストの課題を ①情報完成 ②規則・法則・傾向の適用／検証 ③情報総合 ④意図の汲み取り ⑤照合 ⑥順序の再構成 ⑦事物・人物の特定の7つに分類している（『日本語能力試験出題基準』1994, p.201）（表2.3「日留試と旧日能試のテスト課題の比較」p.31を参照）。鈴木 (2003, p.39) は、この分類に従って日留試の聴解アイテムを分類した結果、ほぼ同じ分類となったと報告している。山本 (2003, p.76) は、社会科学系専門講義科目を担当する教員に、留学生の日本語に対する要望のアンケートをとった。その結果、聴解に関しては、講義やテレビのニュースなどが理解できる高度な日本語聴解力を求めているということがわかった。門倉 (2009, p.4) は、日留試の聴解テスト・アイテムでは、課題で設定された必要情報のみを聞き、そうでないものは聞き流すスキルを求められていると考え、読解の速読・多読と同様の速聴力・多聴力の必要性を述べている。大学では講義、発表、報告など、日常会話ではほぼあり得ない長さの聴解であり、内容も複合的なものを聞かなければならない。したがって、速聴力や多聴力の必要性に加え、短いやりとりの会話だけではなく、論理的なまとまった長さのスピーチの聴解にも力点をおくべきであると主張している。

Richards (1985, pp.198-199) は、アカデミック・リスニングを講義を聞く聴解力とし、日常会話と講義を聞くときの違いを下位スキルで示した。Flowerdew (1994, pp.11-12) も、講義理解のための聴解は日常会話理解とは違い、背景知識や講義の中から必要なもののみを選び出す能力であると述べている。そして、

¹ 堀井 (2006) では留学生に必要な日本語のコミュニケーション能力としてアカデミック・ジャパニーズ、キャンパス・ジャパニーズ、ライフ・ジャパニーズの3つに分類している。詳細は4章図4.2「留学生のコミュニケーション能力」(p.54)を参照。

講義の形式、談話構造、語彙や統語などの違いなど、講義を聞くためのアカデミック・リスニング研究の必要性を強調している(同掲p.25)。平尾(1999, p.14)は、講義聴解は聞いて理解するだけではなく、講義内容についての解説・考察・議論ができる生産的な聞き方であり、会話聴解とは性質を異にすると述べている。平尾(1999, p.15)は講義理解のために、文字資料や板書などの視覚情報の利用や未知語彙の処理能力も必要であるとしている。これらの研究者たちは、アカデミック・リスニングとは講義理解の聴解であり、「学術的・学問時」に行なう聴解であると考えている。「アカデミック・ジャパニーズ・グループ」の見解は、「大学内での場面において必要となる」聴解であり、「学術的・学問時」の場面より広い。日留試の解釈はさらに広く、「留学生が遭遇するあらゆる場面」での聴解である。鈴木(2003, p.39)は、日能試の聴解課題と日留試のアカデミック・リスニングには違いがないと報告し、倉持(2004, p.279)も、日留試の場面が大学生活に関するものが多いということ以外では日能試の1、2級の聴解と変わらないと述べている。これらの研究結果は、日留試で使用している「アカデミック・リスニング」の解釈が非常に広く、日常生活での聴解も含めていることを裏付けている。

本研究では聴解行動をアカデミック・リスニングの視点から分析・考察する。日留試が一般的な日本語能力を測定する日能試と袖を分かち、日本の高等教育機関への留学を希望する人たちに向けた日本語テストという点に注目すると、アカデミック・リスニングは大学の授業を受ける際に支障がない日本語能力と考える。よって、本研究では「アカデミック・リスニング」を「学術的・学問時に行なう聴解」と定義する。

3.2 聴解能力と聴解テスト

まず、聴解能力について様々な研究者の意見をまとめる。

竹蓋(1984, pp.43-44)は、「聴解」の定義を3つの言葉を挙げて説明している。

Hearing : 聴覚によって受動的に聞こえる心理的な感覚

Speech perception : 内容の理解を含まない、音素の識別

Listening : 受動的な感覚と能動的な情報処理、認識力の総合的活用によるメッセージの生成、メッセージ内容の理解や解釈を生成する行為

言語教育研究で扱うのは「listening」である。竹蓋(同掲pp.61-69)は、「行動としての聴解²: listening」の特徴として、①ノイズに強い ②選択的聞き取りができる ③聞き取りにくくても短時間で向上する ④補正する ⑤総合的行動である を挙げている。最後の「総合的行動」とは、音を聞き取るという単純な行動ではなく、以下のような多くのことが同時に行われていることをさす。

1. 音を聞く前から次の音や内容について予測する。
2. 音が同時調音となることや間違いの訂正などをしながら聞く。
3. 聞き取った音をしばらく記憶しておく。
4. 音の組み合わせを既知の語彙リストと照合する。
5. 語彙が集合体となり、発話の切れ目があるとそれが文となるかどうかを確認する。
6. 文として成立すると仮定した場合、メッセージはその場のTPO、超言語的情報、体言語などで、伝達された内容とあっているかをチェックする。
7. もしあっていなかった場合、音の聞き取り、語彙照合、その後の処理に誤りがなかったかを再度

² 受動的な Hearing (話し手は発話を文法規則に従い、音素を並べ、発音器官で信号化する、聞き手はそれを聞いてもとの音素の並びに復元すればよい) の感覚と、能動的な情報処理能力、認識力の総合的活用によるメッセージの生成、さらにはその内容の理解、または解釈を生成する行為を「行動としての聴解」と定義している。(竹蓋 1984, pp.43-44)

確認する。

これに発話が加われば、これから話そうとしている内容や表現方法、発声準備などが重なる。このように「行動としての聴解」は、完全な状況、完全な音声情報で発せられるものを聞くのではなく、様々な問題点を調整して聞く行為のことである。また、竹蓋（同掲pp.68-69）は「聴解」には以下のような条件が必要だと述べている。①生理的条件（聴力があり、集中して聞けるなど） ②心理的条件（知力、思考力、記憶力、興味など） ③言語的条件（音素識別力、語彙力、文法力、意味の理解力など） ④コミュニケーション行動的条件（コミュニケーション機能の理解、経験、聞き分け、推測、総合的な行動ができるなど）。梅村（2003,p.140）は、聴解にとって有効な知識は、①音を正確に弁別する（識別する）ための音声の知識 ②語句や文の意味を理解するための語彙的・文法的知識 ③要点をおさえ論旨を的確に把握するための文章論的知識 が必要だと述べている。梅村の挙げた3点は、竹蓋の「③言語的条件」と同様の内容である。竹蓋は、聴解能力とは、生理的・心理的・行動的なものが伴った言語知識であると捉えている。

阿部他（1994）は、言語理解に使用する知識は、①与えられた言語表現 ②文脈・状況 ③心内知識（言語知識・世界知識³）としている。これらの知識の他に、聞き手の心理的、生理的、物理的な能力や性質、状態も理解に影響すると述べている。聴解とは、これらの言語知識および世界知識を利用して言語表現の理解と言語表現の背後にある状況を理解するのである。その際、一時的・短期的な記憶（working memory）として、心の中に意味の表象（話者の世界のモデルと話題の世界のモデル）を作る。（阿部他1994, pp.11-17）

Buck（2001）も、聴解理解には言語的知識（音韻、語彙、統語、意味、談話構造）と非言語的知識（話題、内容、一般常識）が使われると述べている。理解の仕方には、音、語彙、統語的なものから理解するトップ・ダウンと、トピックや常識などから類推していくボトム・アップの2つがある。言語の応用知識は宣言的知識（declarative knowledge）と手続き知識（procedural knowledge）の2つに分けられる。宣言的知識とは、事実や事柄の知識から何を行うかを作動記憶で認知することであり、それらを宣言的記憶に蓄積する。手続き知識とは、どう行うのかの知識であり、同じことが起こったときに、宣言知識が作動記憶に取り出され、何をすべきかを生産的記憶に送る。これが自動化されると、宣言的記憶を使わなくなる。特に聴解テストでは、生産的知識を常に意識している（図3.1）⁴。聴解では、言語理解だけではなく、テキストの文脈や常識・経験などからの推論、社会的に適切かというコミュニケーションの文脈からも意味の構築を行う。聴解とは、自動的な処理ができること、聞き手のあらゆる知識を使うこと、文脈の解釈をすることである。聴解は、聞き手によって違う推論をたて、違う解釈をすることから、個別的であり、個人的な過程である。（Buck, 2001, pp.1-30）

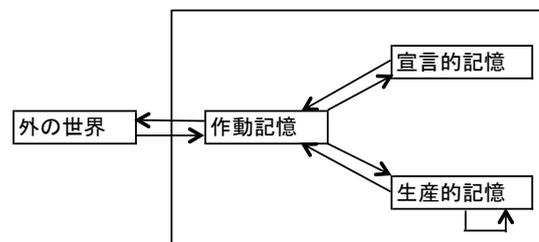


図3.1 ACT-Model (Anderson, 1983, p. 19)

上記の研究者たちの「聴解」の定義に共通してみられるのは、聴解には、言語知識、世界知識、社会

³ 知識、一般的常識をさす。例えば、事物の上位下位関係や部分全体関係、状態や出来事の因果関係や時間的順序関係、社会慣習・風俗・文化・宗教・価値観など、そして専門知識などである。

⁴ ACT(Adaptive Control of Thought) Model は Anderson (1983) *The Architecture of Cognition*, Cambridge, MA : HUP による認知学習モデルのことである。（白畑他 1999 『英語教育用語辞典』 pp.11-12）

言語学的知識が必要だという点である。社会言語学的知識とは、竹蓋がコミュニケーション行動的条件とし、阿部他が文脈・状況、Buckがコミュニケーションの文脈と説明している知識である。外国語の聴解の場合、この知識の有無、あるいは知識量の程度の違いが、理解に大きく影響を与える。文脈や場の解釈、コミュニケーション機能などが母語や自文化と同様である場合と異なる場合、あるいはその言語の文化や社会の理解度の深さにより、聴解の理解力に差がでると言えよう。

次に、聴解テストについて各研究者の見解をまとめる。

聴解テストは言語テストの変遷（2章「2.1言語能力の定義と言語テストの変遷」p.10を参照）と同様、指導法の影響が大きい。竹蓋（1984, pp.187-188）は、言語評価の流れは主観的評価から客観的評価へ、部分的アプローチから統合的アプローチへ、そして概念・機能アプローチやある目的に特化したアプローチへ、さらに学習者の個別ニーズを重視したアプローチへと進んできたと述べている。しかし、この流れは、常に前の考え方や方法に勝るものではなく、前の代で欠けていたものを重視し生まれたものであった。よって、聴解テストは1種類の問題形式ですべてを評価するのではなく、複数の形式で総合的、診断的にすべきであろうと結んでいる。

菅井（2006b, pp.6-11）は、Morrow（1979）が提示したコミュニカティブ・テストの概念と課題が、3つの研究方向に分かれたと説明する。1つ目は、Canale & Swain（1980）の言語的特徴からのコミュニカティブ能力の定義（以下、4点）であり、この定義からテスト設計へと発展した流れである。

1. 文法あるいは言語形式に関する能力（項目別テストで用いられた）
2. 社会言語学的能力（状況にあわせる）
3. 方略的能力（言語能力の欠如を補うための行動）
4. ディスコース能力（文脈の中の修辞関係など）

2つ目は、Munby（1978）、Aitkin（1978）、Richards（1983）、Weir（1993）のコミュニカティブ・アプローチによる下位技能分類への流れである。コミュニカティブ能力の定義のみでは、どのように実際の行動にあらわれるかがわからないため、Morrowが定義した「この受験者は言語を使って何ができるかを知る」ことができない。そこで、目的・場面・人物などを設定したタスクを受験者に与え、その行動から言語能力を推測するという考えが生まれた。彼らはコミュニカティブ活動を行っている際に、どのように言語的な情報処理をしているかという点に着目し、下位スキル（micro-skill/sub-skill）⁵の分類をした。Munby（1978, pp.123-131）は、言語スキルを54グループに分け、さらにそれぞれのグループを下位項目に分け、260項目に分類した。Richards（1983, pp.198-199）は、聴解下位スキルとして、カンバセーショナル・リスニング33項目、アカデミック・リスニング18項目に分類した（分類項目詳細は表3.1「カンバセーショナル・スキルとアカデミック・スキル」p.42を参照）。Weir（1993, pp.98-99）も、聴解下位スキルの分類を行っている（分類項目詳細はp.43を参照）。3つ目はできるだけ自然なコミュニケーションをサンプリングする（sampling）という流れである。このサンプリングにより、現実性の高いタスクが作られれば、テストの妥当性は高くなると考えられた。McNamara（2000邦訳p.29）⁶は、テスト設計の基礎として実際のコミュニケーション上の役割や課題を研究する職業分析（job analysis）が有益であると考え、職業分析から得られたタスクに評価表による基準付けをすることを勧めている。

本研究では、上記の研究の2つ目である下位スキル分類を聴解テストに用いることで「この受験者は言語を使って何ができるかを知る（Morrow）」ことができ、テストで測定する聴解能力を明確に表せると

⁵ macro-skills は「話す・聞く・読む・書く」の4技能である。

⁶ 菅井（2006b, p.11）はMcNamara（1996）*Measuring second language performance*, UK: Longman を参照している。

考え用いることにした。

3.3 聴解下位スキル

武井 (2002, p.43) は、多くの聴解の構成要素研究の結果から、①総合聴解力と下位能力には相関があり、独立性を保ちながら複数の下位能力からなる可能性が高い ②総合聴解力を構成する下位要素は、語彙力、語認知速度、背景知識、音素識別力、記憶力、母語聴解力などが有力であると述べている。聴解能力には、下位能力・下位スキルが関わっていること、本研究で下位スキル分類を用いることから、この節では、下位スキル研究についてまとめる。⁷

Buck (2001, pp.53-54)は、Aitken (1978)が挙げた以下のコミュニカティブ・アプローチの聴解スキルを、聴解の完全なスキルではないかもしれないが、話者の話を処理し理解するためには必ず必要であると述べている。

- ①語彙の理解、わからない語彙の推測 ②統語パターン、形態の型、談話パターン理解
- ③ストレス・非ストレス、イントネーション、句読の理解 ④話者の目的の認識
- ⑤社会的状況、話者の意図、文脈から整理・推測 ⑥話者の聴者に対する態度と議論の話題を認識
- ⑦話者が使用しているテクニックとレトリックの認識

Rost (1990, pp.152-153)は、聴解の下位スキルとして1.~5.の分類の中に22のスキルを挙げている。

1. 知覚：発話内プロミネンスの認知 ①1つ1つの単語音の識別と知覚 ②強弱、ストレスのない母語の削除、音の変化（同化、省略、リエゾン）、音韻変化、異音の知覚 ③言葉をつなぐためのストレスやピッチの使用（情報の塊やリズム・パターンの提示、強調やコントラスト） ④話し手の違いに適応
2. 解釈：話し手の発話命題の意味の構築 ⑤よくわからない語彙の意味の推測 ⑥省略された情報をもとに明示的に述べられていない情報を橋渡的に推察 ⑦複数の命題の関連性を推測
3. 解釈：発話に関連したフレームワークの構築 ⑧アイデア紹介やトピック変更、ポイントの強調と明確化と拡大、反対意見などの談話の指標を認識 ⑨メイン・ポイントと説明的あるいは些末なことを区別して談話のメイン・アイデアやテーマを構築 ⑩概念的レベルで談話の次に来る部分を予測 ⑪談話内から聞き手の助けになる要素の識別 ⑫談話の命題の予測と検証をするために文脈の連続性を維持 ⑬スキーマを完成するために話し手のテキストから手がかりを選択
4. 解釈：発話の際の話し手の意図の解釈 ⑭聞き手に対する話し手の意図を示唆している対人フレームを識別 ⑮「韻律的ジェスチャー（ピッチの高さ、音域、ピッチ・パターン、ポーズ、テンポ）」の変化を認識し、話し手が使ったこれらのジェスチャーの不整合を識別 ⑯話し手の矛盾、不十分な情報、発話の曖昧さの識別 ⑰比喻、皮肉、談話上の一般的な法則の「違反」を識別して、事実と意見を区別
5. 実行：適切な応答のための談話表現の利用 ⑱タスクで使用するために与えられた情報から、顕著な点を選択 ⑲音声ソースから他の形態（例えば、ノートを取るときの書く形式）にする際、情報を削減し、コードを変換 ⑳トピックやアイデアの明確化の必要性を識別 ㉑テキストやその他の情報源からの情報を統合 ㉒話し手に適切なフィードバックを提供

⁷ Munby (1978)の分類は260項目もあるため、ここでは省略する。

Rostの下位スキルとして注目すべき点は「②話し手に適切なフィードバックを提供する」、つまり聞き返しなどのインターアクションも含めていることである。Rostの分類に対し武井(2002, p.43)は、テスト作成の参考にはなるが、あまりに細分化されているためもっと絞り込むべきだと述べている。

Richards(1985, pp.198-199)は、聴解をカンバセーショナル・リスニングとアカデミック・リスニングに分け、それぞれ33と18の下位スキルを挙げている(表3.1)。

表3.1 カンバセーショナル・スキルとアカデミック・スキル(Richards, 1985, pp. 198-199)

(【 】:Richardsの英文リストにはなかったが、筆者が説明を加えた)

下位スキル	カンバセーショナル・リスニング
1.	長さが違う言語の固まりを短時間覚えていられる
2.	目標言語の特徴的な(弁別的な)音が識別できる
3.	言葉のストレスパターンが認識できる
4.	英語の律動的構造が認識できる
5.	発話内の情報のサインである音の強弱とイントネーション機能が認識できる
6.	音を強く言ったり弱く言ったりする時の言葉を明確に捉えられる
7.	音が縮約された言葉の形が認識できる【短縮形】
8.	言葉の境界線がわかる(分かち書き・熟語など)
9.	目標言語の基本的な語順がわかる
10.	会話のトピックで中心的に使用される語彙がわかる
11.	キーワードが見つけられる(トピックや主題が何であるかわかるため)
12.	文脈から言葉の意味が推測できる
13.	内容語・機能語がわかる(会話部分)
14.	主なシンタクスパターンや仕組みがわかる
15.	会話の結束装置がわかる
16.	文法単位の固まりの省略形や文の省略形がわかる【省略とは抜いてしまうこと。つまり文法的にはなくてはならないのだが、抜いてしまう】
17.	文の構成を見つけることができる
18.	構成上、主となる部分、副となる部分が見分けられる
19.	文法型・センテンスタイプの違いで意味の違いを表現することがわかる(特別な意味を表すには違う言い方がある)
20.	場面、参加者、ゴールにより発話のコミュニカティブ機能がわかる
21.	場面、ゴール、参加者、進行具合などの再現あるいは推測ができる
22.	目的、ゴール、設定、進行具合などがわかるために現実の知識や経験を使うことができる
23.	出来事の描写から結果を予測できる
24.	出来事同士の関係やリンクを推測できる
25.	出来事の原因や結果を推理できる
26.	文字どおりの意味なのか慣用句的な言い回しなのか見分けられる
27.	複数の話者たちで会話されている内容を聞きながらトピックや構造上の一貫性を再現したり認識したりできる
28.	談話上の一貫性(文法指示ではなく、推論によってわかる)がわかり、主題、裏付け、与えられた

<p>情報、新しい情報、通念、具体例の関係を見つけられる</p> <p>29. スピーチ速度の違いを処理できる</p> <p>30. 会話におけるポーズ、間違い、言い直しが処理できる</p> <p>31. 意味を理解するために表情やパラ言語や他の手がかりを使うことができる</p> <p>32. 聞き手の目的やゴールによって違う聞き方をすることができる</p> <p>33. 理解しているのかしていないのかということを言語的あるいは非言語的に知らせることができる</p>
<p>下位スキル アカデミック・リスニング（講義を聴く）</p> <p>1. 講義の目的や領域がわかる能力</p> <p>2. 講義のトピックがわかり、トピックの展開についていける能力</p> <p>3. ディスコースの中の複数の構成要素の関係がわかる能力（主題・通念・仮説・補佐的アイデア・例など）</p> <p>4. 講義談話の特徴的なディスコースマーカーがわかる能力（接続詞・副詞・切り出しの言葉・常套句など）</p> <p>5. 関係を推測する能力（理由・結果・まとめなど）</p> <p>6. トピックに関するキーワードを見つけることができる能力</p> <p>7. 文脈から言葉の意味を推測する能力</p> <p>8. 結束性のマーカーを見つけることができる能力</p> <p>9. 大切な情報のある部分にかかるイントネーションの機能がわかる能力（ピッチ・音量・ペース・音の高さ）</p> <p>10. 話題に対しての話者の態度を見極める能力</p> <p>11. 様々な話の形式についていける能力（対話・テープのような聴覚のみ・テレビのような視聴覚）</p> <p>12. アクセントとスピードが違う話についていける能力</p> <p>13. 話の様々なスタイルを知っていること（フォーマル・会話・読み聞かせ・無計画）</p> <p>14. 言語使用の違いを知っていること（書き言葉・話し言葉）</p> <p>15. 関連事項がわかる能力（冗談・余談・漫談）</p> <p>16. 強調のためのノンバーバルサインや態度がわかる能力</p> <p>17. 教室内での約束事を知っていること（ターンテイキング・質問を明確になど）</p> <p>18. 教授や学生の話の機能を認識する能力（警告・提案・推薦・アドバイス・説明など）</p>

Weir (1993)は、言語テスト作成に関して注意すべき点をまとめている。彼はテスト・タスクは直接的であり、現実世界の場面を使用して、できる限り現実性を強調すべきだと述べている。テスト・タスクが現実的であれば、より肯定的なウォッシュバックが得られ、学生がその言語を使って何ができるかを判断しやすいからである（同掲 p.28）。Weir (1993, pp.98-99)は、聴解テストを作成するために聴解力判断のチェックリストを4つのカテゴリーに分け、さらに具体的な下位スキルを17挙げている⁸。

1. 概略や重要な情報などを区別して聞く「意味の直接理解」：①概略を聞く ②主旨または重要な情報を聞く（この中には、議論の展開、主旨と補助部の区別、例と陳述の区別、命題とそれへの反論の区別、事実と意見の区別を含む） ③部分の理解（重要な細部の呼び起こしを含む） ④話者の聞き手あるいはトピックに対する態度／意図

⁸ 邦訳は菅井（2006b, p.8）を引用。

2. 推測やコンテキストからの演繹をする「意味の推測による理解」：⑤推測と演繹（テキストから得られる情報によりコンテンツの評価もする） ⑥発話を社会的／状況的なコンテキストと結びつける ⑦発話のコミュニケーション機能を認識する ⑧コンテキストから、未知の語彙の意味を演繹する
3. 音韻、文法、語彙、文構造を理解する「言語的特徴の補完による意味の理解（微視的言語学⁹ microlinguistic）」：⑨音韻的特徴の理解（ストレス、イントネーションなど） ⑩文法的形式として比較／原因／結果／程度／目的などの概念を理解 ⑪ディスコース・マーカ－の理解 ⑫文構造（文を構成する名詞、動詞とその修飾、否定形）の理解 ⑬文法的結束性と指示表現 ⑭語彙同士の結束性と連語法 ⑮語彙理解
4. 要約、主旨などを抽出できる「聴解と記述（講義内容、電話の会話をノートに書き取る）」：⑯テキスト全体を要約するために重要な部分を抽出する能力／聞いたものの主旨のアウトラインや重要な細部へと絞る能力 ⑰情報が連続している中で、知りたいアイデアやトピックについてのテキストから、関係のあるポイントを選択的に抽出する能力

Buck et al. (1997)¹⁰は、TOEICの多肢選択肢の聴解テスト30アイテムを分析し、14の能力を取り出した。①速いインプットを分類調査する能力 ②低い頻度の語彙を処理する能力 ③より高い語彙密度でテキストを分類調査する能力 ④より複雑な構造を処理する能力 ⑤より長い文節を処理する能力 ⑥より高い情報密度でテキストを分類調査する能力 ⑦聴く目的を決定する短い文節をスキャンする能力 ⑧点在している情報を総合する能力 ⑨冗長な情報を使用する能力 ⑩ワード・マッチング・ストラテジーを使用する能力 ⑪表面的な言葉の関連付けに抵抗する能力 ⑫任意の名称を思い出す能力 ⑬テキストをもとに推論をする能力 ⑭推測するために背景知識を使用する能力

さらに、Buck & Tatsuoaka (1998, pp.141-142)は、短答形式の聴解理解問題を分析し、15能力を抽出した。①タスクを完成するためにどんなタイプに関する情報を検索するかを決定することによってタスクを特定する能力 ②自動的にリアルタイムで比較的速く話されたテキストをスキャンする能力 ③比較的大きい情報負荷を処理する能力 ④中くらいの量の情報負荷を処理する能力 ⑤比較的難解な情報を処理する能力 ⑥情報の位置確認のために前の項目を使用する能力 ⑦明確なマーカ－なしで関連情報を識別する能力 ⑧強調された音（アクセント）を利用するあるいは理解する能力 ⑨比較的速いテキストを自動的に処理する能力 ⑩テキストをもとに推論する能力 ⑪背景の知識をテキスト処理に組み入れる能力 ⑫L1の文字通りの意味ではなくL2の概念で処理する能力 ⑬冗長な情報を認識して使用する能力 ⑭テキスト中に点在する情報を処理する能力 ⑮比較的迅速かつ効率的に応答を構築する能力

聴解の下位スキル分類は聴解テストを作る際に参考になり、聴解とは何かを問うことができる。聴解とは、様々に異なる見方ができる大量の副構成要素を持った多様な過程なのである。聴解テストでは、基本的な言語的情報を取る能力と、広い文脈の中で判断する能力の両方を測らなければならない。(Buck, 2001, p.59)

⁹ 筆者訳

¹⁰ Buck, Tatsuoaka, Kostin & Phelps (1997), The sub-skills of listening: Role-space analysis of a multiple-choice test of second language listening comprehension, In Huhta, Kohonen & Luoma (eds.), *Current developments and alternatives in language assessment*, Tampere: University of Jyväskylä

上記のように、多くの研究者が下位スキルの分類を行っている。しかし、これらの下位スキルは、実際の裏付けの検証が得られていないという批判もある（静・竹内・吉澤2002, p.169）。では、なぜ下位スキルを研究するのだろうか。Buck (2001, p.59)は、言語スキルの複雑な過程をあらゆる視点から眺めることでテストを作る際に参考にできるからであると説明している。

3.4 テストの入力ファセット

Bachman (1990 邦訳 p.392) は、言語テストの得点に影響を与える要因として、①意志伝達言語能力（言語能力・方略的能力・心理生理的機能）②テスト方法のファセット（facets）¹¹（環境・受験の指示・入力・期待される応答・入力と応答の関係）③個人特性（文化背景・背景知識・認知能力・性・年齢）④ランダム因子（CLA¹²構成要素、テストの方法のファセット、個人特性間の交互作用・測定誤差）を挙げている（図 3.2）。

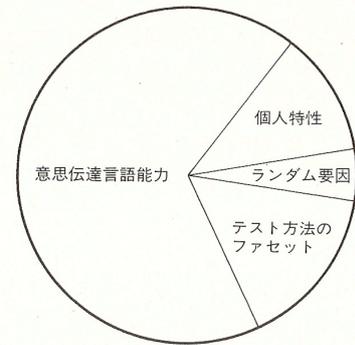


図 3.2 言語テスト得点の変動要因 (Bachman, 1990 邦訳 p. 392)

Bachman (1990 邦訳 pp.137-175) は、テスト方法がテストの成績に影響を与えるとして、

1. テスト環境：場所や道具への慣れ、検査者への慣れ、テストの時間帯、物理的条件
2. 受験の指示：テストの構成、時間配分、指示
3. 入力：形式、言語の性質、ジャンル、組織的特性、語用論的特性
4. 期待される応答：形式、言語の性質、応答の制限
5. 入力と応答の関係：相互的、非相互的、適応的

の5つのファセットを挙げている。この5つを、日留試の聴解テストで考えてみる。

(1) テスト環境

テスト実施場所は、教育機関などの教室を使用する事が多いが、ほとんどの受験者にとっては未知の場所である可能性は高い。聴解テストの場合、スピーカーからの一斉放送であるが、受験者が教室で学習しているときも、CDプレイヤーなどのスピーカーから聞いていると思われるため、慣れていると考えられる。検査者は、テスト実施団体の人たちであり、初見である。テストの時間帯は各国により午前、午後などばらつきがあるが、早朝や深夜という時間帯ではなく、日常的な活動時間帯である。物理的条件については、各受験場所によってかなりばらつきがあるだろう。日中の試験であるため、無音状態ではない。座席とスピーカーの距離により多少の違いはあるかもしれない。空調の有無や状態なども関係する。テスト実施側は、一定の条件を確保するため、騒音がないような場所、放送の音は質が高いこと、光や気温などで受験者に負の影響が出ないよう最大の努力をしている。

(2) 受験の指示

テストの構成だが、当初は「記述・聴解・聴読解・読解」であり、2010年の改定で「記述・読解・聴読解・聴解」と順番が変更された。聴解は集中力が必要なため、疲労がたまってくるところであろうテストの最後に移したことはあまり評価できず、実施側の説明が必要である。聴解テストのアイテム数は、当初

¹¹ テスト得点に影響する測定上の条件（静他 2002, p.79）

¹² communicative language ability の略

は20であったが、2010年からは15と減り¹³、その後のテストではこれで固定されている。テスト・アイテムの正解率、困難度は公開されていないが、表面的に見る限りでは困難度の簡単なアイテムから難しいアイテムの順に出題されているようである。アイテムの長さは1～2分である。一問一答であり、点数に重みづけがない。多肢選択法（4選択肢）で正答は1つである。マークシートの塗り方は、4つの選択肢全てに正しいか正しくないかをマークする方法で、旧日能試の聴解テストの解答方法と同様である。4つすべての選択肢を塗りつぶす解答方法をとっている理由は、聴解テストの場合、正答の選択肢が聞こえた瞬間に多くの受験者の塗る行動がヒントとなることを避けるためである。解答の仕方については例題が与えられている。受験者は、問題形式やマークシートなどを、公開されているテスト冊子で事前に確認できる。

（3）入力 （4）期待される応答

「入力」と「期待される応答」には「形式・言語の性質」のように類似性が見られる。また、入力に対して応えるという相互作用であるため（Bachman, 1990 邦訳 p.146）、（3）と（4）をまとめる。

入力形式は、録音された日本語（目標言語）を聞く聴覚のみである。応答形式は、多肢選択のマークシート（非言語）である。多肢選択法のような特定の型の言語テストは、それ自体がジャンル¹⁴とみなされ、これに慣れている受験者は慣れていない受験者より有利である。選択肢の入力は、日本語（目標言語）で聞く聴覚のみである。1テスト・アイテムの長さは1～2分程度だが、選択肢の長さは1単語のみのように短いものではなく、節から文レベルであり、短期記憶力の優劣が影響する可能性がある。命題内容は、語彙知識の量や文脈化の程度¹⁵に影響される。ただし、聴解テストの談話に一般的知識や概念ではない、あるいは専門的な語彙が含まれている場合は、受験者による既知の情報差の影響を少なくするため、談話内で説明が加えられている。情報の型は「具体的か抽象的、肯定的か否定的、事実的か反事実的」で困難度が変化する可能性がある¹⁶。話題は、受験者によってバイアスがかかってしまうため慎重にならざるをえないが、バイアスを避けようとして、誰もが何度もよく耳にする一般的な話題ばかりを選ぶと、受験者にとっては新情報がなくつまらなく感じる話題になってしまうことが生じる。この問題を解決するには、できる限り多くの数の話題を提示することであるが（Bachman, 1990 邦訳 p.159）、2010年以降の日留試の聴解は15テスト・アイテムと減少し、十分な数とは言いがたい。組織的特性については、アイテム・ライターが作成した文法・結束性・修辭的構造が組織化された談話である。語用論的特性とは真正性であり、日留試は日本での生活場面と、日本の大学内で留学生が遭遇する場面の談話で構成されている。この点では、日本での滞在経験や留学経験を持っている受験者のほうが有利だと考えられる。最後に、社会言語学的特性（方言・言語的変種、言語使用域）についてであるが、日留試の場合、方言や言語的変種は使用していない。言語使用域は、演説や講義などのフォーマルなものから、友人とのおしゃべりのようにくだけたものまで網羅している。2人の会話に登場する人物には、留学生が談話内の登場人物とはなりにくい、また、あまり接しない相手である親子、幼稚園児同士、老人同士などは含まれていないと思われる。講師と学生、先輩後輩など、コミュニティにおける上下関係で言葉遣いが変わるなど、日本の縦社会の知識の有無の影響はあるだろう。したがって、日本人特有の人間関係の捉

¹³ 改定前では聴解と聴読解のテスト時間が長く、集中力が続かないため改善したと、2009年12月に日本学生支援機構のHPで説明している。（2015年3月 http://www.jasso.go.jp/eju/documents/kaitei_09_12.pdf）

¹⁴ 形式的な特性を持っている文体構造。例えば、詩、物語、ことわざ、祈り、演説など（Bachman, 1990 邦訳 p.159）。

¹⁵ 文脈化の程度とは、談話の中の新しい情報と、関連のある既知情報との比である。例えば、登山の話題の場合、登山についての知識や経験がある受験者にとって文脈化の高いものとなる。

¹⁶ 2者のうち（〇〇か〇〇）、後者のほうが一般的には難しい。

え方などを知っている、あるいはそれに近い文化圏の受験者のほうが有利である。聴解テストの音声源はノイズが排除されており、ノイズによる妨害はない。日留試の聴解テストは、与えられたタスク（設問）の正答を多肢選択肢から1つだけ選ぶという形式のため、自由解答のような非制限性はなく、解答の多様性もない。正答を選ぶ時間は一定であり、聴解テスト・アイテムをすべて聞き終えた後で、解答を見直す時間はない。このように人工的な制限が加えられているため、通常の言語使用における多様性のいくつかは失われ、同時に真正性も減少してしまう。

（5）入力と応答の関係

聴解テストでは、与えられた情報に対してインターアクションをおこすことはできない。よって、入力と応答は非相互の関係である。また、日留試の聴解テスト・アイテムは誰に対しても同じものであり、個々の解答によって次のテスト・アイテムの困難度などに変更が生じないため非適応的である。

上記の（1）から（5）でテスト・アイテムが網羅しているのは、（3）入力と（4）期待される応答である。入力形式の種類では、聴解本文の前のプリタスクの有無、場面や状況説明の有無、文字・イラスト・写真などによる視覚情報の有無が挙げられる。また、多肢選択法に関しても、文字提示か音声提示かの違いがある。赤木（2003, p.52）は、様々な聴解テストの設問（場面・状況説明とプリタスク）や選択肢の提示方法の違いを調査した（表3.2）。この表を参考に、日本語大規模テストの特徴を考察する。

日本語テストであり、かつ日本で作成されたテストは日留試、旧日能試、ジェトロビジネス日本語能力試験、新日能試である。SAT II は、アメリカ合衆国の大学入試資格試験として用いられており、日本語テストは語学の分野に含まれる。JPT は、韓国で実施している日本語テストである。日本で作成された聴解テストは新日能試を除くと、設問（場面・状況説明とプリタスク）と選択肢は音声提示であり、タスクは本文の前後で二度聞くことができ、一問一答という共通点を持っている。

初期の日能試聴解テスト・アイテムは本文を聞いたあとに課題がくる。加藤（1992）は、聞く目的を知らない場合、すべてに集中して聞かなければならず、内容を一時保存しておく記憶にかなり頼る聞き方になると述べている。これは受験者にとって負担になる。現実の聴解活動では、目的に応じた聞き方をするのが普通である。同様の状況にするため、聴解テストでは、これから聞く内容についての文脈を与え、内容予測や聞く動機付けを行う。旧日能試が1990年からプリクエスチョン（状況説明とプリタスクを含む設問）の形式をとったことは改善である。プリクエスチョンが、受験者の階層化された知識・スキーマ（schema）の発動を助けるような機能を持ち、この形式の出題で聴解テストは現実度（authenticity）が高まった。聴解の目的（goal）、トピック、出来事の進行（frame）、登場人物の意図的行動（plan）、設定場面での解答（script）といった手続きが、プリクエスチョンによって円滑に把握できるからである。（加藤 1992, p.35）

大坪・三枝（1992, p.20）は、プリタスクをより日常的な聞き方であると評価しているが、テスト・テクニックに頼る受験者も生まれ、負の効果もあると述べている。彼らは、プリタスクが解答への誘導を行い、プリタスクで指定された部分以外を聞かなくさせてしまうと報告をしている。設問が与える影響についての記述が、聴解アイテムを解答する際の認知過程を明らかにしようとした古川他（1992, p.175）の研究にある。4人の対象者に聴解テストを行い、その後どのように解答をしたかを think-aloud で語ってもらった。その結果、設問があると解答のための情報が得られやすく解答しやすいと肯定的に捉えた意見と、設問があると解答となるポイントしか聞かなくなるため望ましくないと否定的に捉えた意見の両方が見られた（同掲 pp.297-359）。

表 3.2 聴解テスト・アイテムの設問、選択肢

(赤木 2003,p.52 を参考に新日能試、TOEFL CBT、IELTS を筆者が付加)

テスト名	1 答か複答か	場面・状況説明	タスクの位置	選択肢の提示方法
日本留学試験	1 問 1 答	有	前後	図・単語程度 (問題 I) 音声 (問題 II)
旧日本語能力試験	1 問 1 答	有	前後	音声
新日本語能力試験 (N1・N2 のみ ¹⁷)	1 問 1 答 1 問 2 答 (統合問題中 1 アイテムのみ)	有 無 (即時応答)	前後 後 (概要理解・統合問題)	文字 (課題理解・ポイント理解) 音声 (概要理解・即時応答・統合問題中 1 問 1 答の 2 アイテム) ¹⁸
ジェトロビジネス 日本語能力試験	1 問 1 答	有	前後	音声
実用日本語検定 会話説明問題	1 問複答	無	後	音声
SAT II (日本語)	1 問複答	有 (英語)	後	文字 (英語)
JPT 日本語能力試験 会話・説明問題	1 問 1 答 / 1 問複答	無	-	文字
TOEFL	Part 1 1 問 1 答	無	後	文字
	Part 2 1 問複答	有	後	文字
	Part 3 1 問複答	有	後	文字
スペイン語検定試験 DELE	Inicial 短い対話 1 問 1 答	無	後	図・写真・単語程度
	Inicial 1 問複答	有	文字	文字
TOEFL CBT	Part A 1 問 1 答	無	後 (文字併用)	文字
	Part B 対話 1 問複答	無	後 (文字併用)	文字
	Part B レクチャー 1 問複答	有 (文字)	後 (文字併用)	文字
IELTS	1 問複答	有 / 無	文字	文字・記述

旧日能試の出題基準 (1994) には、聴解テストの特色の 1 つに「述べられていることを全て聞き取るのではなく、課題の解決にとって最も効果的な聞き方をすることが求められる。具体的には、課題の解決にとって重要な部分のみに注目して聞き、それ以外の情報については聞き飛ばすといったストラテジーの適切な使用が必要とされる」(p.188) とある。この場合、本文の前にタスクがなければならない。タ

¹⁷ N5・N4・N3 には「発話表現」があり、視覚情報として絵・図が与えられ、その場面状況に適切な発話を選択するという形式があるが、初級～初中級レベルであり、本研究のアカデミック・リスニングには該当しないと判断し、表から外した。

¹⁸ 新日能試の聴解テスト・アイテムの具体例は 2 章 2.4.1 (pp.21-27) を参照。

スクを含む設問と選択肢が音声表示のみである旧日能試の聴解テスト・アイテムは、聞くスキルのみを測定するという考えの基に作成された。日本語テストとして代表的な旧日能試が、プリタスクがあり、聴解スキル以外を排除した問題形式だったことで、その後開発されたジェトロビジネス日本語能力試験や日留試が同様の問題形式となった可能性は高い。赤木（2003）は、状況がわからないままの聴解は実際場面では特殊であるとしながらも、聞くべきことを事前に知って聞く場合とそうでない場合が現実には存在するため、プリタスクに関しては次のように述べている。聴解テストでは、聞く目的を事前に知っている場合は必要情報の取捨選択となる。そうでない場合は全体の内容把握と、それを記憶しておいてポスト・クエスチョンに答えるという、より高度な音声処理能力が必要である（赤木 2003, pp.53-54）。

Sherman (1997)は、本文を2回聞かせる聴解テストで ①質問を本文の前にする ②質問を本文の後にする ③質問を1回目と2回目の間にする ④質問無しで内容自由記述をさせる の4つのタイプで結果を調べた。質問を本文の間にはさんだ③のタイプが1番得点が高く、①と②は同程度、④が1番低かった。受験者へのアンケートでは、③が1番答えやすく、次いで①、②、④と続いた。この結果から、聴解テストでは、質問の位置が本文の前後どちらであっても得点には影響しないことがわかり、受験者には、本文の前に質問があるほうが心理的に正の影響を与えていることがわかった。

宮本（2004）は、旧日能試の聴解テスト・アイテムを使用して、本文を聞く前にプリタスクがある場合と、後にポスト・クエスチョンがある場合の違いを実験した。その結果、全体として正答率および識別力に有意差のある違いが見られなかったが（同掲 p.52）、アイテムのタイプにより違いが認められた。絵のあるアイテムと内容把握や推測などのアイテムに関しては違いが見られなかったが、情報の取捨選択のアイテムでは、設問がないほうが識別力が高かった。これは、情報の取捨選択の場合、設問に対する必要情報の部分理解だけで解答できてしまい、全体の内容が理解できる言語能力が高い受験者や、談話の全体理解から何が問われるだろうかと判断する能力を持った受験者の識別ができなかったことを示している。実験後のインタビューでは、受験者の多くがプリタスクがないアイテムは自信が持てないと答えている（同掲 p.51）。Sherman (1997)の結果と同様、プリタスクは受験者の心理的側面に正の影響を与えているが、識別力には負の影響を与えていると言ってよいだろう。

足立（2009a）は日留試の聴解テスト・アイテムを使用し、①設問を音声で聞く（通常の日留試テスト・アイテム） ②設問（状況説明+プリタスク）はないが、場面・状況説明が絵や文字（視覚情報）で提示されている ③設問がない¹⁹ の3つの問題形式による聞き方の違いを研究した。その結果、①はプリタスクの答えを探しながら聞かすが、②と③では、聞き手は何が大切な情報かを考えながら聞いていることがわかった。プリタスクがある場合、心理的に正の効果がある反面、聞き方に関しては設問の解答のみに集中し全体を聞かない、テストテクニックに頼ってしまうという負の効果²⁰も見られた。②と③ではほぼ差がなく、どのような状況下での談話であるかは本文からわかるため②で提示した絵や文字情報の必要性がさほど感じられないというインタビュー結果であった。受験者にとって、プリタスクの有無が重要であり、状況場面説明は必要度が低いことがわかった。このときの研究協力者は日本語の上級レベルの留学生であり、聴解テストに慣れていたため、聴解テストの談話内容から場面状況が推測できたと考えられ、今回の結果から全ての聴解テストに状況場面説明が不必要であると結論づけることはできない。

¹⁹ 「設問がない」テスト・アイテムは本文の前にもなく、本文終了後にポスト・クエスチョン（タスク）が音声で与えられる。

²⁰ 最後だけ理解できれば解答が得られる、4選択肢の場合最初の選択肢に正答がくることは少ないなどの多肢選択法の聴解テストにのみ通用する知識を利用し、聴解能力以外の能力に頼りすぎることをさす。

これらの研究結果から、聴解テストにおけるプリタスクの有無、あるいは設問（状況説明＋プリタスク）の有無について、さらなる研究の必要性を感じた。本研究の6章と7章では、聴解テスト時における聴解行動を分析・考察するが、その際、設問の有無による聴解行動の違いという点にも注目し、6章では設問があるテスト・アイテム、7章では設問がないテスト・アイテムの場合の聴解行動について論じる。

最後に、選択肢の提示方法が音声のみの場合と、視覚情報の場合の違いについてまとめる。島田(2006)は、音声のみの選択肢と、選択肢を見ながら聴解テストをした場合を比較した結果、テスト・アイテムの難易度(困難度)は下がるが、識別度(識別力)に変化が見られないと報告している。赤木(2003, p.53)は、表3.2の結果を分析し、選択肢を音声のみで聞くテストより視覚提示をしているテストのほうが多い理由は、聴解に集中する時間的な限界や正答ではない選択肢を聞かせる非現実性が原因ではないかと述べている。聴解テストは受験者に記憶への負担をかけるため、記憶に頼らない選択肢の視覚提示により、テスト結果の信頼性があがる可能性がある。また、選択肢が視覚情報で与えられた場合、選択肢を聞く時間が削られ、テスト時間全体が短縮できる。

テスト時間全体の短縮の利点は、既に設定されているテスト時間内でより多くのテスト・アイテムを出題できることにある。テスト・アイテム数を多く出題することで、テスト結果の誤差(内容の文脈化の程度やまぐれ当たりによるものを含む)が軽減できる。現在、日留試の日本語テスト時間は2時間程度である。限られた時間内に、より多くのテスト・アイテムを出題することができるテスト形式の変更が、今後の課題と言えるだろう。

3.5 聴解の質的研究方法

聴解の質的研究では、プロトコル分析がよく使用される(O'Malley & Chamot 1990、古川他 1992、Wu 1998、Goh 2002、Vandergrift 2003、横山 2004, 2005a, 2005b)。プロトコル分析は対象者が自らを内観し、言語報告をした記録をデータとする(海保・原田 1993)。対象者が発話したすべてのもの、例えば、言い誤りやポーズや間投詞を、時間順に文字化したものがデータとなる。言語に表されていない部分をも、データとして丁寧に分析者が言語データに付加していくことで、対象者の認知過程の推測が可能となる(同掲 p.99, p.101)。よって、プロトコル分析は、対象者の認知過程の構築、認知処理の仕方などの分析に適している方法である(Ericsson & Simon 1984、海保・原田 1993)。プロトコルのデータ解析では、頻度分布図や相関図を描いたり、データのばらつきを計算する探索的データ解析と、条件の違いによる比較から検定を行う確定的データ解析がある(海保・原田 1993, p.111)。これらの方法を使用することで、聴解の困難度に影響を与える要因を調べたり、形態の違い(会話や雑音が入ったものなど)や形式の違い(複数選択問題、真偽問題など)などによる聴解への影響の研究などに用いられている(Green 1998, p.30)。しかし、プロトコルのデータは非常に個人的な内容が豊富である反面、個人差も大きい(Green 1998, p.21)。そのため、ケース・スタディのような研究に適しているが、本研究のように聴解テスト時の聴解行動全体の動きを、個人差を超えて捉えようという試みにはあまり適さない。対象者それぞれの行動の範疇や関連性を見ていく必要はあるが、個人差は全体の行動に統括されなければならない。

このような研究に適している研究方法として、修正版グラウンデッド・セオリー・アプローチ(Modified Grounded Theory Approach; 以下 M-GTA)が挙げられる。M-GTAとは、データに基づいた意味解釈を行い、概念の生成とそれらの関連性をまとめてある理論を創る方法であり(木下 2007, p.35)、それらによる統合的な説明図を創る(同掲 p.75)。

木下(2007, p.67)は、M-GTAに適した研究は「人間を対象に、ある“うごき”を説明する理論を生成す

る方法」であると述べている。佐藤（2008,p.108）は、M-GTAのパラダイムは以下の3つの構成要素からなるとしている。

1. 条件：現象の構造に関わるもの。なぜどこで、どうして、いつそのような現象が起きているのか、という問いに対する答え
2. 行為・相互行動：現象のプロセスに関わるもの。上記のような条件のもとで、誰（と誰）がどのような行動を、どのような形でとったのかという問いに対する答え
3. 帰結：個人の行動や複数の人々の間の相互行為が結果として、どのようなことをもたらしたのか、という問いに対する答え

本研究は、聴解テスト時に対象者がどのような聴解行動、つまり“うごき”をとったのかをまとめるものであり、上記のパラダイム 2.「聴解問題を聞いて解答を得るまでにどのような行為をとったのか」という問いに対する答えを得ようというものである。以上のことから、本研究では M-GTA を分析方法とした。聴解テスト時の聴解行動の分析方法については、6章「6.2 研究の枠組み」(p.118)にて詳述する。

第2言語の聴解過程におけるモデルの作成は少ないが(武井2002,p.66)、日本語教育では田中他(1986)の中間言語の聴解モデルと、横山(2005a,2005b)の対面理解における理解構築の流れ図がある。

田中他(1986)は、聴解理解は仮説検証をしながら行われると考えた。聴解過程では、情報がインプットされると「聴解の理解のための語彙的知識」「談話構造法」「百科事典的知識」をリファレンス辞書として活用する。外国人は、語彙的知識に関する辞書に載っていない語彙が、日本人と比べると非常に多い。外国人の場合、自らが持っている語彙的知識の辞書に載っていても、わかる言葉とわからない言葉の間にぼんやりとわかる言葉が存在する。このぼんやりとわかる言葉を、推理によってあいまいな語彙的知識の辞書に入れ、このあいまいな言葉を聞き返しによって検証するのである。田中他は外国人の聴解では、既知語彙の知識だけではなく、未知語彙の推測が重要であることを指摘した。

横山(2005a,pp.50-56)は、音声の聴取を記憶し、その後再生するまでの理解過程を図で示した。まず、問題がどのように認知されたかを分析し、その問題の対象を「単語・文・段落・テキスト」と「問題無し」に分類した。問題がある場合、「テキスト外の既有知識」と「テキスト内の文脈」をモニターして、問題解決の手がかりを求める。その後、問題を確認・検証・修正をするために、「質問・推測・放置・保留」という行動をとる。その結果、「未解決・誤解・理解達成」の1つに到達する。問題が未解決の場合は、そのまま記憶するか、問題解決の「保留」に貯蔵し、新たな解決法の過程をとる場合もある。その後、「誤再生・再生無し・正再生」のどれかを行うことになる。聴解の理解過程で問題解決をする際、様々な対象をモニターし、様々な方策を経るという聞き手の多様性が図から見てとれる。問題が未解決の場合は、その後の方向性を示し、対面理解における聴解過程は一過性でないことを明らかにした。

田中他と横山は対面会話における聴解過程研究であり、日本語テストの聴解モデル研究は筆者の知るかぎり皆無である。本研究では、聴解テストにおける聴解行動を分析し、モデル化することで、聴解時に使用する言語知識やストラテジー、聴解時に起こる問題、心理的な側面が、時間経過と共にその後の行動にどう影響を与えるかを可視化する。

4章 聴解テスト・アイテムの基礎的データ

この章では、本研究の1つ目の目的「日本留学試験聴解テスト・アイテムの特徴を明らかにする」。

4.1 基礎的データのための分析枠組み

3章「3.4 テストの入力ファセット」(p.45)で記述した Bachman (1990 邦訳 p.392)の「テスト方法のファセット」の「入力」「期待される応答」「入力と応答の関係」を見るために、日留試の聴解テスト・アイテムの構成について述べる(図4.1)。赤木(2003, p.49)は聴解テスト・アイテムの構成を、場面や状況を表す部分「状況説明」、課題の部分「設問」、聴解スクリプト「本文」、解答部分「選択肢」と呼称している。本研究では、日留試のテスト形式から、「設問」が本文前と本文後に表れるため、本文前の課題を「②プリタスク(pre-task)」¹、本文後の同じ質問を「⑤ポスト・クエスチョン(post question)」とした。本文前の場面や状況を表す「①場面・状況説明」と課題「②プリタスク」、①と②をあわせた③を「設問」とし²、談話部分は「④本文」、「⑥選択肢」とした。

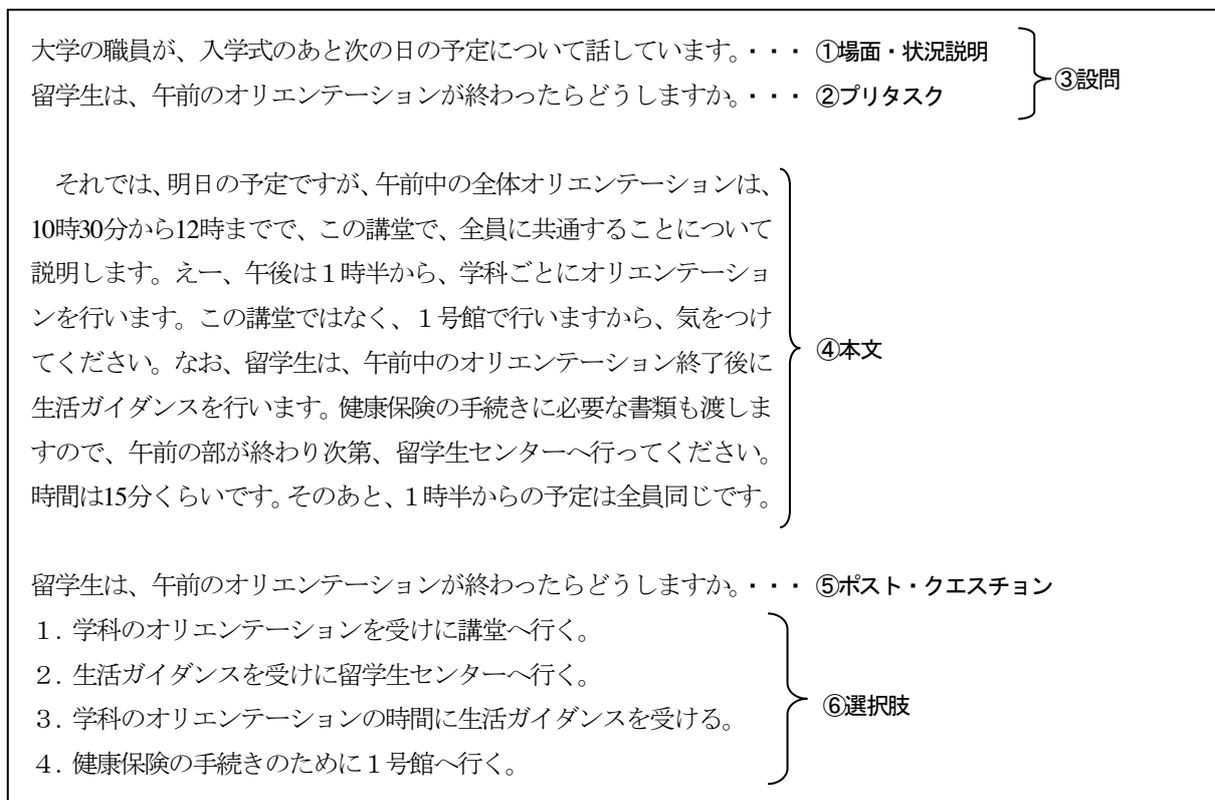


図4.1 日留試聴解テスト・アイテムの構成 (2006年度(平成18年度)2回目 聴解テスト1番)

次に、テスト・アイテム内容を分析するための項目を決める。菅井(2003, pp.40-43)はアイテム・ライティング時に注意すべき点として場面、登場人物、トピック、難易度(困難度)、倫理観、真正性

¹ プリタスクをプリクエスチョン(pre-question)という言い方もあるが、本文を聞く前に与えられた課題と考えプリタスクとした。

² 6章と7章で設問あり・なしの聴解行動を比較するため、①と②をあわせた名称が必要だった。

を挙げている。「登場人物」は「場面」と連動している。例えば、「講義」場面であれば、教師のみあるいは教師と学生、「サークル活動」場面であれば学生同士あるいは先輩後輩であるため、「登場人物」を「場面」の下位場面とした。本研究は、聴解テスト・アイテムのスク립ト分析であり、「困難度」には注目していないため除く。「倫理観」と「真正性」については、倫理観に反するテスト・アイテム、真正性を大幅に欠いたテスト・アイテムは出題されてはいないと考え、分析項目からは除く³。よって、「場面」と「トピック」を分析項目とした。

西川・小牧(2002)は、コミュニケーションを分類する際にコミュニケーションの過程の何に注目するかで分類できると述べている。その例として、コミュニケーションの目的、人数、動機的側面、メッセージ、チャネルを挙げている。人数に注目した場合は、「個人的コミュニケーション：1人です(頭の中で)」、「対人コミュニケーション：2人です」、「グループ・コミュニケーション：多人数です」、「パブリック・コミュニケーション：多人数に対して相互作用がない」の4つに分類される。動機的側面に注目した場合は、コミュニケーションを目標達成の手段とする「道具的コミュニケーション」と、コミュニケーションをすること自体が目的の「欲求充足的コミュニケーション⁴」の2つに分類される。コミュニケーション過程のメッセージに注目した場合は、「バーバル・コミュニケーション」と「ノンバーバル・コミュニケーション」である。チャネルに注目した場合は、一方通行の「一方的コミュニケーション」と双方向でできる「相互的コミュニケーション」に分類できる。(西川・小牧2002, pp.4-7)

「人数・動機的側面・メッセージ・チャネル」の4点で分析することを検討した結果、「人数」のみの分類とした。その理由は、本研究では聴解テスト・アイテムを使用するため「メッセージ」はバーバルのみであること、プリタスクがあるテスト・アイテムであるため「動機的側面」では「道具的コミュニケーション」の談話以外はなく、「チャネル」は「人数」に包括されると考えたからである。日留聴解テストにプリタスクがあることから、「理解の種類」「情報を得る理由」「コミュニケーションの目的」を分析項目に加える。さらに、本文談話のコミュニケーションに関わる項目として、「談話形態」「談話内の機能」を分類項目としてとりあげることにした。以上のことから、8項目で分析することにし、これを分析のための作業枠組みとした。

- ①理解の種類 ②場面 ③形態 ④トピック ⑤人数 ⑥コミュニケーションの目的 ⑦機能
⑧情報を得る理由

以下、個々の項目について詳述する。

①理解の種類

理解の種類は、「全体理解」と「部分理解」の2つに分ける。全体理解は、本文の談話を理解する際、談話全体の内容を把握するように聞く、あるいは談話の概要を聞く聞き方をいう。部分理解は、全体ではなく、焦点を絞って聞く聞き方である。部分理解には、時間、場所、目的、理由などを聞く目的があり、理解の種類は聴解テスト・アイテムのプリタスクが影響すると考えられる。

³ 「倫理観」「真正性」は分析項目とはしないが、アイテム分析の際に、問題があると考えた場合は記述する。

⁴ 「欲求充足的コミュニケーション」とは、人と話すことによって自分の欲求不満を解消する、ストレスを発散する、緊張をほぐすなどである。

⁵ 本研究では、「全体理解」と「部分理解」を「設問」の「プリタスク」から判断した。今回は、「理解の種類」として2分類のみの分析を行ったが、今後さらに詳しくどのような「設問」であるかを分析していく予定である。

②場面

テストを開発する際、受験者の目標言語使用領域 (target language use) を考えなければならない (Alderson, Clapham & Wall 1995, p.13、Bachman & Palmer 1996 邦訳 pp.50-51)。目標言語使用領域は、「テスト受験者がテストそのもの以外で出会う可能性が高く、言語能力についての推測を一般化した、一連の特定の言語使用課題」である (Bachman & Palmer 1996 邦訳 p.51)。日留試の目標言語使用領域は1章の図 1.1「日留試の測定対象能力」(p.5) から判断すると、授業、大学内の授業以外の場面、留学生の大学外での日常生活の3つが実生活領域 (real-life domain) と考えられる。また、日留試は知識ではなくスキルを重視し、スキルは生活スキルと学習スキルの2つに分けられている。生活スキルには「言語運用能力」と「社会適応能力」があり、大学外の日常生活と大学内での対人関係や異文化調整能力などが挙げられている。学習スキルは、「事務手続き処理能力」「学習活動スキル」「研究活動スキル」の3つに分けられている。この分類の仕方では、学習スキルに「言語運用能力」および「社会適応能力」が必要ないかのように

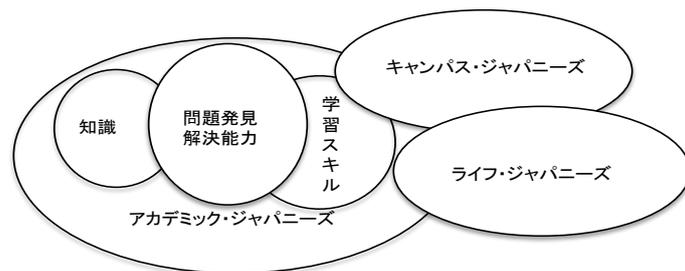


図 4.2 留学生のコミュニケーション能力 (堀井 2006, p. 116)

見えてしまう。しかし、学習スキルにおいてもこの2つの能力は必要である。この問題点を堀井(2006)も指摘している。堀井(2006, p.116)は日留試で用いた「スキル」は「能力」のことでありと考へ、「アカデミック・ジャパニーズ」は高等教育機関で勉学に必要な日本語能力」と定義し、アカデミック・ジャパニーズは「基礎知識、問題発見解決能力、学習スキル」の3つから成り立っていると考へた。堀井は、日留試が測定しようとしている留学生のコミュニケーション能力を明確にするために、「アカデミック・ジャパニーズ」「キャンパス・ジャパニーズ」「ライフ・ジャパニーズ」の3つに分けた(図 4.2)。

「アカデミック」という用語の解釈については、3章「3.1 アカデミック・リスニング」(p.36)で述べた通り、研究者によって多様である。本研究では堀井が分類した「アカデミック・ジャパニーズ」「キャンパス・ジャパニーズ」「ライフ・ジャパニーズ」が使用される場を3つの場面とし、「アカデミック場面」「キャンパス場面」「ライフ場面」と分類することにした。「アカデミック場面」は、狭義の「講義を聞く」だけではなく、大学の授業の形式に従いいくつかの場面を含むものにした。例えば、学生が講義を聞くという一方向のみの講義、教師と学生の相互作用が高い参加型授業、学生が主体となって行うゼミなどがある。本研究では「講義」「参加型授業」「ゼミ」などを含め、大学の教室内の学習場面をアカデミック場面とした。場所が教室外であっても、教師に学術的な内容の質問や相談をする、あるいは学生同士で相談する、議論するなど学術的活動であり、すべてアカデミックな場面とした。さらに、授業中、教師から授業の進め方、試験の方法などの説明も授業という枠組みとしてアカデミック場面とした⁶。キャンパス場面は、大学の施設使用や事務的手続きなど大学生活に必要な談話、サークル活動や大学の行事などについての談話場面である。アカデミック場面とキャンパス場

⁶ 3章 (p.38) で本研究での「アカデミック・リスニング」を場面を問わず「学術的・学問時」に行なうリスニングの意味で使用し、と定義した。「アカデミック場面」も同様に「学術的・学問時」の場面とした。

面の違いは談話内容や目的による。例えば、教師が授業内でレポートについて説明している場合は「学習に直接必要な情報収集」であるためアカデミック場面だが、学生が授業後教師にレポートの日程の「再確認」をしている場合はキャンパス場面とする。ゼミで発表の方法や内容を学生同士が相談している場合は「学術的内容」のためアカデミック場面だが、ゼミ発表のために調査に行く集合時間の相談はキャンパス場面である。ライフ場面は友人とおしゃべりや衣食住などの生活上で必要な活動、娯楽などを含めたものとし、談話をしている場所が大学内と考えられても、学術・研究、学習に必要なではないと判断したものを含めた。

③形態

談話形態は、社会言語学的知識と関わっている。公的な場面と私的な場面では、使用語彙や言葉遣いなどが変わる。Savignon (1997 邦訳 p.18)は、コミュニケーション能力は文脈で特定化された能力であるとし、その文脈を理解し適切な言語使用ができるためには適切な言語使用域 (register) と文体 (style) を選択しなければならないと述べている。本研究では、留学生の生活に関連する言語使用域として、①大学の授業 ②大学の事務的関連 ③私的関係 (友人・家族) ④社会的関係 (目上・初対面、職場の上下関係) ⑤大学以外の公的関連 (職場・公共の場・メディア) の5つを考えた。文体は普通体・丁寧体 (敬語を除くです・ます体)・敬体 (です・ます体を除いた敬語) の3つである。

④トピック

トピックは目標言語使用領域および真正性に関わる。菅井 (2003, p.41) は言語特徴に注目してしまうと真正性が低いものになりがちであると注意を促している。例えば、試験範囲の説明は留学生が実生活で体験する聴解として真正性が高いが、選挙の地域集会のお知らせを聞くことは、演説という言語特徴はあるが真正性が低い。本研究ではトピックを各テスト・アイテムのタイトルと位置づけ、本文の内容がある程度推測できるための項目とした。トピックは8項目の基礎的データの中で1番自由度が高く、多岐にわたるためトピックによる分析は行わなかった。

⑤人数

人数は「個人的コミュニケーション：1人です (頭の中で)」、「対人コミュニケーション：2人です」、「グループ・コミュニケーション：多人数です」、「パブリック・コミュニケーション：多人数に対して相互作用がない」の4つで分類する。

⑥コミュニケーションの目的

西川・小牧 (2002, pp.4-5) は、コミュニケーションの目的の一般的理由として、様々な文脈に共通するものは「情報を提供する」「影響する」「社会化する」の3つであると述べている。「情報を提供する」とは、コミュニケーションの根幹であり、教えたり知らせたりすることである。これによって知識を得たり理解を深めることができる。「影響する」とは、相手の行動や態度を変容させることである。「社会化する」とは、相互に作用することで、人間関係を作っていく。この3つの目的は通常単独ではなされず、複合的に行なわれる。本研究では以上のことからコミュニケーションの目的を「目的1：情報」「目的2：影響」「目的3：社会化」の3つとした。アイテムによって単一目的、あるいは複数目的となることが予想される。

⑦機能

相手にどのように伝えるかという点では、コミュニケーション形態だけではなく、どのような意図を持って伝えるかも必要である。表現意図の分類は、日本語教育学会コース・デザイン研究委員会(1991, pp.157-158)を参考としたが、以下の5つに大別している。

1. 談話の構造に作用を及ぼす機能（談話を始める・発展させるなどの機能、定形の慣用的挨拶など）
2. 情報を授受する機能（情報を提供する・求める・確認するなどの機能）
3. 聞き手に働きかける機能（注意を引く、誘う、提案するなど）
4. 相手の働きかけに応える機能（賛成・反対する、許可を与えるなど）
5. 自己の感情を表す機能（感謝する、同情するなど）

この5つの機能のなかで、「1.談話の構造に作用を及ぼす機能」の「談話を始める・発展させるなどの機能」に関しては、聴解テストの談話であることから、談話は自動的に始まり、発展させるには談話の長さが足りない。また、「定形の慣用的挨拶」はテスト・アイテムの自然な談話構成に必要なものではあるが補助的な機能のみを担うと判断し、1.はコミュニケーション機能の分類から外し、2.から5.までの4つとする。

⑧情報を得る理由⁷

通常、我々が何かを聞く場合には理由がある。聴解テストの場合は、プリタスクによって聞く理由が明らかにされる。本研究では、プリタスクおよび本文の情報を知る理由は何かという視点で分類することにした。例えば、講義を聞いている場面でプリタスクが「調査で注意すべき点は何だと言っていますか」の場合、情報を知る理由は「学術・研究のため」である。また、次の授業内容について話している場面でプリタスクが「次回の授業で何をしていますか」の場合、情報を知る理由は「学習のため」である。本研究ではアカデミック・リスニングをより詳細に見るため「学術・研究のため」「学習のため」の2つにわけ、次にキャンパス場面で多く見られるであろう「大学生活のため」、そしてライフ場面の「生活のため」、また、啓蒙的な内容や趣味、社会的な人間関係などがあると考え、「自分のため」「人間関係のため」を加え6つで分類した。

分析対象は、日留試が始まった2002年度から2013年度までの11年間、年2回の実施であるので、計24回の公開された聴解テスト・アイテム440である。図4.3(次頁)は、実際の聴解テスト・アイテムの分析例である。分析方法は、各聴解テスト・アイテムの設問と本文を「4.1 基礎的データのための分析枠組み」で挙げた8項目で分類し(分析例)、その後各テスト・アイテムの特徴を立体的に示すアイテム・バンクのモデルを提示する。

⁷ 「情報を得る理由」は課題として与えられた聞く目的のことだが、プリタスクが特定の聞く目的であり、ここではそれと区別するために「理由」とした。

- ①理解の種類：部分理解
- ②場面：キャンパス（下位場面：学生課関係）
- ③形態：説明
- ④トピック：入学ガイダンス
- ⑤人数：パブリック
- ⑥コミュニケーションの目的：目的1情報-行動、目的2影響-行動変容
- ⑦機能：情報を授受する機能-説明する、聞き手に働きかける機能-命令する
- ⑧情報を得る理由：大学生活のため

大学の職員が、入学式のあと次の日の予定について話しています。留学生は、①午前のオリエンテーションが終わったらどうしますか。

それでは、明日の予定ですが、午前中の全体オリエンテーションは、10時30分から12時までで、この講堂で、全員に共通することについて③⑦説明します。えー、午後は1時半から、学科ごとにオリエンテーションを行います。この講堂ではなく、1号館で行いますから、⑦気をつけてください。なお、留学生は、午前中のオリエンテーション終了後に生活ガイダンスを行います。健康保険の手続きに必要な書類も渡しますので、午前の部が終わり次第、⑥⑦留学生センターへ行ってください。時間は15分くらいです。そのあと、1時半からの予定は全員同じです。

図 4.3 日留試験聴解テスト・アイテムの分析例（2006 年度（平成 18 年度）2 回目 聴解テスト 1 番）

* は分析する際に注目した部分で、前の番号は分析項目①～⑧に対応している

4.2 結果と考察

4.2.1 理解の種類

理解の種類は、プリタスクにより決定される。全体理解と部分理解の2分類の予定だったが、実際のテスト・アイテムのプリタスクには4つのタイプが見られた（表 4.1⁸）。

では、日留試験のプリタスクのタイプを具体的に見ていく。プリタスクが「いつ」「どこで」「何を」などを問う例1のようなテスト・アイテムの場合は「部分理解」となる。

- 例1 ・留学生は、午前のオリエンテーションが終わったらどうしますか。
 ・次回の授業で何をするとおっしゃいますか。

「複数部分理解」はプリタスクで「いつどこで」のように複数問う場合と、「最も～」「1番～」のように比較する部分も聞かなければ判断できない場合である。

- 例2 ・この先生は、ティッシュペーパーとトイレットペーパーの違いは何だと言っていますか。
 ・このトラックを使った場合、今までの方法に比べてよくなるのはどんな点ですか。

表 4.1 理解の種類（2002-2013）

	アイテム数	%
部分理解	318	72
複数部分理解	61	14
全体理解	36	8
部分・全体理解	25	6
計	440	100

⁸ 表 4.1～表 4.24 の「%」の数値は、四捨五入で表示しているため、計が 100%にならない場合もある。

例2の最初のプリタスクは「ティッシュペーパー」と「トイレットペーパー」の2種類について理解しなければならぬため複数部分理解となる。例2の2番目のプリタスクは「今までの方法」と「新しい方法」の2種類の理解が必要なため複数部分理解となる。

表4.1(前頁)からわかるように、72%が「部分理解」、14%が「複数部分理解」であり、両方あわせて86%がプリタスクを正確に理解できれば、その情報がある部分のみの理解で解答が得られるアイテムであることがわかった。ただし、解答を得るための情報選別をする際には、理解を助けるための情報も必要である。また、結束性は前後の文脈を理解していないとわからない場合もあるため、純粋に部分理解のみでよいという意味ではないが、プリタスクに関連する部分を断片的に捉えただけで解答が得られる可能性はある。このようにプリタスクがあるために聞き方を限定してしまうことの不自然さ、あるいはテスト・テクニックの影響については、検討すべきである。

「全体理解」は、プリタスクが「言いたいことは何か・話題は何か」などで、全体を聞いて判断しなければならぬものである。36アイテム中アカデミック場面が28場面であり、そのほとんどが講義であったことは予想に反しない。このタイプのアイテムは2008年から増加傾向にあるのだが、回により数にばらつきがある(例:2010年1回目5アイテム、2010年2回目2アイテム)。2012年は1回目には出題がなく、2回目も1アイテムのみであり、2013年も1回目には出題が見られなかったが、2回目には2アイテムが出題されている。このことから、日留試では、談話の概要を表出させる能力に関しては必要ではあるが、重要視していないと推測される。アイテムの長さが2分以下程度であることから、「全体理解」を問うほどの内容を含む談話作成が難しいとも考えられる。

「部分・全体理解」は、プリタスクに「これからどうしますか・まず何をしますか」などがあり、本文を聞かないと「これから」「まず」になるのかがわからないものである。得るべき情報は「何を」「どうする」という部分理解であるが、プリタスクからは起点がわからないため、どこに情報があるのか予測できず全体を聞くことになる。このタイプのアイテムは2002年から2005年まで毎回2アイテムから5アイテム程度出題されていたが、その後は1アイテム程度に減少した。数にばらつきはあるが、定期的に出題されていると言ってよいだろう。しかし、このタイプはプリタスクを聞いた時点で「これから」がいつを指すかわからないため、課題としては完結していないと言える。なお、2011年以降は出題されていないため、聴解テストにおいて、時間の起点が曖昧な談話に問題点が見られたのではないかと考えられる。上記の時間的な起点ではなく、「部分・全体理解」と分類したアイテムが2013年2回目に1つ出題されている。設問では「～興味深いと言っていることは何ですか。」という部分理解を問うアイテムなのだが、問われている「興味深いこと」が談話の最後にならないとわからない。このため、どこに情報があるのか予測できず全体を聞くことになる、という「部分・全体理解」と同様の行動が必要と考え、ここに分類した。談話展開の中で「〇〇に興味を持っているらしい」と情報理解したのち、「興味深い」というキーワードを発するのを待つことになり、単純な部分理解に比べ情報処理過程に負荷がかかる。この負荷は、アカデミック・リスニングの特徴の1つとして捉えられるのではないだろうか。今後、更なる出題が望まれる。

以上のように、日留試の聴解テストの場合、プリタスクで問われている部分の情報を聞き取れるという能力が重視されていることがわかった。嶋田(2005,p.81)は、日留試は課題達成のテストであると述べており、聴解テストでは課題(プリタスク)の理解が重要であるとした。ポスト・クエスチョンが本文のあとに再度配置されているのは、テスト・アイテムの目的としてプリタスクの解答を得ることが重要視されているからであろう。松沢(2002, pp.78-79)は、第二言語学習におけるタスクの定義を3つに分けている。1つ目は、目標外国語の教室内でのすべての活動をタスクとする。2つ目は、

学習者が目標外国語を使用して実世界で行う可能性のある行動・活動である。3つ目は、学習者が目標外国語を使用して実際に行動・活動して目標を達成するものである。3つ目の定義は、言語学習の教室内外は問わず、実際に目標言語を使用してある目標を達成すればそれもタスクとするという立場で、2つ目の定義を含んでいる。日留試の聴解テストにおけるプリタスクは、3つ目のタスクの定義である。テスト・アイテムは2つ目のタスクの定義で作成されたものであるが、テスト受験者は与えられたタスクに従って目標外国語を使用してそのタスクを達成する行動をとる（解答する）ので、3つ目のタスクの定義と重なる。伊東（2008a, p.147）も、プリタスクによって課題を達成する聞き方をすると述べている。プリタスクが部分理解を要求していることから、日留試は課題達成が目的のテストであり、問題発見解決能力を測定する目的はないと考えられる。

4.2.2 場面

2002年度から2013年度の聴解テスト・アイテム全体で「アカデミック・キャンパス・ライフ場面」の割合を見てみる（表4.2、図4.4次頁）と、アカデミック、ライフ、キャンパスの順に多い。場面の割合を経年で見てみると（図4.5次頁）、初年度の2002年はアカデミック場面は少なかったが、徐々に増加していることがわかる。それに伴ってキャンパス場面とライフ場面は減少しているが、2010年から2013年までのアカデミック場面は60%～75%とあまり変化が見られない。2010年に日留試の改定が行なわれており、改定以降アカデミック場面が占める割合はほぼ安定してきたと言ってよい。この変化は、日留試の重視する日本語能力が、日常生活ではなく大学における勉学に関わる場面に移行してきたと捉えられる。留学生の1日の生活に占める授業の割合は非常に高く、この移行は現実に即していると言える。

門倉（2003a, p.125, 2003d p.212）は公開用試験問題⁹を分析し、講義や教師から学問的指導を受けるような狭義のアカデミック・ジャパニーズから見ると、日留試のテスト・アイテムは掲示を読んだり事務的手続きをするなどのキャンパス・ジャパニーズが多すぎると苦言している。堀井（2003, pp.118-119）は、2002年度2回目の日留試の聴解テストとTOEFL（Model Test4）の聴解テストを比較した結果、授業に関するアイテムが日留試は23%でありTOEFLが54%であることから、今後は授業関係のトピックを増やすように提言している。門倉は公開用試験問題を、堀井は2002年度の問題を分析しているため、キャンパス場面が多いことの指摘は初期の日留試に対してのものである。図4.5からわかるように、聴解テスト・アイテムに占めるアカデミック場面は増加し続け¹⁰、大学留学を希望する受験者の能力測定の場面設定としての妥当性が確保されたとみてよいだろう。

⁹ 日本国際教育協会は2002年に日留試が実施される前、どのようなテストであるかを『日本留学試験公開用問題』という冊子で公開した。

¹⁰ 日留試の聴解テスト・アイテムでアカデミック場面が増加した理由は、推測の域を出ない。「4.4 改訂前2002年版と2010年改訂版の比較・考察」（p.76）で2010年改訂前と2010年改訂後を比較しながら考察を行っている。

表 4.2 場面の割合
(2002-2013)

場面	アイテム数	%
アカデミック	197	45
キャンパス	102	23
ライフ	141	32
計	440	100

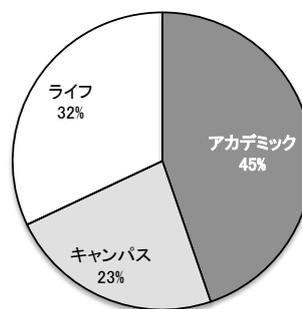


図 4.4 場面の割合 (2002-2013)

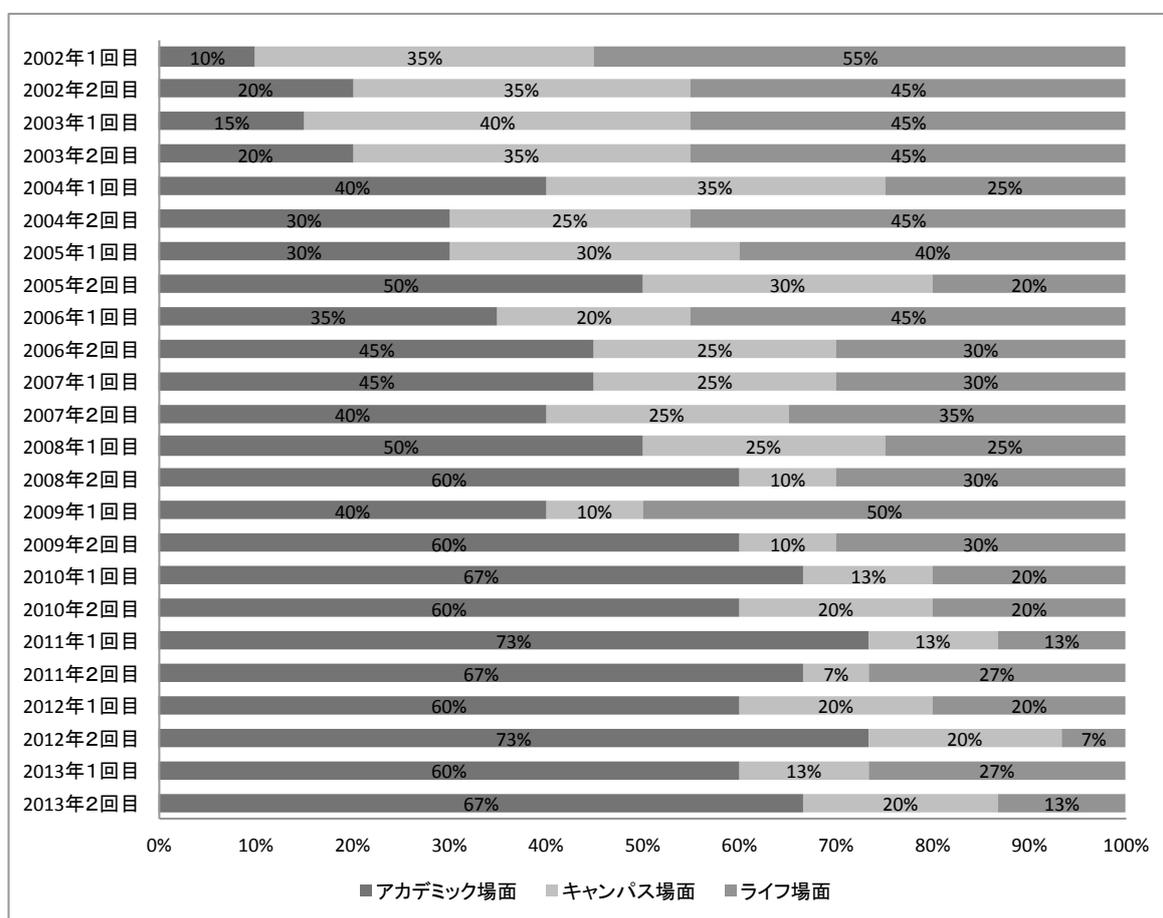


図 4.5 場面の割合の推移 (2002-2013)

表 4.3 アカデミック・ジャパニーズ・スキル・シラバス (堀井 2005, p. 28)

アカデミック・ジャパニーズ・スキル・シラバス			
アカデミック・ジャパニーズ	講義	授業前	予習：(課題) 図書を読む (スキミング・スキミング速読)
		参加	聴く (聴解)
			ノートテイク
			板書読み・資料読み (聴読解)
			質問に答える、質問をする
		授業後	復習：わからなかった点を調べる
	課題 (宿題) 遂行		
	ゼミ	授業前	情報収集 (図書館・コンピュータ利用)
			レジュメ・スクリプトづくり
		参加	他の発表を聞く：質問力、コメント力
			発表 (プレゼン) をする：質問、ディスカッションを仕切る
	授業後	レポートにまとめる	
	実験		実験をする
試験	試験前	情報収集 (試験時間、場所、条件、内容など)、まとめ学習	
	試験後	評価チェック	
レポート	提出前	情報収集 (テーマ、字数、用紙、提出方法などの形式など)	
	提出	アカデミック・ライティング	
キャンパス・ジャパニーズ	教務課関係		履修要項の理解・登録など
	学生課関係		掲示板 (休講、諸連絡、呼出、行事) 読み取り 諸手続き (長期欠席、休学など) 遂行
	教員と	面接など	アポ取り、質問、伝達、お礼など
	職員と		問い合わせなど
	学生と		授業についてのやりとりなど
	サークル活動		先輩後輩関係語
ライフ・ジャパニーズ	友人などと		ため語
	アルバイト先		敬語・業務遂行
	地域社会		公共施設 (役所、駅、病院など) 利用、買い物、近所づきあいなど

アカデミック・キャンパス・ライフ場面をさらに詳細にみていく。堀井 (2005, pp.25-28) は日留試の日本語シラバスをわかりやすいシラバスで提示できないかと考え「アカデミック・ジャパニーズ・スキル・シラバス」を作成した (表 4.3)。この表を参考に3場面の下位場面を分類し、この表の項目では分類できなかったものは、項目を追加した。

授業形態としては聞くことが中心の「講義」と、教師と学生の相互作用が高い「授業」があるので「授業¹¹」を追加した。表4.4からわかるように、アカデミック場面のほとんどを「講義」場面が占める。学生が大学で過ごす時間の大半は講義および授業であること、また様々な専門分野の留学希望者が受験することを考えると、なるべく多くの分野から出題することで公平性が高くなる。そのために1テストの聴解テスト・アイテムに多くの「講義」場面があることは納得できる。しかし、「講義」を聞く場合、日留試のように1～2分程度の短く切り取った談話でよいのかという疑問が生じる。実際に講義を聞く際、学生は教師が話している内容から重要なポイントは何か、記憶すべきか、記録すべきか、聞き流してよいのかなどを判断しながら聞いている。そして、講義が終わった時点で今日の講義の内容の概要を得る。聴解テストにおいて、このような聴解行動を行なうためには1～2分の談話では短い。一問一答の制約とテスト実施時間が限られているため短い談話の出題となっていることは理解できるが、真正性は低い。聴解テスト・アイテムの一律性について、門倉（2003a,p.124、2003d,p.211）は日留試の一問一答の形式が問題（タスク）の内容をワンパターンにしていると批判している。因（2005,p.271）は、日留試の聴読解分析においてであるが、同様に一問一答のため多様性はないと述べている。そのため、多くの問題が常識や合理的な推論ではなく、聴解の正確さを問うものばかりで総合的な判断能力を測定するテスト・アイテムは少ないと苦言を呈した。堀井（2003,p.119）も、困難度が低い原因の1つに談話の長さを挙げ、複数設問をつけることを提案している。

キャンパス場面の下位場面は表4.5である。「学生と、サークル活動、課外活動」は授業以外の場であり、学生が勉学だけではなく様々な活動に参加していることを留学生に知らせる波及効果がある。しかし、職員との談話や学生課、教務課に関する場面が少なく、日本の大学の施設やシステムなどを伝えるには不十分である。日本語は、外国語学習の対象としてはマイナーな言語である。日本の高等教育機関についてもあまり知られていないだろう。日留試のように、海外で年に2回実施されるテストによる波及効果は大きい。大学のキャンパス内が見えるようなテスト・アイテムを出題する意義は大きいと言える。テストにおける各場面の割合の推移から（図4.5 p.60）、アカデミック場面増加に伴い、キャンパス場面は激減していることが見てとれる。日本の高等教育機関のシステムなどの知識が十分に広まったという

表 4.4 アカデミック場面の下位場面
(2002-2013)

アカデミック場面	アイテム数	%
講義	157	80
ゼミ	15	8
授業	13	7
レポート・論文	7	4
実験・調査	4	2
試験	1	1
計	197	100

表 4.5 キャンパス場面の下位場面
(2002-2013)

キャンパス場面	アイテム数	%
学生と	56	55
教員と	16	16
サークル活動	9	9
職員と	7	7
課外活動	6	6
学生課関係	6	6
教務課関係	2	2
計	102	100

¹¹ この「授業」はp.51で定義した「参加型授業」のことである。

判断からだろうか。あるいは、日本国内受験者と海外受験者に日本の高等教育機関のシステムおよびキャンパス内の知識の有無に差があることを考慮したためだろうか。後者に関しては、キャンパス内の知識を問うのではなく、単なる場面の提示であれば問題はない。1回のテスト・アイテム数は決められているため、アカデミック場面が増加したことからキャンパス、ライフ場面が減少するのは当然であるが、留学生にとってキャンパス場面は「日本」の大学の疑似体験であり、ライフ場面より重要であるとする。今後も各場面の推移には注目したい。

最後に、ライフ場面を分析・考察する。表4.6をみると「友人などと」が52%を占めている。ライフ場面における友人との会話理解は、日留試の測定対象能力(p.31 図1.1)の「対人関係構築・維持・修復能力」としてあげられている。今回、分類したアイテム「友人との会話」では、「対人関係の維持」(26アイテム)はあるが「構築・修復」はなかった。実際に留学生が抱えている「対人関係」の問題点は文化・社会的な摩擦や誤解などが原因のものもあるだろう。そのような状況の談話は、留学生には役に立つのではないだろうか。

次に「メディア」が20%だが、今はインターネットで情報を得ることが多く、テレビもニュースや情報番組にはテロップが多く、文字情報を見るだけで十分理解できるため、今後、メディア聴解の重要度は低くなると予想される。「講演・意見」は最も留学生が出会う可能性が低い場面である。実際、分類された聴解テスト・アイテムの内容には会社経営や旅行などについてで留学生に絶対に必要というものではなかった。「地域社会」では留学生会館の場であったり留学生が地域住民から何かを頼まれるなど、留学生と関連づけているテスト・アイテムもあるが、このような状況に接する可能性は低いと考える。留学生会館など施設使用についてや交流会などのトピックがあるが、大学キャンパス内の施設利用や新入生歓迎会でも置き換えられる。他は医療、交通、買い物などである。これらは上記したように初級レベルを習得していれば、たとえ聴解内容が100%理解できなくとも他の手段で解決できる。「仕事」のアイテムは、場面・状況説明が「男性が女性に～」で始まり、談話内容も社内の会話のようであったため「仕事」に分類した。場面・状況説明で「男子学生が女子学生に～」と設定すれば、キャンパス場面に分類される。このようにライフ場面は登場人物を学生、大学関係者にすることで、キャンパス場面に変更ができるアイテムが多い。では、留学生にとって、ライフ場面で頻度が高かつ内容理解が必須である場面は何であろうか。菅井(2003,p.40)は大学入学の予備教育の場面設定として、例えばアルバイトを設定するのは間違いだと述べている。実際に留学生がアルバイトをする可能性は高いが、留学はあくまでも学業をするためであり、アルバイトを推奨するような内容はふさわしくない。留学生は移住し仕事をする外国人とは違い、ライフ場面もかなり特殊ではないだろうか。例えば、アパートにいる時間に比べ圧倒的に大学内にいることが多いことなど、ライフ場面も留学生に特化した場面であることが望まれ、かつ留学の目的にそったものが望まれる。

表 4.6 ライフ場面の下位場面 (2002-2013)

ライフ場面	アイテム数	%
友人などと	73	52
メディア	28	20
講演・意見	22	16
地域社会	17	12
仕事	1	1
計	141	100

4.2.3 形態

形態に影響があるとして、①大学の授業 ②大学の事務的関連 ③私的関係(友人・家族) ④社会的関係(目上・初対面、職場の上下関係) ⑤大学以外の公的関連(職場・公共の場・メディア)と5つに分けて分類を始めたが、形態に特徴があると考えられる談話は結果的に10分類となった(表4.7次頁)。授業が多いのは講義場面が多いからである。講義の形態としては丁寧体が通常のため、留

学生には聞きやすいと考えられる。次は私的会話である。ほとんどが友人とおしゃべりのため、友人言葉や省略形などを用いた談話であるが、流行語（若者言葉を含む）は含まれていない。社会的会話は、例えば、学会開催、研究のためのデータ収集などで大学外部の人と接触したときにする談話と、教師や先輩に質問などをしたときである。このような場面では敬語が使用されるが、ビジネス場面とは違い、ある程度丁寧であれば問題はなく、敬語使用に関する能力に注目して測定しているとは考えにくい。事務的会話が少ないのは、キャンパス場面である大学施設関連、事務手続きなどの項目が少ないからである。学生課や教務課の職員は学生に対して敬語を使うことが多く、形式的な手続きに従っていることも多いため、留学生には情報として事前に知っておいてもらいたい。波及効果をねらい常時出題が望まれる。「説明」はオリエンテーション、授業の進め方、施設利用方法などで、23 アイテム中21 アイテムがモノログであった。2 アイテムはダイアログであり、司会者の質問に専門家が答える、留学生が面接試験で留学理由を説明する場面であった。最後の場面のような「説明」は留学生にとって頻度の高い談話形態であり、モノログだけではなく会話形式で聞き返しや質問などが入った説明場面も必要である。その他には「談合」「電話」「放送」「ニュース」が出題されており、数は少ないものの多様性が見られた。

4.2.4 トピック

今回、トピックはテスト・アイテムのラベル化という位置づけのみとした。今後、講義場面で扱うトピックの専門性により、自然科学、社会学、文学などの分類をすることで、トピックの偏りを防ぐことができ、アイテム・ライティング時には有用性があるだろう。

4.2.5 人数

人数は、相互性のある2人の会話の「対人間」、モノログで不特定多数にむかってされる「パブリック」、2人以上で相互作用を伴う会話の「グループ」、回想など誰かに聞かせる目的がないものは「個人内」とした。結果は「対人間」と「パブリック」でほぼ二分していることになった（表4.8）。

「対人間」のテスト・アイテムは友人とおしゃべりと教員への質問・相談が多い。友人とおしゃべりは場面の考察時に触れたが、留学生にとって必然性の高い内容が望まれる。留学生にとって教員との会話は必然性があり、今後も様々な状況下での出題が望まれる。

ここで、テストではなく現実に目を向けてみる。現在大学生の抱えている問題として、必然性が低いコミュニケーションを避ける傾向が見られる。携帯電話の普及、インターネットの発達により直接話すのではなくメールを使用する、ツイッターで自ら発信する・受信する、またフェイスブックで情報を共有することを、コミュニケーションをとっていると錯覚する。確かにコミュニケーションはしているのだが、直接対面してコミュニケーションをとっているわけではない。対面してのコミュニケーションに必要な、相手の反応を取り込み変化させて「話す・聞く」をおこなう行動が少なくなりつ

表 4.7 形態 (2002-2013)

形態	アイテム数	%
授業	176	40
私的会話	116	26
社会的会話	52	12
事務的会話	16	4
説明	23	5
演説・講演	27	6
談合	12	3
電話	12	3
放送	4	1
ニュース	2	0
計	440	100

表 4.8 人数 (2002-2013)

人数	アイテム数	%
パブリック	225	51
対人間	209	48
グループ	3	1
個人内	3	1
計	440	100

つある。三宅 (2006) は、日本人大学生間のコミュニケーションが少ないことに気づき、対面コミュニケーションの大切さを教えるべきだと述べている。対面コミュニケーションに必要なのは「人にわかる話し方、人の話がわかる聞き方」とし、「場の共有」を用いて説明している。「場の共有」は「空間 (space) の共有」と「経験 (experience) の共有」があり、大学生の問題は「空間の共有」をしていても「経験の共有」をしていないことにあると指摘する。そして、日留試が発したアカデミック・ジャパニーズ教育に対して、「既存の大学の教育に適応させればいいのか」(同掲p.192)と疑問を投げかけた (三宅 2006, pp.190-199)。日留試で出題されている友人との会話は、既に親しい友人であるため、「人間関係の維持」と考えられる。三宅がいう対面コミュニケーションの難しさは「人間関係を築く」ところにある。そのような聴解テスト・アイテムは、留学生に日本の大学生が抱える問題点の認知と、それを打破するための対面コミュニケーション・スキルの必要性を促す。大学教育そのものを問い直す動きがある中で、留学生も含め、どのような大学生に育って行くことが望まれているのかを発信していく義務が、日留試のような大規模テストにはある。

今回の結果では、グループ・コミュニケーションの少なさが顕著であった。グループ・コミュニケーションに分類された3アイテムはすべて3人の会話であり、2005年以降出題されていない。音声のみで3人の発話を聞き分けることは、声質がかなり違っていても難しいのだろう。2人の対面コミュニケーションよりターンの取り方などは複雑になるが、このような談話構造の理解も必要である。新日能試では、統合理解の中に3人の会話のテスト・アイテムがある。グループ・コミュニケーションのテスト・アイテムは真正性も高いため、聴解テストで画面などを利用し、誰が話者であるかわかるようになることを、近い将来に期待したい。最後に、個人内発話もほぼなかったが、コミュニケーション理解重視であるため、必要度は低い。

4.2.6 コミュニケーションの目的

コミュニケーションの目的の結果は、表4.9である。西川・小牧 (2002, p.4) が「情報を提供する」はあらゆる場面で見られると言っていたが、日留試の聴解テスト・アイテムすべてに「情報」を提供する (目的1) という目的が見られた。「情報を提供する」ことで知識を得る、理解を深める、行動する、問題解決をする、と考え、その他を加えて下位項目とした。「知識」が1番多く、ほとんどが「講義場面」であった。「問題解決」は例えば「病院の予約時間の交渉」「ヒーターの故障原因説明」など問題 (トラブル) を解決するものである。情報を提供することで、目的2である相手への「影響」、「行動変容・態度変容・意見変容」と複合しているアイテムも多かった。「問題解決」は、お互いのネゴシエーションが必要であり、日本社会ではどのような流れで解決にいたるのかという社会文化的な内容も含んでいる。「理解」はメディアや講演会などを聞くことで、「講義」のように特に「知識」を得る目的で聞いているわけではない。また「友人の怪我の状況」や「友人の意見」などを理解するものもあった。「行動」は、情報を与えることで相手に行動を起こさせ

表4.9 コミュニケーションの目的 (2002-2013)

目的1 情報	アイテム数	%
知識	199	45
問題解決	87	20
理解	70	16
行動	36	8
その他	41	9
必要	7	2
計	440	100

目的2 影響	アイテム数	%
行動変容	82	77
態度変容	14	13
意見変容	10	9
計	106	100

目的3 社会化	アイテム数	%
社会システム	9	60
人間形成	6	40
計	15	100

る「行動変容」と複合したものと、既にしなければならぬことは決まっており、その「条件」や「説明」などの情報を与える単独のものもあった。例としては、「オリエンテーション」「課題説明」「授業の進め方」などで、機能としては「説明」「指示」「依頼」「命令」などである。これらは「行動変容」と複合しているアイテムが多い。「その他」は、例えば「映画の感想をお互いに言う」「体験談を聞く」「意見を聞く」などであり、相手から情報を引き出すというものである。目的1と2の「情報」と「影響」が複合的になっているアイテムは、会話参加者の相互作用があり、社会文化的な要素が見せられるテスト・アイテムである。

目的3の「社会化」はアイテム数としては少なかった。「社会システム」は既存の社会システムに対する提案や環境について、「人間形成」は講演会で経験談などにより聞き手に対して啓蒙するというものであった。講演会は特に留学生にとって身近な場ではないが、卒業生の話や授業として偉人賢人について調べる・話すなどの場面とすると留学生の大学生活に近づく。

4.2.7 機能

聴解テスト・アイテムは談話であるため、機能は1つとは限らずほとんどが複数の機能を持っていた。「4.2.6 コミュニケーションの目的」で、「情報を提供する」がすべての聴解テスト・アイテムに見られたように、「情報を授受する機能」もすべてのアイテムで見られた（表4.10）。「情報を授受する機能」だけのアイテムもあったが、ほとんどは他の機能と組み合わせたり談話が成り立っている。ここではどんな機能が多く使われていたかという点から分析・考察をする。

・情報を授受する機能

「講義する」のように一方通行のものと「質問する・応える」のように双方向のものがあった（表4.10）。「話のきっかけを作る・応ずる」はテスト談話を始めるきっかけとしてその機能を使っているという印象が強い。「語る」は講演会の場で自分の体験談を話すようなものだが、聞く機会は高くない。例えば、著名人が大学を訪れ講演をしたとして、その講演の内容が理解できずとも大学生活には支障をきたさない。内容の頻度、必然性、必要性の観点からの見直しが必要である。「報告する」「発表する」は大学の授業やゼミに多く見られ、それを知るよい機会であり、今後増加を期待したい機能である。

表4.10 情報を授受する機能（2002-2013）

情報を授受する機能	アイテム数	%
講義をする	152	35
質問する・応える	88	20
説明する	41	9
話のきっかけを作る・応ずる	37	8
教える・知る	36	8
相談する・アドバイスする	27	6
語る	23	5
知らせる	22	5
教える・学ぶ	7	2
発表する	5	1
報告する	2	0
計	440	100

・聞き手に働きかける機能

「聞き手に働きかける機能」は、話し手からの一方通行なものである（表 4.11）。

「提案する」が多くなった結果の理由として、選択肢を4つ作らなければならないという形式上、談話内に惑わし選択肢になりうる情報を入れなければならないからという意図が感じられる。例えば、談話構成の中で、話し手が1つの提案だけではなく複数の提案をし、聞き手が1つを選ぶという構成になっていると4つの選択肢が作りやすい。4つの選択肢を作らなければならないというテスト形式に、談話構成あるいは談話展開が束縛されることは、本来あってはならない。このような不自然な内容や展開の談話を避けるために、長い談話と

複数の課題がある形式を利用することも1案である。「許可を得る」「注意する」は日本の社会文化的ルール、また大学の授業でのルールも含まれ、留学生には必要度が高いと考えられる。

表 4.11 聞き手に働きかける機能 (2002-2013)

聞き手に働きかける機能	アイテム数	%
提案する	58	47
命令する	17	14
依頼する	15	12
誘う	12	10
要求する	11	9
文句を言う	3	2
励ます	2	2
愚痴る	2	2
許可を得る	1	1
注意する	1	1
注意を引く	1	1
計	123	100

・相手の働きかけに応える機能

「相手の働きかけに応える機能」は、「聞き手に働きかける機能」と対になって1つの談話に現れることが多い（表 4.12）。教師からの提案に対してはほぼ「(提案を)承諾し」、友人同士では、「(提案に)賛成・反対する」も見られた。「提案する」に対して、「提案の根拠を聞く」「提案の結果を考える」「提案をシミュレーションする」なども応え方の1つとなりえ、より多様な展開があったほうがよい。

聴解テストで、受験者が聞き手の立場になって聞いていると考えた場合、パブリック・コミュニケーションのアイテムでは不特定多数の聞き手の1人として聞くことになり、

自らその談話に参加する可能性が低いという点で、実際の聴解とテスト聴解との差は小さい。対人間コミュニケーションのアイテムの場合、第三者として聞く「盗み聞き（後述）」の場合と、談話の中の聞き手役（例えば、プリタスクが大学生同士の会話のような場合、自分が談話内の聞き手になったと想像して聞く）になって聞く場合とがある。後者の聞き手役としての聴解も、実際に聞き手になって談話に参加できるわけではないため、実質的にテスト聴解は「盗み聞き」である。

ここで、「盗み聞き (eavesdrop)」の説明をする。「盗み聞き」とは、「他の人達の会話を立ち聞きす

表 4.12 相手の働きかけに応える機能 (2002-2013)

相手の働きかけに応える機能	アイテム数	%
承諾する	48	54
賛成・反対する	21	24
拒否する	4	4
断る	3	3
解決法を言う	3	3
礼を言う	2	2
詫びる	2	2
いいわけする	2	2
許可を与える	2	2
認める	1	1
文句を言う	1	1
計	89	100

る、自分たちに向けられていないモノログを聞く」ことであり、第二言語学習時に録音された談話を聞く練習に対しての表現である。岡崎・岡崎 (2001,p.67) は、「盗み聞き」について、日常生活ではこのような聞き方をしないことから、教室外の聴解行動に転移ができないとしている。Flowerdew & Miller (2005)は、通常の生活では表立って行われぬ「盗み聞き」は会話に参加しそれをコントロールする権利がないことから、学習者の聴解は実際のものより難しくなると指摘をした。このような考えに対し、Field (2008,p.61)は、「盗み聞き」の練習は、以下の5つの点から実際の場面での聴解力を高める効果があると反論している。

1. 母語話者の会話を聞くことは、短い発言と交代で話すことを聞き取る練習になる。
2. リスニングの対象となる会話は、聞き手と話し手がどのように話をするかを示す L2 会話の見本となる。
3. 実生活で聞き手が「盗み聞きする」ことがないというのは、誤りである。L1 の聞き手も、周りの会話をモニターすることはある。ただ、あまり注意して聞いていないので理解していることに気づいていないだけである。
4. すべての会話が、参加者が2人いて、同程度の参加をしていると考えるのも誤りである。Field との個人的な話において、Gillian Brown は、4、5人のグループで話している場合、その中の誰か1人は、70%あるいはそれ以上の時間、「盗み聞き」をしている可能性があり、だれかが話し終わる度に返事をしなければいけないというプレッシャーを必ずしも感じていない、と指摘している。
5. 「盗み聞き」の練習は、学習者に、いつでも応える準備をしていなければいけない「話し手を目の前にした会話」よりも、L2に注意を払い練習するために必要な機会を多く与える。

(富田他 2011, pp.94-95)

例えば、テレビドラマを見ているときは、出演者の談話を盗み聞きしている。「自分たちに向けられたモノログ」ではないニュースや演説でも、盗み聞きしているうちに興味がわくということもある。Field が挙げている 4.のように2人あるいはグループで会話をしているところへ自分が合流した場合、しばらくは「盗み聞き」の状態が続く。このように、我々は実生活でも盗み聞きをしている。テスト聴解が「盗み聞き」であることに問題はなく、テスト受験そのものが学習効果をあげることを忘れてはいけない。

・自己の感情を表す機能

「自己の感情を表す機能」が含まれているアイテムは、少なかった (表4.13)。「感情」は、ある出来事あるいは特定の物や人に対して起こる。個人差が大きく、受験者に負の感情 (怒り、悲しみ、不快など) を呼び起こすことは望ましくない。よって、テスト・アイテムに含まれることは少なく、今後も増加はないと考えられる。しかし、「怒る」(表4.13) 以外の感情表現は、日本の社会文化的要素も含まれ、よい人間関係を築ききっかけにもなり、継続的な出題が望まれる。

表 4.13 自己の感情を表す機能 (2002-2013)

自己の感情を表す機能	アイテム数	%
感想を言う	12	60
怒る	5	25
祝う	1	5
感謝する	1	5
感心する	1	5
計	20	100

・個人・社会化に関する機能

上記の4つには分類できなかつたため、追加した機能である(表4.14)。「人間関係を維持する」とは聞き手に徹して聞くような場合である。例えば、話し手が愚痴を言っているのを相づちのみで聞いており、自らの感情も表さず、話し手とのインターアクションを回避しているような状態である。「意見を言う」は誰かに聞いてほしいという聞き手に対する働きかけが談話から読み取れず、情報は発しているが独り言のように感じられる談話である。「話し合う」や「議論する」は聞き手と話し手とに分けられなかつた談話である。「啓蒙する」「諭す」は人の社会化に関する機能と考えた。

本文談話のコミュニケーションの機能には、多くの下位項目があり、1つの談話内に複数の機能が含まれることもあった。しかし、講義は「情報を授受する機能」の「講義をする」だけである。確かに、講義理解は留学生にとって重要ではあるが、談話機能が1つしかない聴解アイテムが多くなることは、それだけ画一化した内容となり、あまり望ましいとは言えない。今後も多種多様なコミュニケーション機能を持った談話の出題を望む。

4.2.8 情報を得る理由

日留試の聴解テスト・アイテムすべてに「情報を授受する機能」が含まれていた。その情報を得る理由が表4.15である。「学術・研究のため」「学習のため」の出題率合計が55%であるが、留学生に特化したテストであるため、さらにも増えてよいと考えられる。「生活のため」は衣食住、娯楽などであり、「自分のため」は、誰かの体験談や講演会などで自己成長を促すような内容を聞くことであった。

4.3 テキスト・バンクの提案

本節では、4.2で行なった個々のアイテムの基本データの分析をもとに、アイテム・ライターがテスト・アイテムを作成する際に参考になるのではないかと考え、テキスト・バンク構築の提案をする。テキスト・バンクとはアイテム・バンク(次節で詳述)に通常データとして蓄積される「アイテムの特性パラメータ(識別力や困難度)」と「固定カテゴリ(テスト実施日、場所、レベルなど)」はなく、聴解テスト・アイテムのスクリプト、つまりテキストとその特徴のみをデータとして蓄積しておくものと定義する。

4.3.1 アイテム・バンク

1章の1.1.(2)(p.2)で述べたが、日留試は日本語大規模テストで初めてアイテム・バンクを利用した。本研究では、聴解テスト・アイテムの分析と考察結果から聴解テストのテキスト・バンク構築を試みるため、ここでアイテム・バンクについて記述する。

菅井(2003,p.107)はアイテム・バンクが大規模テストに適している理由として、次の3点を挙げて

表 4.14 個人・社会化に関する機能 (2002-2013)

個人・社会化に関する機能	アイテム数	%
人間関係を維持する	30	37
意見を言う	26	32
話し合う	14	17
啓蒙する	8	10
議論する	3	4
諭す	1	1
計	82	100

表 4.15 情報を得る理由 (2002-2013)

情報を得る理由	アイテム数	%
学術・研究のため	182	41
学習のため	60	14
自分のため	58	13
人間関係のため	50	11
大学生活のため	45	10
生活のため	45	10
計	440	100

いる。①大規模テストのように大量のテスト・アイテムが必要な場合、一度出題したテスト・アイテムをアイテム・バンクに保管し、リサイクルすることが行われる。②一度出題したテスト・アイテムは、統計的に分析を行えば項目困難度と識別度（識別力）がわかる。③識別度（識別力）の低いアイテムは排除するか修正を行い、使用可能なテスト・アイテムを問題の種類（読解、聴解など）、場面、タスク・タイプ、ジャンル、難易度（困難度）などをコンピュータのデータ・ベースに保管しておけば、必要なテスト・アイテムのデータを瞬時に取り出し、フォームを組むことができる。

アイテム・バンクに必要な情報として、アイテムそのもの（問題文、アイテムのテキスト、選択肢など）の他に、Hughes（2003 邦訳 p.249）は以下の情報が必要だとしている。

1. 検索するための情報。例えばアイテムの内容、使用に適したクラスのレベル、シラバスや教科書内でのそのアイテムの段階、使用しているテストテクニック、配点などに関するもの。
2. 正答情報と採点方法に関する指示。
3. そのアイテムの測定統計的な数値情報。例えばむずかしさのレベル、識別力の指標。これらは以前の試行から得られたデータである。
4. そのアイテムに関する注釈的情報（作成された日時、試行された日時など）。

また、伊東（2008a,p.2、2011,p.58）はアイテム・バンクの機能を5つ挙げ、アイテム・バンクに蓄積する情報の使用目的を明確にしている。

蓄積機能：作成した問題項目を、コンピュータ内に蓄積し、保存することができる。また、問題項目のほか、過去の試験結果のデータや項目分析結果も蓄積できるので、将来の問題項目作成を効率よく行うことができる。

抽出機能：出題基準や評価対象領域、また困難度などの条件に基づき、必要な数の項目を抽出して試験問題を構成・作成することができる。

組み替え機能：抽出条件が同じでも、設問の組み合わせや選択肢の組み合わせが可能となるので、異なる試験問題を作成することができる。

加工機能：蓄積・保管されている試験や問題項目を、測定目的に応じて編集・加工・削除することができる。また、既に蓄積されている項目を基礎に、全く新しい問題項目を作成することができる。

製版機能：抽出された問題項目を、パソコン上でレイアウト・編集ができる。

伊東（2011）は上記に挙げたアイテム・バンクの5つの機能をふまえて、文法テストのアイテム・バンクを設計した。その際に使用した分類項目は、①項目の蓄積 ②項目の特性パラメータ¹²の算出・登録 ③階層カテゴリ（アイテムとその分類にかかわる）の構造化 ④固定カテゴリ（アイテムとその特性にかかわる）の構造化 の4つであった。③はアイテム・バンクからテスト・アイテムを再利用したり加工したりする際に、取り出すための形式や内容面を階層的に分類する。④は項目困難度、識別力、実効選択肢数の情報である。（伊東 2011, pp.64-66）

¹² 項目応答理論（item response theory）は、「能力の高い受験者はよりそのアイテムに正答する確率が高い」という考え方を前提とし、個々のテスト・アイテムを受験した者の応答パターンからアイテムの特性値（パラメータ）と受験者の能力値を推定する。ある能力値を持つ受験者があるアイテムに正答する確率を表したものが項目特性曲線であり、これによってアイテムの困難度、識別力（識別度）、当て推量などがわかる。（石川他 2011, pp.41-43）

日留試は1年に2回、複数フォームのテストを作成しなければならないため、大量のテスト・アイテムが必要である。大量のテスト・アイテムを毎回作成するのは困難を伴うが、蓄積機能、抽出機能、組み替え機能を利用することで、時間と労力の削減ができる。また、既に困難度や識別力がわかっていることで信頼性が確保できる。加工機能は文法、語彙、漢字など、言語知識を測定する場合などに有効である。製版機能は試験的にテストを作成し、スペースやバランスなどをすぐに検討できることで時間的な節約が可能になる。したがって、アイテム・バンクの構築は、大規模テストには欠かせないと言ってよい。

4.3.2 聴解テスト・アイテム作成のためのテキスト・バンク構築

日留試の聴解テストのテキスト・バンクを構築する際、Hughes の挙げた必要情報の「1.検索するための情報」と伊東の「③階層カテゴリ」をどのようにデータ化できるかを記述する。

4章で分析した結果、1アイテムに独立して存在する項目（例：理解の種類、場面、形態、人数、情報を得る理由）と、複合的に存在する（コミュニケーションの目的、機能）項目があった。聴解テスト・アイテムは上下左右に様々な特徴が絡み合っており、複合的に存在する項目もあったため、階層カテゴリの構造を立体的なモデルとして作成することにした。テキスト・バンクは本文テキストを対象とするため、プリタスクに関わる「理解の種類」と「聞く目的」は除外する。本文の特徴の「コミュニケーションの目的」は、すべてのアイテムが目的1の「情報を提供する」目的を持っているため、アイテムの固有差が見えにくいと考え、これも除外した。「検索するための情報」を「場面、形態、人数、機能」の4項目とし、アイテムを分類、整理する。

まず、コミュニケーション形態を各場面ごとに見ると表 4.16、表 4.17、表 4.18（次頁）となる。次に、このコミュニケーション形態を縦軸にし、3場面（アカデミック・キャンパス・ライフ）を横軸にして、テスト・アイテムを分類すると、テキスト・バンクの外枠ができる（図 4.6, p.73）。

表 4.16 アカデミック場面の形態 (2002-2013)

形態	アイテム数	%
授業	176	89
私的会話	7	4
社会的会話	0	0
事務的会話	0	0
説明	13	7
演説・講演	0	0
談合	1	1
電話	0	0
放送	0	0
ニュース	0	0
計	197	100

表 4.17 キャンパス場面の形態 (2002-2013)

形態	アイテム数	%
授業	0	0
私的会話	46	45
社会的会話	24	24
事務的会話	10	10
説明	5	5
演説・講演	3	3
談合	7	7
電話	6	6
放送	1	1
ニュース	0	0
計	102	100

表 4.18 ライフ場面の形態 (2002-2013)

形態	アイテム数	%
授業	0	0
私的会話	66	47
社会的会話	25	18
事務的会話	6	4
説明	5	4
演説・講演	24	17
談合	4	3
電話	6	4
放送	3	2
ニュース	2	1
計	141	100

図 4.6 (次頁) から、アカデミック場面のコミュニケーション形態はほとんどが「授業」であることがわかる。これはアカデミック場面の下位場面がほとんどが「講義」だったことから予想はつく。しかし、「講義」場面がすべて「授業」形態ではない。例えば、2 つの講義場面のアイテムについて見てみる。アイテム①は 2006 年 2 回目 3 番「サクラの交配」というアイテムである。①は「場面：アカデミック、形態：授業、人数：パブリック、機能：情報を授受する」というデータを持っている。これは図 4.6 の①の箱に入る。アイテム②は 2005 年 2 回目 7 番「コミュニケーショントレーニング」というアイテムで、②は「場面：アカデミック、形態：説明、人数：パブリック、機能：情報を授受

する・聞き手に働きかける」というデータを持ち、②の箱に入る¹³。

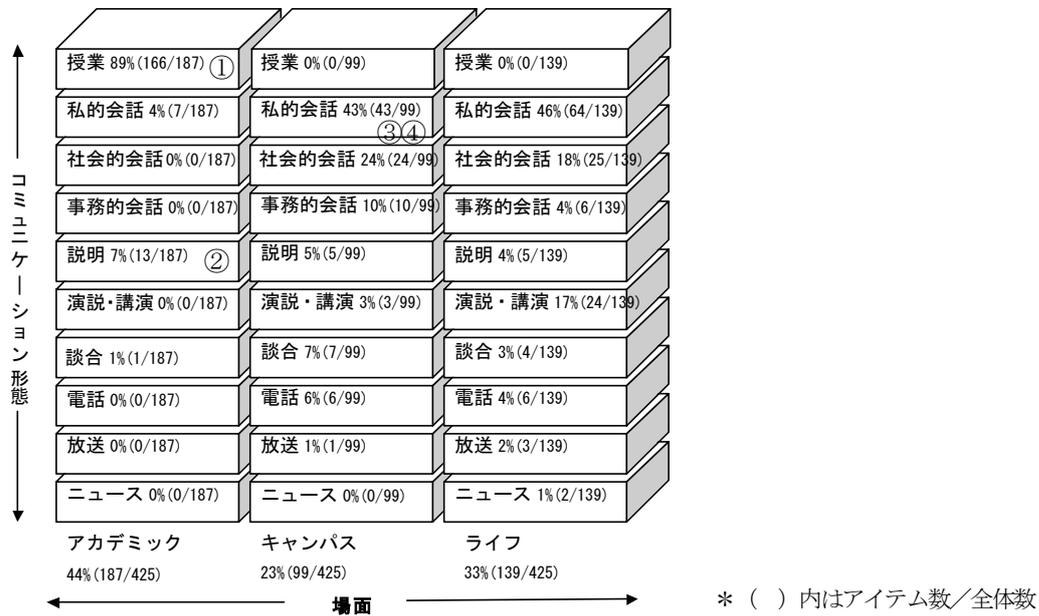


図 4.6 聴解テスト・アイテムのテキスト・バンク (2002-2013)

¹³ アイテム詳細は参考資料5「テキスト・バンクの説明時に使用した聴解テスト・アイテム①と②」(p.197)を参照。

さらに、テキスト・バンクを詳細に見て行くために、キャンパス場面から2アイテム¹⁴を選び説明する。③2006年2回目2番「高校生の大学生活体験」は、「場面：キャンパス、形態：私的会話、人数：対人間、機能：情報を授受する・聞き手に働きかける・相手の働きかけに応える」、④2006年2回目17番「人形を使った育児体験」は、「場面：キャンパス、形態：私的会話、人数：対人間、機能：情報を授受する・自己の感情を表す・個人社会化に関する」というデータを持ち、同じ箱に入っている(図4.6、図4.7)。図4.7は③と④が入っているキャンパス場面のみを拡大したものである。キャンパス場面の横軸は「人数」で4分割される(図4.7)。この箱を拡大すると図4.8(次頁)になる。

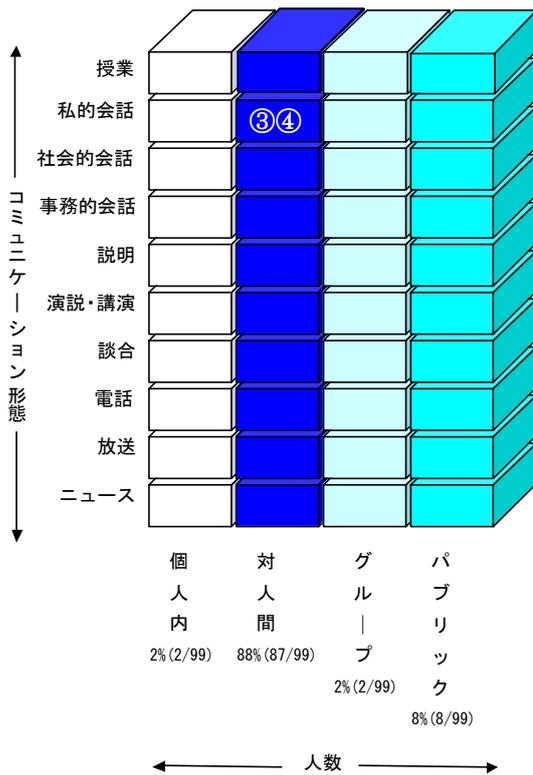


図4.7 聴解テスト・アイテムのテキスト・バンク キャンパス場面・人数(2002-2013)

¹⁴ アイテム詳細は参考資料6「テキスト・バンク分類箱の説明時に使用した聴解・テスト・アイテム③と④」(p.197)を参照。

図 4.8 の縦軸は「コミュニケーション形態」、横軸は「場面」と「人数」、奥行き軸が「機能」である。

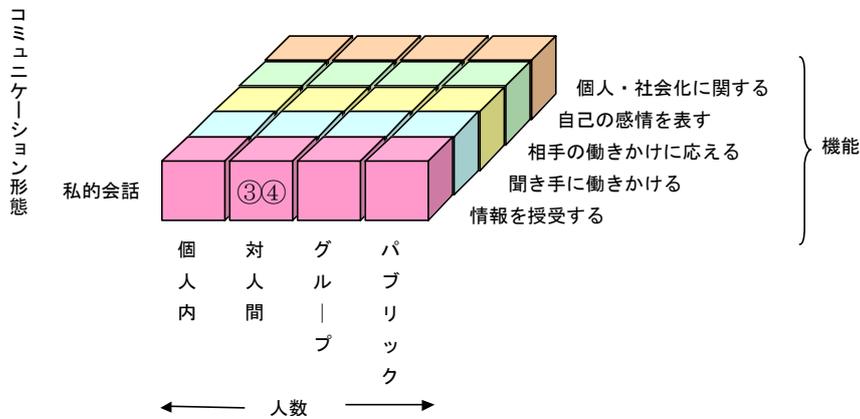


図 4.8 聴解テスト・アイテムのテキスト・バンク 私的会話・キャンパス場面・人数・機能 (2002-2013)

テスト・アイテム③と④はやはり同じ箱に入っている (図 4.8)。この箱を上から見てみると図 4.9 となる。

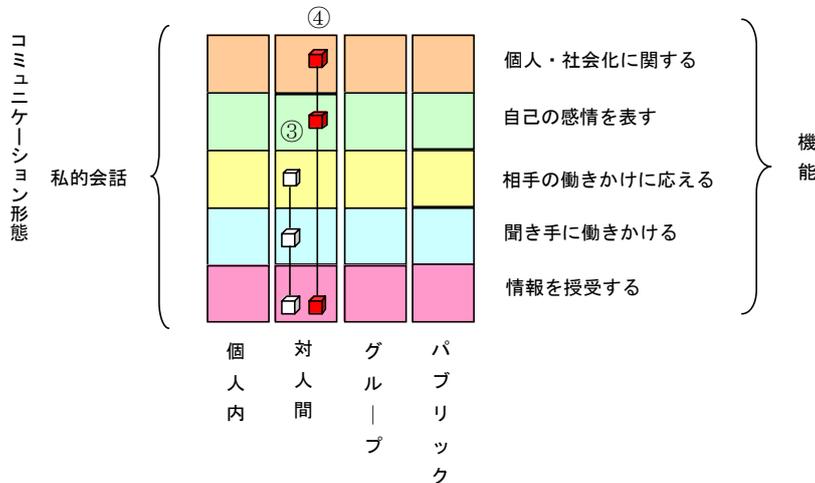


図 4.9 聴解テスト・アイテムのテキスト・バンク

私的会話・キャンパス場面・人数・機能 鳥瞰図 (2002-2013)

図 4.9 から、③と④は「私的会話」で「対人間」コミュニケーションである同じ箱に入っているが、違う機能の箱に入っていることがわかる。アイテム・バンクを構築する際には、このような談話の詳細なデータは不必要かもしれない。しかし、このデータは談話が多面的な情報を持っていることを示しており、それが理解にどのような影響があるかという聴解研究の基礎データとなる。また、アイテム作成の際に、多面的な談話作成のためのチェック項目を与え、多種多様な談話作成に役立つのでは

ないだろうか。テキスト・バンクに蓄積されたデータを土台にしてアイテム・ライターがアイテムを作成することで作成スピードが向上し、多様な視点から作成することが可能になる。談話の内容や談話構成の偏りを減らすことにも貢献できるだろう。

4.4 改定前2002年版と2010年改定版の比較・考察

最後に、聴解に関係する 2002 年のテキスト・タイプと、2010 年改定版の談話の種類の違いを比較し、今回の分析結果（2013 年まで）をふまえて考察する。

2002 年 試験に含まれるテキスト・タイプ（日本国際教育協会 2003,p.247）

1. 対人関係作り・維持に関するもの
2. 指示・命令に関するもの
3. 事物や人物の記述・メタ言語に関するもの
4. 架空の世界に関するもの
5. 放送メディア（ラジオ・テレビなど）に関するもの
6. 公共アナウンスに関するもの
7. 街頭アナウンスに関するもの
8. スピーチ
9. 講義、演習、実験、フィールドワークなどに関するもの
10. 打ち合わせ、討論（会議、シンポジウム、意見交換など）に関するもの
11. 面接、インタビューに関するもの
12. その他

2010 年改定版 談話の種類（日本学生支援機構 2010,p.292）

1. 講義、講演
2. 演習や調査活動に関わる発表、質疑応答、意見交換
3. 学習上または生活上の相談ならびに指導、助言
4. 実務的・実用的な談話など

2002 年当初のテキスト・タイプの内容は、生活全般を広く網羅しており、アカデミック・リスニングは、「9.講義、演習、実験、フィールドワークなどに関するもの、10.打ち合わせ、討論（会議、シンポジウム、意見交換など）に関するもの」の2つである。2010 年版の談話の種類では、アカデミック・リスニングとして、「1.講義、講演 2.演習や調査活動に関わる発表、質疑応答、意見交換」と、全体の4分類のうちの半分を占めるようになった。本章で分析した 2002 年から 2010 年への場面の推移からも（図 4.5 「場面の割合の推移」 p.60 を参照）、大学の講義や授業での談話が重視されるようになってきたことがわかる。2010 年版の「3.学習上または生活上の相談ならびに指導、助言」は、キャンパス場面である。2002 年版の「2.指示・命令に関するもの」に近い状況をさしていると思われる。2010 年版の「4.実務的・実用的な談話など」がライフ場面である。「実務的」は、アルバイトの場面や 2002 年版の「11.面接、インタビューに関するもの」にあたると思われる。「実用的」は、2002 年版の「3.事物や人物の記述・メタ言語に関するもの 4.架空の世界に関するもの 5.放送メディア（ラジオ・テレビなど）に関するもの 6.公共アナウンスに関するもの 7.街頭アナウンスに関するもの 8.

スピーチ」ではないかと考えられる。2002年版では細かく分類されていたものが、2010年版では「4. 実務的・実用的な談話など」と1つにまとめられたことから、アカデミック場面が重視され、ライフ場面が減少傾向にあると判断してよいだろう。この改定から読み取れることは、留学生にとって講義聴解が重要であることを強調し、内容の妥当性をあげた点である。聴解テストの困難度に関して、Shohamy & Inbar (1991)は、テスト・アイテムのテキスト・タイプが ①ニュース放送 ②講義講演 ③協議談話 の順に困難度が下がるという研究結果を得ている。日留試は、聴解テストの講義場面を増加することで、テスト全体の困難度をあげる効果を狙ったと考えられる¹⁵。この改定は、留学生を受け入れる機関が、留学生に講義理解能力と高い聴解能力を求め、そのニーズに、日留試の作成機関が応えた結果であろう。また、研究者たちのアカデミック・ジャパニーズの研究成果の影響も否定できない。

ここで、日留試の改定前と改定後のテスト・アイテムについて考察するために、聴解以外の変更について触れる。日留試の実施機関である日本学生支援機構は、日本語科目の変更について以下のように説明している（日本学生支援機構HP、日本留学試験「日本語」改定についてより¹⁶）。

- ① 聴解・聴読解と読解の得点範囲及び試験時間を変更し、バランスのとれた領域構成となるよう改善
- ② 読解問題に長文及び複問（1つの文章に対し複数の問題をつけたもの）を導入し、大学等で求められる読解力を測定するという観点から、出題形式を改善
- ③ 聴解・聴読解の試験時間を見直し、試験時間が長く受験者の集中力が続かないという問題点を改善
- ④ 記述で出題される課題のタイプを拡充し、採点基準及び得点表示を改善。これにより、大学等における記述の成績の利用を促進

上記の変更点について具体的に記述する。2010年の改定後、聴解テスト・アイテムが20アイテムから15アイテムに減少したことに加え、聴読解テスト・アイテムも20アイテムから12アイテムに減少している。一方、読解は20アイテムから25アイテムに増加した。3分野の点数配分については、2002年は聴解120点、聴読解120点、読解160点であったが、2010年改定後は、聴解・聴読解で200点、読解200点と変更された。テスト・アイテムの増減および点数配分の変更は、聴解と読解のバランスという①と③の理由から理解できるが、テスト実施の順番が、「記述・読解・聴解・聴読解」から「記述・読解・聴読解・聴解」となったことについては、説明がなく、③に「集中力が続かない」とあるが、一番集中力が必要な聴解が、テストの最後になったことは解せない。②の読解のテスト形式の変更は、大学生の学習状況により即しており、読解力の強化を図った改良点と言ってよい。では、聴解能力が軽視されてきたのかということとは言えない。聴解と聴読解の点数が1つにまとめられたことから、聴読解は聴解能力を測定しているテストと考えてよいだろう。聴読解は、聴解能力と読解能力の複合力を測定する目的で始められたが、テスト時間の制限からか、聴読解で読ませる読解文は非常に短く限られており、当初から聴解能力測定の比重が高かった。2010年改定後の聴読解のテスト・アイテムは、2002年当初の頃よりさらに文章提示が減り、グラフや図が多用されている。聴読解の「読」

¹⁵ アカデミック場面が増加し始めたのは、図45 (p.60) からわかるように、2002年から始まっており、2010年の改定を境に急増したわけではない。しかし、2002年当初から毎回見られたアカデミック場面のアイテム増加は、2010年の改定後は、全体の6～7割と一定になり、改訂を期に、アカデミック場面の出題割合を確立したと言える。

¹⁶ 2015年3月 <http://www.jasso.go.jp/eju/kaitei.html>

は、聴解から得た情報を元に視覚情報の中から解答するという役割が強くなっており、いずれ聴読解という分類がなくなる可能性が高い。聴解テストには、視覚情報があるタイプとないタイプがあるという言い方のほうが、一般的に理解されやすい。以上のことから、2010年改定後の日留試は、聴解能力と読解能力を等しく測定するという方針に変わり、内容および形式とも、より真正性が高まったと言える。今後、真正性を追求するため、ある程度の長さのテキストを読んだ後に聴解をする、読解と聴解の複合形式の「聴読解」の必要性の議論が起こることを期待したい。

2010年改定前と改定後の違いを、データの比較により分析する。

①理解の種類

表 4.19 理解の種類の変遷 (改定前と改定後)

改定前(2002-2009)			改定後(2010-2013)		
	アイテム数	%		アイテム数	%
部分理解	230	72	部分理解	88	73
複数部分理解	47	15	複数部分理解	14	12
全体理解	20	6	全体理解	16	13
部分・全体理解	23	7	部分・全体理解	2	2
計	320	100	計	120	100

改定前と改定後の違いは、全体理解が6%から13%と増加し、部分・全体理解がほとんど出題されなくなったことである(表 4.19)。アカデミック場面が増加した結果(表 4.20)、講義理解に必要な全体理解も増加している。

②場面

表 4.20 場面の比較 (改定前と改定後)

改定前(2002-2009)			改定後(2010-2013)		
場面	アイテム数	%	場面	アイテム数	%
アカデミック	118	45	アカデミック	79	66
キャンパス	83	23	キャンパス	19	16
ライフ	119	32	ライフ	22	18
計	320	100	計	120	100

図 4.5 「場面の割合と推移」(p.60) で、場面の推移を示したが、改定前と改定後の全体に占める割合を比較する(表 4.20)。数値上では、改定後のアカデミック場面の増加、キャンパス場面とライフ場面の減少という特徴が見られるが、アカデミック場面は、2002年当初から徐々に増加し始めており、改定時の2010年を境に大きな変化を見せたわけではない。改定前に、アカデミック場面の増加が見られ、改定後はある一定の割合(6割~7割)で出題されている。

③形態

表 4.21 形態の比較（改定前と改定後）

改定前(2002-2009)			改定後(2010-2013)		
形態	アイテム数	%	形態	アイテム数	%
授業	100	31	授業	76	63
私的会話	89	28	私的会話	27	23
社会的会話	44	14	社会的会話	8	7
事務的会話	15	5	事務的会話	1	1
説明	23	7	説明	0	0
演説・講演	21	7	演説・講演	6	5
談合	12	4	談合	0	0
電話	10	3	電話	2	2
放送	4	1	放送	0	0
ニュース	2	1	ニュース	0	0
計	320	100	計	120	100

形態は、場面の影響が顕著である。アカデミック場面では講義場面が増加しており、それに伴って形態の「授業」も増加している。「演説・講演」が減少したのは、内容が講義・授業に特化されてきたからである。キャンパス場面が減少しているため、「事務的会話」や「説明」が減少し、ライフ場面の減少から「電話」「放送」「ニュース」も減少した。外部の人と接触したときの「社会的会話」も減少しており、外部の人と接触するような学会参加や就職関連の場面は、大学入学時や1、2年生時には発生頻度が低いからだと考えられる。「談合」が改定後になくなったのは、大学生活ではあまり見られない形態だからであろう。形態に多様性が見られた改定前と比べ、改定後は「授業」にかなりの比重が置かれるようになったことがわかる。

④トピック

トピックに関しては、本研究では各テスト・アイテムのタイトルとしてラベルの機能を持たせただけであり、分析項目としては扱わなかった。ここでは、2002年から2009年までと（表4.22 pp.81-84）、改定後の2010年から2013年までのテスト・アイテムのトピックの一覧（アイテムのタイトル）を載せ（表4.23 pp.85-86）、改定前と改定後に異なった特徴が見られるかを考察する。なお、トピックは場面と関連があるため、場面も入れてある。本研究ではアカデミック・リスニングに注目しているため、アカデミック場面のトピックのみを対象とする。アカデミック場面のアイテムから、レポートや論文に関する説明など、「学習のため」に情報を授受するアイテムは省き、狭義のアカデミック・リスニングである「学術・研究のため」の内容のトピックのみを灰色にし、考察の対象とした。

改定前に比べ改定後では「学術・研究のため」のトピックの割合が非常に多い。これはアカデミック場面の増加（表4.20 p.78）と、アカデミック場面の中でも「授業形態」の増加（表4.21 上載）、「講義をする機能」の増加（表4.26 p.88）の結果を反映している。改定前のトピックには、「マラソンの死点、臭覚の順応、体内時計」など身体に関するトピック、「カルチャーショック、心理学行動、報酬とやる気」などの心理に関するトピック、「屋上緑化、防犯システム、携帯電話の普及」などの社

会生活に関するトピック、「鳥の観察、七色の虹、魚の保護」などの自然に関するトピック、「産業の発展、販売戦略、商品の宣伝方法」などの経済に関するトピック、「紙、ボールペン構造、プラスチック」などの物品に関するトピックなど、身近なものが多い。生活の中で身近ではなく、多少特化したトピックとしては「マングローブ」のように生活の中ではあまり聞かない語彙や、「微生物への石鹼の影響、メタ記憶」など語彙が専門的であるトピックが見られたが、数は少ない。

改定後のトピックも「筋肉トレーニング、目の錯覚、骨の役割」などの身体に関するトピック、「うそ、子どもの証言、教室で座る位置」などの心理に関するトピック、「情報化社会、景観、高齢化社会におけるスーパーマーケット」などの社会生活に関するトピック、「小さい魚、鳥の実験、蝶の飛び方」などの自然に関するトピック、「銀行の取り組み、弁当屋の販売戦略、商品の値段」などの経済に関するトピック、「新ペットボトル、携帯電話使用、照明器具」などの物品に関するトピックなど、身近な内容である。改定前と比較すると「ゲラドヒヒ、ホタルイカ、チンパンジー」など特化した語彙、「言葉の浄化作用、種の発芽、二酸化炭素と穀物、睡眠と記憶、内圧外圧」などの専門的な語彙のトピックが増えてきた。講義関連のアイテムが増加したことに伴い、専門的な語彙も増加してきてはいるが、大半は人や人の生活に身近なトピックであり、その方向性は改定前と改定後では同様であると言える。Bachman (1990 邦訳 p.159)は話題に関して受験者にバイアスがかかってしまうために慎重にならざるをえないが、身近で一般的な話題ばかりを選ぶと新情報がなく受験者にはつまらなく感じるが生じると述べている。日留試のトピックは受験者にとっては日常的に聞かないような特化したトピックや、専門的な内容に関するものもあり、工夫がみられる。

表 4.22 聴解テスト・アイテムのトピック一覧改定前 (2002-2009)

場面のア: アカデミック キ: キャンパス ラ: ライフ

施行年	回	問題番号	トピック	場面	施行年	回	問題番号	トピック	場面
2002	1	1	飛行機予約	ラ	2003	1	1	試験科目	キ
2002	1	2	ゼミ日時	キ	2003	1	2	アンケート回収数	キ
2002	1	3	伝言	ラ	2003	1	3	課題	キ
2002	1	4	時計のプレゼント	ラ	2003	1	4	集中講義	キ
2002	1	5	引っ越し祝い	ラ	2003	1	5	電車停止	ラ
2002	1	6	履修届	キ	2003	1	6	スピーチコンテスト	ラ
2002	1	7	図書館の延長	キ	2003	1	7	ボランティア活動	ラ
2002	1	8	本屋	ラ	2003	1	8	手紙の切手	ラ
2002	1	9	連休の手伝い	ラ	2003	1	9	練習さぼり	キ
2002	1	10	ミーティング日程	キ	2003	1	10	授業指示	キ
2002	1	11	主役	ラ	2003	1	11	図書貸し出し	キ
2002	1	12	講義課題	キ	2003	1	12	睡眠	ラ
2002	1	13	後輩と再会	ラ	2003	1	13	印刷	ラ
2002	1	14	演劇発表会	キ	2003	1	14	発表のしかた	ア
2002	1	15	健康相談	ラ	2003	1	15	感受性の器	ラ
2002	1	16	体内時計	ラ	2003	1	16	生活と環境問題	ラ
2002	1	17	研究関連文献探し	ア	2003	1	17	エネルギー問題	キ
2002	1	18	テレビ報道	ラ	2003	1	18	自転車の貸し出しサービス	ラ
2002	1	19	朝顔	キ	2003	1	19	カルチャーショック	ア
2002	1	20	経済学	ア	2003	1	20	マングローブ	ア
2002	2	1	留守録レポート	キ	2003	2	1	待ち合わせ	ラ
2002	2	2	海外旅行	ラ	2003	2	2	寮の利用方法	ラ
2002	2	3	先生宅での食事	キ	2003	2	3	歯医者予約	ラ
2002	2	4	学生証紛失	キ	2003	2	4	アパートの場所	ラ
2002	2	5	レポート締め切り	キ	2003	2	5	暖房停止	キ
2002	2	6	昼ご飯	ラ	2003	2	6	相談室	キ
2002	2	7	定食間違い	ラ	2003	2	7	新しい学生証	キ
2002	2	8	アパートのうるさい住人	ラ	2003	2	8	電話連絡	ラ
2002	2	9	授業が厳しい	キ	2003	2	9	レポートアドバイス	キ
2002	2	10	突き指	ラ	2003	2	10	論文提出	ア
2002	2	11	テーブルマナー講習会	ラ	2003	2	11	授業内容	キ
2002	2	12	レストラン経営	ラ	2003	2	12	テニス肘	ラ
2002	2	13	スーパーの配達サービス	ラ	2003	2	13	子供の英語教育	ラ
2002	2	14	マラソンの死点	ア	2003	2	14	講演会	キ
2002	2	15	英語のクラス分け	キ	2003	2	15	ガムの効果	ア
2002	2	16	歯医者予約	ラ	2003	2	16	米の凶作	ラ
2002	2	17	アリの割合	ア	2003	2	17	研究計画	キ
2002	2	18	ゼミ	キ	2003	2	18	屋上緑化	ア
2002	2	19	研究雑誌執筆募集	ア	2003	2	19	本	ラ
2002	2	20	カウンセリング	ア	2003	2	20	コンピュータにできること	ア

施行年	回	問題番号	トピック	場面	施行年	回	問題番号	トピック	場面
2004	1	1	映画	ラ	2005	1	1	ボランティア活動	キ
2004	1	2	レポートチェック	キ	2005	1	2	レポート提出	キ
2004	1	3	レポートの書き方	ア	2005	1	3	テニス試合	ラ
2004	1	4	本の借り出し	キ	2005	1	4	乗り物台数	キ
2004	1	5	テレビ放送	ア	2005	1	5	小学校授業参加	ラ
2004	1	6	アパートの住人への苦情	ラ	2005	1	6	座席	ラ
2004	1	7	授業内容	キ	2005	1	7	生ゴミ処理	ラ
2004	1	8	クラブ	キ	2005	1	8	心理学行動	ア
2004	1	9	宿題	ア	2005	1	9	集中講義	キ
2004	1	10	ホームステイ	ラ	2005	1	10	講演会準備	キ
2004	1	11	接客	ラ	2005	1	11	レンタル自習室	ラ
2004	1	12	レポート提出方法	ア	2005	1	12	臭覚の順応	ア
2004	1	13	就職活動	キ	2005	1	13	第二外国語選択	キ
2004	1	14	防犯システム	ア	2005	1	14	失敗事例のデータベース	ラ
2004	1	15	公開講座	キ	2005	1	15	望遠鏡	ラ
2004	1	16	経済学	ア	2005	1	16	地熱発電	ア
2004	1	17	てんとう虫	ア	2005	1	17	川遊び	ラ
2004	1	18	産業の発展と災害	ア	2005	1	18	ツユクサ	ア
2004	1	19	時間の感覚の実験	キ	2005	1	19	販売戦略	ア
2004	1	20	悪臭に関する苦情	ラ	2005	1	20	環境問題	ア
2004	2	1	交流会	ラ	2005	2	1	パソコンフェア	キ
2004	2	2	レンタル	ラ	2005	2	2	市民講座	ラ
2004	2	3	研究室	キ	2005	2	3	救急病院	キ
2004	2	4	夕食会	キ	2005	2	4	食品分類	ア
2004	2	5	志望理由	キ	2005	2	5	アンケート調査	キ
2004	2	6	レポートで読む本	キ	2005	2	6	待ち合わせ	キ
2004	2	7	気温	ラ	2005	2	7	コミュニケーショントレーニング	ア
2004	2	8	学割	ラ	2005	2	8	育児時間	ア
2004	2	9	パソコン	ラ	2005	2	9	怪我	キ
2004	2	10	授業選択	キ	2005	2	10	講義内容	ア
2004	2	11	流行	ラ	2005	2	11	ボールペン構造	ア
2004	2	12	ボランティア	ラ	2005	2	12	ワークシェアリング	ア
2004	2	13	都市開発	ラ	2005	2	13	モニター	ラ
2004	2	14	紙	ア	2005	2	14	市民への知らせ方	ラ
2004	2	15	鳥の観察	ア	2005	2	15	報酬とやる気	ア
2004	2	16	障害者に対する心遣い	ラ	2005	2	16	照明と眠り	ラ
2004	2	17	子供の絵	ア	2005	2	17	職業の選び方	キ
2004	2	18	コンビニ経営方式	ア	2005	2	18	課題	ア
2004	2	19	害虫駆除	ア	2005	2	19	魚の特徴	ア
2004	2	20	漫画で教育	ア	2005	2	20	都市政策	ア

施行年	回	問題番号	トピック	場面	施行年	回	問題番号	トピック	場面
2006	1	1	単位取得	キ	2007	1	1	教室予約	キ
2006	1	2	よい文章の書き方	ラ	2007	1	2	ジョギングの持続	ア
2006	1	3	引っ越し	ラ	2007	1	3	犬を飼う理由	ラ
2006	1	4	授業欠席	キ	2007	1	4	友人へ本返却	キ
2006	1	5	読書	ア	2007	1	5	印象に残るスピーチ	ア
2006	1	6	経験	ラ	2007	1	6	テレビ放送サービス	ラ
2006	1	7	プラスチック	ア	2007	1	7	コンサート券の譲渡	ラ
2006	1	8	コミュニケーション教育	ラ	2007	1	8	酸素購入	ラ
2006	1	9	会社経営	ラ	2007	1	9	海外からの病害虫	ア
2006	1	10	体内時計	ア	2007	1	10	ファミリー・フレンドリー企業	ア
2006	1	11	子供の食生活	ラ	2007	1	11	アンケート調査の目的	キ
2006	1	12	授業	キ	2007	1	12	エコツアー	キ
2006	1	13	アスファルト	ラ	2007	1	13	気象条件と歴史	ア
2006	1	14	ボランティア	キ	2007	1	14	司会者の仕事	ラ
2006	1	15	ディベート	ア	2007	1	15	魅力的な会社	ラ
2006	1	16	バリアフリー	ア	2007	1	16	地方都市開発	ア
2006	1	17	動物園の人気の秘密	ラ	2007	1	17	ミーティングに遅れる	キ
2006	1	18	ビジネス	ア	2007	1	18	生活の場所	ア
2006	1	19	ロボット	ラ	2007	1	19	生活習慣・社会性の教育	ア
2006	1	20	植物の花の時期	ア	2007	1	20	ユニバーサルデザイン	ア
2006	2	1	入学ガイダンス	キ	2007	2	1	他人の視点の自分	キ
2006	2	2	高校生の大学生生活体験	キ	2007	2	2	1・2年合同演習	ア
2006	2	3	サクラの交配	ア	2007	2	3	朝食と学力	ラ
2006	2	4	説明書	ラ	2007	2	4	保険証	ラ
2006	2	5	特急電車停車のニュース	ラ	2007	2	5	レポート評価	キ
2006	2	6	野菜の取り方	ア	2007	2	6	介護ロボット	ア
2006	2	7	コンピュータソフトの購入	キ	2007	2	7	運動能力	ア
2006	2	8	ゼミ課題	ア	2007	2	8	七色ボールペン	ラ
2006	2	9	日記	ラ	2007	2	9	寮かアパートか	キ
2006	2	10	アンケート調査	ア	2007	2	10	自然解説者学校	ラ
2006	2	11	動物園と農家の協力関係	ア	2007	2	11	肥満治療薬	ア
2006	2	12	新型トラックの機能	キ	2007	2	12	七色の虹	ア
2006	2	13	林業の木材の再利用	ラ	2007	2	13	新年度の授業	キ
2006	2	14	アリの行動	ア	2007	2	14	エネルギー消費	ア
2006	2	15	ディスカッションの方法	ア	2007	2	15	ネット百科事典	キ
2006	2	16	雨水を貯める貯水槽の取り組み	ラ	2007	2	16	川魚	ラ
2006	2	17	人形を使った育児体験	キ	2007	2	17	売り上げ向上	ア
2006	2	18	車いすでの海外旅行	ラ	2007	2	18	読書の仕方	ラ
2006	2	19	環境ビジネス	ア	2007	2	19	子どもに対する経済教育	ラ
2006	2	20	微生物への石鹼の影響	ア	2007	2	20	技術	ア

施行年	回	問題番号	トピック	場面	施行年	回	問題番号	トピック	場面
2008	1	1	健康診断	キ	2009	1	1	橋	ラ
2008	1	2	個人指導	キ	2009	1	2	サクラ	ア
2008	1	3	少子化と企業	ア	2009	1	3	成績評価	ア
2008	1	4	環境保全	ア	2009	1	4	日本人の食生活	ラ
2008	1	5	水道水	ラ	2009	1	5	心をつかむ話し方	ラ
2008	1	6	外国後学習	キ	2009	1	6	自転車通学	ラ
2008	1	7	江戸時代	ア	2009	1	7	舌と臭い	ア
2008	1	8	老人施設	ア	2009	1	8	見学申し込み	キ
2008	1	9	情報取り入れ	ア	2009	1	9	街路樹	ラ
2008	1	10	家の解体	ラ	2009	1	10	環境税	ア
2008	1	11	血圧測定	ラ	2009	1	11	カラス	ア
2008	1	12	ゼミ申し込み	ア	2009	1	12	中高年の大学入学者	ラ
2008	1	13	大学という場所	ア	2009	1	13	薬の開発	ア
2008	1	14	研究室のイス	キ	2009	1	14	見出し	ラ
2008	1	15	先生との連絡	キ	2009	1	15	水分補給	ア
2008	1	16	携帯電話の普及	ア	2009	1	16	食事の曜日	ラ
2008	1	17	ミツバチ	ア	2009	1	17	若者に注目	ア
2008	1	18	環境問題への取り組み	ラ	2009	1	18	有名人の政治家	ラ
2008	1	19	コンピュータと人の知能	ア	2009	1	19	深夜営業	キ
2008	1	20	宇宙でのチーム	ラ	2009	1	20	猫の野生化	ラ
2008	2	1	発表準備	ア	2009	2	1	授業登録	キ
2008	2	2	本の貸し出し	キ	2009	2	2	考古学の魅力	ラ
2008	2	3	商品の宣伝方法	ア	2009	2	3	指導者の新しい役割	ア
2008	2	4	パソコンの置き場所	ラ	2009	2	4	試験時持ち込み	ア
2008	2	5	機械製品の工夫	ラ	2009	2	5	色	ア
2008	2	6	環境破壊	ア	2009	2	6	食べ物の輸入	ラ
2008	2	7	インターネットの生活	ア	2009	2	7	インターネットタイム	ラ
2008	2	8	施設開放	ラ	2009	2	8	魚の保護	ア
2008	2	9	通訳	ラ	2009	2	9	読書ボランティア	ラ
2008	2	10	ドイツ語の授業	キ	2009	2	10	遺失物届け	キ
2008	2	11	赤ちゃんの反応	ラ	2009	2	11	ほめかた	ア
2008	2	12	葛藤	ア	2009	2	12	森づくりボランティア	ラ
2008	2	13	絵本	ア	2009	2	13	果実の大きさ	ア
2008	2	14	視覚機能	ア	2009	2	14	一般の店と図書館	ア
2008	2	15	文学	ア	2009	2	15	科学を学ぶ	ラ
2008	2	16	アホウドリ	ラ	2009	2	16	老化	ア
2008	2	17	化学物質	ア	2009	2	17	ツバメの行動	ア
2008	2	18	デジタルペン	ア	2009	2	18	エレベーター空間	ア
2008	2	19	方向感覚	ア	2009	2	19	メタ記憶	ア
2008	2	20	クモの餌取り	ア	2009	2	20	植物生育条件	ア

表 4.23 聴解テスト・アイテムのトピック一覧改定後 (2010-2013)

場面のア：アカデミック キ：キャンパス ラ：ライフ

施行年	回	問題番号	トピック	場面	施行年	回	問題番号	トピック	場面
2010	1	13	レポートアドバイス	キ	2011	1	13	レポート	キ
2010	1	14	睡眠	ア	2011	1	14	骨の役割	ア
2010	1	15	わかる	ア	2011	1	15	学校内の空間	ア
2010	1	16	1人暮らし	ラ	2011	1	16	環境問題レポート	キ
2010	1	17	種の発芽	ア	2011	1	17	共同時の人選	ア
2010	1	18	交流範囲	ラ	2011	1	18	美術館発表	ア
2010	1	19	情報化社会	ア	2011	1	19	新分野の心構え	ア
2010	1	20	言葉の浄化作用	ア	2011	1	20	自動車アンケート	ア
2010	1	21	クジラの特徴	ア	2011	1	21	鳥の実験	ア
2010	1	22	地域改革	ラ	2011	1	22	デザインに大切な事	ラ
2010	1	23	筋肉トレーニング	ア	2011	1	23	ノーベル賞	ラ
2010	1	24	将来の進路	キ	2011	1	24	電子メール	ア
2010	1	25	苦情が与える影響	ア	2011	1	25	高齢化社会におけるスーパーマーケット	ア
2010	1	26	目標達成	ア	2011	1	26	子どもの証言	ア
2010	1	27	植物ホルモン	ア	2011	1	27	包装の種類	ア
2010	2	13	プロジェクター貸し出し	キ	2011	2	13	大学祭の説明	キ
2010	2	14	道の駅	ア	2011	2	14	鳥の歌い方	ア
2010	2	15	うそ	ア	2011	2	15	新ペットボトル	ア
2010	2	16	入学願書	キ	2011	2	16	味覚障害	ラ
2010	2	17	小さい魚	ア	2011	2	17	科学研究の役割	ア
2010	2	18	日記	ラ	2011	2	18	スポーツの新しい考え方	ア
2010	2	19	銀行の取り組み	ア	2011	2	19	服を着ること	ラ
2010	2	20	農薬の代わりに虫	ラ	2011	2	20	高齢者の介護	ア
2010	2	21	科学を利用した商売	ア	2011	2	21	ホタルイカ	ア
2010	2	22	授業アンケート	キ	2011	2	22	猫ボランティア	ラ
2010	2	23	工場見学	ラ	2011	2	23	生活資料館	ア
2010	2	24	二酸化炭素と穀物	ア	2011	2	24	通信販売交換	ラ
2010	2	25	景観	ア	2011	2	25	弁当屋の販売戦略	ア
2010	2	26	変動費	ア	2011	2	26	携帯電話使用	ア
2010	2	27	目の錯覚	ア	2011	2	27	健康過小評価	ア

施行年	回	問題番号	トピック	場面	施行年	回	問題番号	トピック	場面
2012	1	13	異文化体験	ラ	2013	1	13	大学院進学	キ
2012	1	14	新入生オリエンテーション	キ	2013	1	14	口頭発表	ア
2012	1	15	議会事務局	ア	2013	1	15	製造業者	ア
2012	1	16	映画待ち合わせ	ラ	2013	1	16	サークル	キ
2012	1	17	ブランド志向	ア	2013	1	17	蝶の飛び方	ア
2012	1	18	本について	キ	2013	1	18	子どもの科学教室	ラ
2012	1	19	小鳥の生態	ア	2013	1	19	スポーツ選手の体の特徴	ア
2012	1	20	料理教室	ラ	2013	1	20	レポートに足りない部分	ア
2012	1	21	おからの処分	ア	2013	1	21	機械の事故	ア
2012	1	22	商品の値段	ア	2013	1	22	ボランティアを長く続けるため	ラ
2012	1	23	言語聴覚士	キ	2013	1	23	木を切ること	ラ
2012	1	24	チンパンジーの道具使用	ア	2013	1	24	会議での黒板の使い方	ラ
2012	1	25	教室で座る位置	ア	2013	1	25	商品のデザイン	ア
2012	1	26	川の改修工事	ア	2013	1	26	樹木の形	ア
2012	1	27	社会的犯罪予防	ア	2013	1	27	言葉の伝わり方	ア
2012	2	13	中学校の先生	ア	2013	2	13	映像メディア論の授業	キ
2012	2	14	ひまわり	ア	2013	2	14	子ども向け展覧会	ア
2012	2	15	ペンギンの呼吸	ア	2013	2	15	マニュアル	ア
2012	2	16	読んでいない本	キ	2013	2	16	文化祭のポスター	キ
2012	2	17	水泳の練習方法	ア	2013	2	17	ことわざ	ア
2012	2	18	害虫を防ぐ方法	キ	2013	2	18	高校生の読書	ラ
2012	2	19	睡眠と記憶	ア	2013	2	19	チンパンジーのえさやり	ア
2012	2	20	スピーチコンテスト原稿	キ	2013	2	20	授業準備	ア
2012	2	21	プラネタリウム	ア	2013	2	21	IT企業評価	キ
2012	2	22	雑誌の読書量	ア	2013	2	22	照明器具	ア
2012	2	23	船を誘導	ラ	2013	2	23	交通の便と観光	ア
2012	2	24	商店街活性化	ア	2013	2	24	記憶	ラ
2012	2	25	健康のための食生活	ア	2013	2	25	辛い刺激	ア
2012	2	26	高齢者支援	ア	2013	2	26	内圧外圧	ア
2012	2	27	ふくろう	ア	2013	2	27	ゲラダヒヒ	ア

⑤人数

表 4. 24 人数の比較 (改定前と改定後)

改定前(2002-2009)			改定後(2010-2013)		
人数	アイテム数	%	人数	アイテム数	%
パブリック	144	45	パブリック	81	68
対人間	170	53	対人間	39	33
グループ	3	1	グループ	0	0
個人内	3	1	個人内	0	0
計	320	100	計	120	100

人数においての変化は改定後に「グループ」と「個人内」の談話がなくなったことである（表 4.24 前頁）。「グループ」は 2004 年と 2005 年に出題された以降はない。「個人内」は 2002 年、2005 年、2007 年に 1 回ずつ出題されていた。改定後は講義のアイテムの増加に伴い「パブリック」が増加している。

⑥目的

表 4.25 で、顕著な差が出ているのが、「目的 1 情報」の「知識」で、改定前の約 2 倍の割合になっている。これは、モノログの講義場面のアイテムが増えたためである。授業形態は増加したが、インターアクションがある授業やゼミのアイテムが増えず、教師から学生への一方向の情報の流れの講義のアイテムだけが增加している。日本の大学の講義は、教師と学生のインターアクションは少なく、ニーズも高いと言える。しかし、日留試の聴解テスト・アイテムのように、短く切り取られたような講義の一部分を理解できたとして、講義の聴解能力が高いと言えるのだろうか。改定後の数値から、講義理解が重要であるというメッセージは伝わるが、講義理解をするための聴解に必要な能力、あるいはテストで測定している能力は、改定前と改定後に変化はないと思われる。

改定後に減少している目的に「問題解決」がある。日本語による問題解決能力は留学生には重要である。アイテム数が限定されているため、何かが増加すれば、何かが減少してしまうことは必然であるが、アカデミック場面で「問題解決」を目的とした内容のアイテムを期待したい。

表 4.25 コミュニケーションの目的の比較（改定前と改定後）

改定前(2002-2009)			改定後(2010-2013)		
目的 1 情報	アイテム数	%	目的 1 情報	アイテム数	%
知識	123	38	知識	76	63
問題解決	79	25	問題解決	8	7
理解	54	17	理解	16	13
行動	29	9	行動	7	6
その他	30	9	その他	11	9
必要	5	2	必要	2	2
計	320	100	計	120	100

改定前(2002-2009)			改定後(2010-2013)		
目的 2 影響	アイテム数	%	目的 2 影響	アイテム数	%
行動変容	63	76	行動変容	19	83
態度変容	12	14	態度変容	2	9
意見変容	8	10	意見変容	2	9
計	83	100	計	23	100

改定前(2002-2009)			改定後(2010-2013)		
目的 3 社会化	アイテム数	%	目的 3 社会化	アイテム数	%
社会システム	8	67	社会システム	1	33
人間形成	4	33	人間形成	2	67
計	12	100	計	3	100

⑦機能

表 4.26 機能の比較（改定前と改定後）

改定前(2002-2009)			改定後(2010-2013)		
情報を授受する機能	アイテム数	%	情報を授受する機能	アイテム数	%
講義をする	85	27	講義をする	67	56
質問する・応える	70	22	質問する・応える	18	15
説明する	36	11	説明する	5	4
話のきっかけを作る・応ずる	26	8	話のきっかけを作る・応ずる	11	9
教える・知る	34	11	教える・知る	2	2
相談する・アドバイスする	21	7	相談する・アドバイスする	6	5
語る	20	6	語る	3	3
知らせる	22	7	知らせる	0	0
教える・学ぶ	3	1	教える・学ぶ	4	3
発表する	1	0	発表する	4	3
報告する	2	1	報告する	0	0
計	320	100	計	120	100

機能については、「情報を授受する機能」「聞き手に働きかける機能」「相手の働きかけに応える機能」「自己の感情を表す機能」「個人・社会化に関する機能」の5つに分類して分析を行ったが、「情報を授受する機能」以外のアイテム数が少ないため、「情報を授受する機能」のみの比較を行った（表 4.26）。「講義をする」の増加が顕著であり、これは講義のアイテムが増加した影響である。一方、2人がインターアクションを行う機能が減少しており、講義のようなモノログが増加したため、ダイアログのアイテムが減少した結果である。「発表する」が、多少増えており、大学生が自ら産出して伝えていく能動的な行動のよいモデルとなるが、数としてはまだ少ない。

⑧情報を得る理由

表 4.27 情報を得る理由の比較（改定前と改定後）

改定前(2002-2009)			改定後(2010-2013)		
情報を得る理由	アイテム数	%	情報を得る理由	アイテム数	%
学術・研究のため	108	34	学術・研究のため	74	62
学習のため	49	15	学習のため	11	9
自分のため	48	15	自分のため	10	8
人間関係のため	37	12	人間関係のため	13	11
大学生活のため	37	12	大学生活のため	4	3
生活のため	41	13	生活のため	8	7
計	320	100	計	120	100

情報を得る理由も、講義場面のアイテムが増加したため、「学術・研究のため」の増加が見られる（表 4.27）。留学生にとって、「学習のため」に聞くアイテムが、もっと必要ではないだろうか。日本の大学で勉強する際の決まりごと、例えば、レポートの書き方、先生への質問の仕方、専門書の使

い方など、多種多様な場面・内容が存在する。アカデミック・リスニングは、講義を聞く能力に代表されるが、講義一辺倒のテスト・アイテムを並べることが、日留試の目的ではないはずである。講義を聞いて理解できるだけでなく、わからないときにどう対処するかといった、学術、学習場面で留学生が抱えている問題の解決につながるようなアイテムの出題を望む。

アカデミック場面が増加したことは、日留試の目的に沿う改定ではあるが、モノローグの講義アイテムだけが増加してしまうと、テスト・アイテムが画一化され、測定能力も一律化してしまう。講義以外の「学術的・学問時」に関わるアイテムの増加も必要である。

4.5 結論と今後の課題

テスト・アイテムを分析した結果、聴解テスト・アイテムの8つの項目、①理解の種類 ②場面 ③形態 ④トピック ⑤人数 ⑥コミュニケーションの目的 ⑦機能 ⑧情報を得る理由 すべてに、談話の場面が影響を与えていることがわかった。これは、言語使用域(場面)によって、言語のスタイル(style)が決まってくるからである。日留試の聴解テスト・アイテムの特徴は、広義なアカデミックから狭義のアカデミックに移行しつつある。では、聴解テストが、講義理解に特化したアイテムだけで構成されるようになるのかということ、そうではないと考えられる。日留試は、言語能力の測定のみではなく、日本で生活し、日本の高等教育機関で学習するために必要な情報を、渡日前に留学希望者に伝え、留学生がスムーズに日本の大学生活に適應できるようにする役割も担っているからである。

4章では、実施された日留試の聴解テスト・アイテムの分析と考察を行い、その特徴をまとめ、聴解テスト作成のための1モデル(テキスト・バンク)を提案した。第1章「1.2 本研究の目的と意義」(p.7)で述べたが、テキスト・バンクをアイテム作成時に活用できるのではないかと考えたからである。日留試のアイテム・ライターは日本語教師や学校教師が多く、彼らには経験と知識が期待されている。彼らは、言語教育のテスト問題、練習問題の作成経験を持ち、作成知識もあり、テストのニーズを理解して適切な場面を想定できる。しかし、アイテム・ライター個々の経験と知識だけに頼るには限界があり、また恣意的になりがちである。そのため、論理的なアイテム作成の枠組みがあれば、アイテムに一定の質を保たせることが可能となる。既にコンピュータによるアイテム作成の試みが行われており¹⁷、さらに項目自体を自然言語処理と談話構造分析を元にコンピュータで作る試みもある(菅井 2003, p.107)。人が発する自然談話をコンピュータがどこまで作れるのかという疑問はあるが、人的リソースだけに頼ってアイテム作成をする時代から、体系的に作成する時代に移行しつつあると言ってよいだろう。

もう一点、アイテムを分析した結果、留学生に必要な日本語能力を測定するために、日本の大学の講義の進め方、大学生活および学習に必要な手続きなどについての情報が必要であると感じられた。以下に、留学生にとって必要な日本語能力の調査・研究結果を、いくつか記載する。

札野・辻村(2003)は、学部の大学教官・教員に留学生が必要な日本語は何かというアンケート調査を行った。調査前に ①基本的な生活のための日本語能力 ②人間関係を豊かにするための日本語能力 ③本来の活動に必要な日本語能力 の3つを検討し、①は大学入学以前に習得すべきものとして除外し、②と③の具体的な能力を調査した。その結果、「②人間関係を豊かにするための日本語能力」

¹⁷ 菅井(2003)は例としてUCLAのWebLAS(<http://weblas.ucla.edu>)と「Web問題作成ツール」(入江2002, <http://www.iwai-h.ed.jp/~irivw/>)を挙げている。

は、日本語でしか（英語が通じない）目的が果たせない場面や、相手とのコミュニケーション・ニーズが高いときに必要であることがわかった。例えば、学内では、事務室、図書館、食堂で働く人達とのやりとりは英語が使えない場合が多く、手続きなどの書類関係も同様である。学外では、アルバイト先での指示、顧客とのトラブル、住まいでの生活上の注意や契約など、金銭が絡み、また文化的背景がないと対応が難しい場面もあるので高い日本語能力が必要である。日本語能力が必要な場面として、学内では、教員に質問したり学生にわからないところを尋ねるなどを挙げているが、学部の違いや教員や学生の英語力、留学生の性格などで必要度にはばらつきが生じる。例えば、理科系は文科系に比べ英語ができればよいという考えが強く、日本人学生への質問は留学生に日本語能力がなくても日本人学生が真剣に対応するので問題はない、留学生の性格がオープンであれば日本語能力のカバーはできるなどが挙げられた。学内学外とも、世間話は必要度が低かった。「③本来の活動に必要な日本語能力」で、上位のほとんどを占めたのが「発表をする、書く、作成する」などの能動的な行動であった。上位10項目までに「課題」と「講義・ゼミ」関連が4つずつ入り、行動達成をするための日本語能力が求められていることがわかった。専門用語に関する項目は、37項目中下位のほうに位置しており、教官たちは、高校で習うような幅広い分野の用語は必要なく、自分の専門分野の用語が必修であるという意見であった。このように、日本の大学の教官・教員が期待する日本語能力にはかなり幅がある。しかし、高い日本語能力が留学生生活を成功させるのではない。コミュニケーションをとろうとする積極的な態度や、専門分野の知識能力が重要であると結んでいる。

日本語教育振興協会基礎日本語教育研究プロジェクトチーム（2001）は、留学生が大学で必要となる日本語能力は課題達成であるとし、大学内での学業、キャンパス生活の場面やトピックから具体的な課題74場面を列挙し、その課題を達成するために23項目のスキル・知識を挙げている。その中で聴解と関わるもの¹⁸は、「講義を聞きながらノートを取る、的確に質問をする、指示に従って簡潔に口頭発表する、論理的に意見を述べる、いろいろなレベル・スタイルの話し方に対応する、他の学習者と学業上でのコミュニケーションを取る、要点を聞く、文脈の中で分からない言葉の意味を推測する、大きな全体の流れを理解する、指示を理解し様式にしたがって書類などを書く、わからない時困った時の対処の方法を見つける」の11項目であった。中居（2003）も、「大学での学習・研究生活」「大学でのキャンパス生活」「地域社会生活・日常場面」で、留学生の言語行動場面25を挙げ、それぞれに言語行動、状況・話題、課題達成能力を列挙している。

上記の調査・研究に共通しているのは、留学生はライフ場面における基本的な生活日本語の能力をすでに習得しているという考え方である。生活場面における基本的な日本語能力測定のアイテムは、初期の日留試に見られたが、徐々に減少している。今後、講義理解に加え、質問、発表、ノート取りをするという能動的な行動の聴解アイテムが増加してくることを期待する。世間話のような「おしゃべり」は必要度が低いようだが、本章の分析結果では、日留試の聴解アイテムには「おしゃべり」が多かった。中には、留学生があまり関心を抱かない話題の「おしゃべり」もあった（例：町に貯水槽を作った理由など）。非常にアイテム数が多かった「おしゃべり」に関しては、トピックを学習・学術に関連したものにすることで、留学生にとっての必要度をあげることができる。ライフ場面では、サバイバル的な場面は必要なく、文化・歴史・生活習慣などの理解が浅いためにトラブルが生じたり、不快感を感じたりする場面を想定し、問題解決をする、円滑なコミュニケーションをする内容を出題することで、波及効果、学習効果を狙うのも1つの方法である。日留試は生活場面もテスト範囲に含

¹⁸ 話すスキルは、それ以前に聴解理解の必要性があると考え、聴解と関わるものとした。

めている。金田（2009）は、外国人が日本で生活する際に遭遇する場面をリスト化しており、このような資料を参考に、アイテム作成をするのもよい。日留試が日本の大学生活への入り口を担っていること、そして、その後の4年間の大学生活も含んでいることを反映したテスト・アイテムが必要である。

日留試は、言語知識を測定する部分的測定（discrete-point item）を排除したコミュニカティブ・テストであり、熟達度テストである。Bachman（1990 邦訳 p.346）は、熟達度テストにおいて真正性は熟達度と同義であると述べている。言語テストの重要な課題として「直接（direct）、運用力（performance）、機能的（functional）、意思伝達重視（communicative）、真正性の高い（authentic）¹⁹」点が注目されてきた（同掲 p.343）。これらの点を多く持っているのが口頭面接や作文であるが、テストとしては採点に時間がかかり、実施も大変であるため、特に大規模テストではあまり行われない。大規模テストが、真正性より効率性を優先するからである（同掲 p.341）。聴解テストでは、「運用力」は直接測定できないが、「機能的、意思伝達重視、真正性の高い」テスト・アイテムを作成することは可能である。「真正性の高い」テスト・アイテムを作成する1つの方法として、目標言語を実際に使用している談話を標本化することが考えられる。言語使用の事象は状況に依存し、固有の価値を持っている（同掲 p.353）。

Bachman の提案する目標言語の実際の使用場面を考えてみよう。例えば、留学生が講義を聞いている際にわからない言葉がでてきたとする。その際に、どのような談話が展開されるだろうか。手を挙げて講師に質問する、隣の学生に質問する、隣の学生の辞書を借りる、ノートに書き留めておいて授業後に親しい友人に質問するなど、複数の談話が予想される。隣の学生との関係（友人、知人、見知らぬ人）、講義中の態度（質問をよくする、一切質問しない）、講師との面識（初対面、以前から知っている）、教室の大きさ、学生の人数、留学生の教室内での位置（座席）、留学生の個人的要素（性別、年齢、母語など）も談話の因子となる。談話理解における発話内能力²⁰や社会言語学的能力も必要である。現実世界での談話は、語用論的な要素を含み、非常に複雑で多面的である。

Bachman（1990）は、テストにおける真正性について、実生活法²¹の支持者・批判者両者の議論をふまえて、テスト受験者とテスト・アイテムの相互作用（interaction）の産物であると結論づけた。テスト・アイテムの分析と、受験者の特性に関わる因子、例えば、言語能力、認知的特性、人格特性、実世界の知識、テスト経験など、の分析が、真正性の高いテスト作成に役立つ。テスト・アイテムの分類や分析は、アイテム・バンク構築の基礎となり、再使用の際には便利なラベルとなる。また、新しくテスト・アイテムを作成する際に、アイディアを与え、テスト全体のバランスを見る際にも有効である（Bachman, 1990 邦訳 p.359）。本研究はテスト・アイテムだけの分析であったが、今後受験者の因子も分析対象とし、双方の相互作用についての研究が必要である。

¹⁹ 言語テキストに使われているタスク、素材、状況が、目標言語を使用する現実でのタスク、素材、状況に近いほど真正性が高いと言われる。（静他 2002, p.47）

²⁰ 発話内能力とはスピーチ行為（speech acts）の概念で説明される。例えば、「ここは寒いよ」という発話には「ヒーターをつけてよ」という依頼であったり「風邪をひいてしまうから子どもは入れないほうがいい」という忠告であったりする。（Bachman, 1990 邦訳 p.103）

²¹ テスト行動がどの程度実生活の行動を映し出すのかの度合いを考慮し、1番の関心事としている方法。

5章 聴解テスト・アイテムのスク립ト分析

この章では、本研究の2つ目の目的「日本留学試験聴解テストが留学生に期待している聴解能力を、アカデミック・リスニングの視点から明らかにする」ことを試みる。

5.1 研究の枠組み

Weir (1993, p.28)や Buck (2001, p.85)は、真正性の高いテキストを、コミュニケーション・テストで使用することを推奨している。McNamara (2000 邦訳 pp.38-40)も、実際の言語使用場面の談話および課題を分析することで、真正性の高いテキストと課題が得られるとしている。しかし、テストを実施する際に、実際の言語使用場面を使用することにはいくつかの不都合が生じる。例えば、実際の談話を切り取って、聴解をさせた場合、録音時に雑音が多く録音マイクの性能により声のボリュームが統一できないなどの音の問題が生じる、会話が非常に冗長になるため聴解テスト時間が長くなることでテストとしての実施性に影響を与える、不正確な言語形態や談話構成が混在することで受験者に混乱を与えるなどが考えられる。また、テストでは、視覚情報はなく、聞き返しができない、解答方法は多肢選択技法を使用することが多いなど、実際の聴解場面では起こりえない状況下で聴解をしなければならない。McNamara (2000 邦訳 pp.40-41)は、聴解テストにおいて、特に応答形式の真正性は、テストの実施上、また受験者全員への公平性を考えると、妥協せざるを得ないと述べている。日留試の聴解テスト・アイテムも、実際の言語使用場面の会話ではなく、アイテム・ライターによって作られたスク립トである。しかしながら、人工的に作成されたアイテムの良い点は、意図的に測定したい言語能力要素を本文スク립トに効率的に入れられることである。そこで、聴解テストのスク립ト分析をすることで日留試が測定したい言語能力が見えてくるのではないかと考えた。受験者が聴解時にどのような能力を使用しているのかを明らかにする研究ではなく、アイテム・ライターが聴解テスト・アイテムのスク립トを作成する際の心理や意図を明確にすることで、日留試がどのような聴解能力（アカデミック・リスニング）を期待しているのかを明らかにすることである。よって、本研究では聴解テスト・アイテムの作成時のスク립トに注目した。スク립トに含まれる言語能力要素を分析することから、日留試が留学生に期待している能力を明らかにするものである。

Bachman & Palmer (1996 邦訳 p.78, p.89)は、言語テストの設計と開発をする際、言語テストの結果を解釈する際、既存のテスト課題に含まれている言語能力を判断する際には、言語能力のすべての範囲の構成要素を意識すべきだと述べている。本研究では、日留試の聴解テスト・アイテムの課題に含まれている言語能力を判断するために、Bachman & Palmer (1996, p.79) (表 5.1 次頁) の言語能力の構成要素を使用する。彼らが言語能力として考えた言語知識の下位項目である ①文法的知識 ②テキストについての知識 (以下：テキスト知識) ③機能的知識 ④社会言語学的知識 の4項目と、⑤方略的能力 の1項目をあわせ5項目とした。そして、より具体的に聴解能力として捉えるために、Richards (1985, pp.198-199) のカンパセショナル・リスニングの下位スキルと、アカデミック・リスニングの下位スキルを5項目に分類した (表 5.2～表 5.6 pp.97-100)¹。Richards のリスニング・スキルについては3章表 3.1 「カンパセショナル・スキルとアカデミック・スキル」 (p.42) にあげているが、ここでもう一度載せる (表 3.1 次頁)。

¹ 日留試の日本語シラバスより、大学内および大学外の生活場面における日本語を測定範囲としているため、カンパセショナル・スキルとアカデミック・スキル両方の下位スキルを使用することにした。

表 5.1 Bachman & Palmer (1996)による言語能力の構成要素

<p>言語知識 (邦訳 p.79)</p> <p>構造的知識</p> <ul style="list-style-type: none"> ・ 文法的知識 (語彙、統語、音韻／書記体系) ・ テキストについての知識 (結束性、修辭的会話的構造) <p>語用論的知識</p> <ul style="list-style-type: none"> ・ 機能的知識 (概念的機能、操作的機能、学習的機能、想像的機能) ・ 社会言語学的知識 (方言や変種、言語使用域、自然な表現や慣用的な表現、文化的指示および比喩的表現) <p>方略的能力 (邦訳 p.82)</p> <p>(目標設定 アセスメント 計画)</p>
--

表3.1 カンバセーショナル・スキルとアカデミック・スキル (Richards, 1985, pp.198-199) ²

【 】:Richardsの英文リストにはなかったが、筆者が説明を加えた)

<p>下位スキル カンバセーショナル・リスニング</p> <ol style="list-style-type: none"> 1. 長さが違う言語の固まりを短時間覚えていられる 2. 目標言語の特徴的な (弁別的な) 音が識別できる 3. 言葉のストレスパターンが認識できる 4. 英語の律動的構造が認識できる 5. 発話内の情報のサインである音の強弱とイントネーション機能が認識できる 6. 音を強く言ったり弱く言ったりする時の言葉を明確に捉えられる 7. 音が縮約された言葉の形が認識できる【短縮形】 8. 言葉の境界線がわかる (分ち書き・熟語など) 9. 目標言語の基本的な語順がわかる 10. 会話のトピックで中心的に使用される語彙がわかる 11. キーワードが見つげられる (トピックや主題が何であるかわかるため) 12. 文脈から言葉の意味が推測できる 13. 内容語・機能語がわかる (会話部分) 14. 主なシンタクスパターンや仕組みがわかる 15. 会話の結束装置がわかる 16. 文法単位の固まりの省略形や文の省略形がわかる【省略とは抜いてしまうこと。つまり文法的にはなくてはならないのだが、抜いてしまう】 17. 文の構成を見つけることができる 18. 構成上、主となる部分、副となる部分が見分けられる
--

² Richards (1985, pp.198-199) (表 3.1, p.42) の下位スキルであるカンバセーショナル・スキルと、アカデミック・スキルを Bachman & Palmer (1996) の5つの構成要素により分類したものである。なお、下位スキルの番号は、Richards のオリジナル文献の番号である。オリジナル文献を参照した際に混乱しないよう、同じ番号を使用した。

19. 文法型・センテンスタイプの違いで意味の違いを表現することがわかる（特別な意味を表すには違う言い方がある）
20. 場面、参加者、ゴールにより発話のコミュニケーション機能がわかる
21. 場面、ゴール、参加者、進行具合などの再現あるいは推測ができる
22. 目的、ゴール、設定、進行具合などがわかるために現実の知識や経験を使うことができる
23. 出来事の描写から結果を予測できる
24. 出来事同士の関係やリンクを推測できる
25. 出来事の原因や結果を推理できる
26. 文字どおりの意味なのか慣用的な言い回しなのか見分けられる
27. 複数の話者たちで会話されている内容を聞きながらトピックや構造上の一貫性を再現したり認識したりできる
28. 談話上の一貫性（文法指示ではなく、推論によってわかる）がわかり、主題、裏付け、与えられた情報、新しい情報、通念、具体例の関係を見つけられる
29. スピーチ速度の違いを処理できる
30. 会話におけるポーズ、間違い、言い直しが処理できる
31. 意味を理解するために表情やパラ言語や他の手がかりを使うことができる
32. 聞き手の目的やゴールによって違う聞き方をすることができる
33. 理解しているのかしていないのかということを言語的あるいは非言語的に知らせることができる

下位スキル アカデミック・リスニング（講義を聴く）

1. 講義の目的や領域がわかる能力
2. 講義のトピックがわかり、トピックの展開についていける能力
3. ディスコースの中の複数の構成要素の関係がわかる能力（主題・通念・仮説・補佐的アイデア・例など）
4. 講義談話の特徴的なディスコースマーカーがわかる能力（接続詞・副詞・切り出しの言葉・常套句など）
5. 関係を推測する能力（理由・結果・まとめなど）
6. トピックに関するキーワードを見つけることができる能力
7. 文脈から言葉の意味を推測する能力
8. 結束性のマーカーを見つけることができる能力
9. 大切な情報のある部分にかかるイントネーションの機能がわかる能力（ピッチ・音量・ペース・音の高さ）
10. 話題に対しての話者の態度を見極める能力
11. 様々な話の形式についていける能力（対話・テープのような聴覚のみ・テレビのような視聴覚）
12. アクセントとスピードが違う話についていける能力
13. 話の様々なスタイルを知っていること（フォーマル・会話・読み聞かせ・無計画）
14. 言語使用の違いを知っていること（書き言葉・話し言葉）
15. 関連事項がわかる能力（冗談・余談・漫談）
16. 強調のためのノンバーバルサインや態度がわかる能力
17. 教室内の約束事を知っていること（ターンテイキング・質問を明確になど）
18. 教授や学生の話の機能を認識する能力（警告・提案・推薦・アドバイス・説明など）

Richards(1985)が示した上述のリスニング・スキルから、本研究の目的を勘案し、以下の5つを省いた。①基礎的な言語知識であり、留学生の言語知識として要求されていないと判断したもの ②音に関連しているため、スクリプト分析ではわからないもの ③視覚に関連しているため日留試の聴解テストでは測定できないもの ④記憶に関わるためテストでは測定しないもの ⑤日留試の聴解テストの本文には含まれないと判断したもの。①に関しては日留試が言語知識を個別的には測定しないテストであることも省いた理由の1つである。以下が、①～⑤で省いたリスニング・スキルである。

・省いたリスニング・スキル

①基礎的な文法や語彙など

カンバセーショナル・リスニングの下位スキル

8. 言葉の境界線がわかる (分かち書き・熟語など)
9. 目標言語の基本的な語順がわかる
10. 会話のトピックで中心的に使用される語彙がわかる
13. 内容語・機能語がわかる
14. 主なシンタクスパターンや仕組みがわかる
16. 文法単位の固まりの省略形や文の省略形がわかる
17. 文の構成を見つけることができる
18. 構成上、主となる部分、副となる部分が見分けられる

アカデミック・リスニングの下位スキル

13. 話の様々なスタイルを知っていること (フォーマル・会話・読み聞かせ・無計画)

②音の理解

カンバセーショナル・リスニングの下位スキル

2. 目標言語の特徴的な (弁別的な) 音が識別できる
3. 言葉のストレスパターンが認識できる
4. 日本語の律動的構造が認識できる
5. 発話内の情報のサインである音の強弱とイントネーション機能が認識できる
6. 音を強く言ったり弱く言ったりする時の言葉を明確に捉えられる
7. 音が縮約された言葉の形が認識できる
29. スピーチ速度の違いを処理できる

アカデミック・リスニングの下位スキル

9. 大切な情報のある部分にかかるイントネーションの機能がわかる能力 (ピッチ・音量・ペース・音の高さ)
12. アクセントとスピードが違う話についていける能力

③視覚が必要なこと

カンバセーショナル・リスニングの下位スキル

31. 意味を理解するために表情やパラ言語や他の手がかりを使うことができる

アカデミック・リスニングの下位スキル

16. 強調のためのノンバーバルサインや態度がわかる能力

④記憶に関わること

カンバセーショナル・リスニングの下位スキル

1. 長さが違う言語の固まりを短時間覚えていられる

⑤聴解テストであるため

カンバセーショナル・リスニングの下位スキル

30. 会話におけるポーズ, 間違い, 言い直しが処理できる
31. 意味を理解するために表情やパラ言語や他の手がかりを使うことができる
33. 理解しているのかしていないのかということと言語的あるいは非言語的に知らせることができる

アカデミック・リスニングの下位スキル

11. 様々な話の形式についていける能力(対話・テープのような聴覚のみ・テレビのような視聴覚)
14. 言語使用の違いを知っていること(書き言葉・話し言葉)

以下の節で、Bachman & Palmer (1996, p.79) (表 5.1 「Bachman & Palmer による言語能力の構成要素」 p.93 参照) の①文法的知識 ②テキスト知識 ③機能的知識 ④社会言語学的知識 ⑤方略的能力 に上記の項目を削除した Richards (1985, pp.198-199) のリスニング・スキルを分類する (表 5.2～表 5.6 pp.97-100)。カンバセーショナル・スキルには㊸、アカデミック・スキルには㊹をつけ、Richards (1985, pp.198-199) (表 3.1 pp.42-43, pp.93-94) のオリジナル文献の番号をつけた。なお、表の左側には各知識・能力の記述を試みた。

5.1.1 文法的知識

文法的知識とは、「個々の発話または文がどのように構成されているか (Bachman & Palmer 1996, 邦訳 p.79)」ということで、比較的独立した能力であり (Bachman 1990, 邦訳 p.100)、語彙、統語、音韻、書記体系の知識からなる。日留試は、文法項目を個々に測定するテストではないこと、カンバセーショナル・リスニングの下位スキルは基礎的文法であり留学生は既に習得している日本語能力であることを鑑み、本研究の対象から外した(・省いたリスニング・スキル pp.95-96)。残ったアカデミック・リスニングの下位スキル「㊹4. 講義談話の特徴的なディスコース・マーカーがわかる能力 (接続詞・副詞・切り出しの言葉・常套句など)」だけを分析対象とする (表 5.2)。

Richards は英語を対象としているため、日本語のディスコース・マーカーとの違いがある。英語では接続詞「but, because」や、副詞「when」は文頭に来るが、日本語では同じ機能を持つ接続助詞の「～が、～と、～なら」、格助詞「～から」、形式名詞「～とき」は文末にあり独立しておらず、マーカーとして判断が困難なため、接続詞の「でも、実は、もちろん」など文頭にきたもののみをマークした。

表 5.2 スクリプト分析におけるリスニング・スキルの枠組み①文法的知識

文法的知識の記述	Richard の下位スキル
発話または文の構成要素の理解	㊹4. 講義談話の特徴的なディスコース・マーカーがわかる能力 (接続詞・副詞・切り出しの言葉・常套句など)

5.1.2 テキスト知識

テキスト知識は「テキストをなすために発話または文がどのように構成されているか (Bachman & Palmer 1996 邦訳 p.79)」で、結束性および修辭的会話的構造の知識である。Bachman (1990 邦訳 p.101)

は結束性と修辭的會話的構造の知識を以下のように説明している。結束性は新旧情報の配列規約、指示、置き換え、省略、接続、語彙的結束性といった意味論上の諸関係を明確に特徴づけるための方法である。修辭的會話的構造の知識は、テキストの概念的構造全体と関係しており、テキストの持つ効果とも関係している。叙述、記述、比較、分類、過程分析などの展開の一般的方法や、パラグラフ内での情報配置（トピックセンテンス、第一の主要な支持文、第2的支持文、結論、話題轉換のための文など）などの多くの規約を含む。また、會話的言語使用として注目させる、トピックの取り上げ、トピックの展開、會話の維持などの規約を含む。

以上を踏まえ、表 5.3 のように分類した。日本語の結束性とは、前後の文章理解が必要な文法的結束性の指示詞「こ・そ・あ」や、形式名詞などの理解である (㊦24. ㊦8.)。

修辭的會話的構造の知識に「トピック」に関する記述があるため、「㊦11. ㊦27. ㊦27. ㊦2. ㊦6.」はテキスト知識とした。トピックに関連した能力では、「トピックが確認でき、展開についていける能力」として4つ (㊦11. ㊦1. ㊦2. ㊦6.)、テキスト全体の理解が必要である「それぞれの文章の役割と関係をテキスト全体から見つける能力」として4つを分類した (㊦27. ㊦28. ㊦3. ㊦5.)。

テキスト理解を補助する能力として、「未知語彙の説明やヒントが、テキストのどの部分で与えられたかを推測できる能力」もテキスト知識とした (㊦12. ㊦7.)。

表 5.3 スクリプト分析におけるリスニング・スキルの枠組み②テキスト知識

テキスト知識の記述	Richard の下位スキル
前後の文章理解が必要な文法的結束性の指示詞「こ・そ・あ」や、形式名詞などの理解	㊦24. 出来事同士の関係やリンクを推測できる
	㊦15. 會話の結束装置がわかる
	㊦8. 結束性のマーカーを見つけることができる能力
それぞれの文章の役割と関係をテキスト全体から見つける能力	㊦27. 複数の話者たちで會話されている内容を聞きながらトピックや構造上の一貫性を再現したり認識したりできる
	㊦28. 談話上の一貫性（文法指示ではなく、推論によってわかる）がわかり、主題、裏付け、与えられた情報、新しい情報、通念、具体例の関係を見つけられる
	㊦3. ディスコースの中の複数の構成要素の関係がわかる能力（主題・通念・仮説・補佐的アイデア・例など）
	㊦5. 関係を推測する能力（理由・結果・まとめなど）
トピックが確認でき、展開についていける能力	㊦11. キーワードが見つけられる（トピックや主題が何であるかわかるため）
	㊦1. 講義の目的や領域がわかる能力
	㊦2. 講義のトピックがわかり、トピックの展開についていける能力
	㊦6. トピックに関するキーワードを見つけることができる能力
未知語彙の説明やヒントが、テキストのどの部分で与えられたかを推測できる能力	㊦12. 文脈から言葉の意味が推測できる
	㊦7. 文脈から言葉の意味を推測する能力

5.1.3 機能的知識

機能的知識は、「発話または文およびテキストが、言語使用のコミュニケーション上の目標にどのように関係づけられるか (Bachman & Palmer 1996 邦訳 p.79)」である。彼らは、機能的知識に概念的機能、操作的機能、学習的機能、創造的機能の4つを挙げているが、Bachman (1990)にそれぞれの詳しい説明があるので以下にまとめる。概念的機能は、自分が意図したことを相手に伝えるための表現で、命題を表現したり、知識や感情を交換するときに使用される。操作的機能は、自分の回りの世界に影響を与えるもので、道具的機能としての命令、指示、申し出、約束などがあり、制御的機能として規律や規則などを含み、相互作用的機能として人間関係の形成や維持などがある。学習的機能は、自分の回りの世界に関する知識を広めるもので、指導、学習、問題解決、意図的記憶などである。創造的機能は、感性から生じた感覚、例えば美意識のために環境を創造したり広めたりすることである。娯楽、芸術、読書などの他、冗談や空想、隠語も含まれる。(Bachman, 1990 邦訳 pp.106-108)

機能的知識では「場面、登場人物に関わる機能の理解」として3つ (㊸20. ㊸21. ㊸18.)、「談話の機能に自らの経験による既知概念をプラスして予測や推理をする能力」として2つ (㊸23. ㊸25.) を分類した。

表 5.4 スクリプト分析におけるリスニング・スキルの枠組み③機能的知識

機能的知識の記述	Richard の下位スキル
場面、登場人物に関わる機能の理解	㊸20. 場面、参加者、ゴールにより発話のコミュニケーション機能がわかる ㊸18. 教授や学生の話の機能を認識する能力 (警告・提案・推薦・アドバイス・説明など) ³
	㊸21. 場面、ゴール、参加者、進行具合などの再現あるいは推測ができる
談話の機能に自らの経験による既知概念をプラスして予測や推理をする能力	㊸23. 出来事の描写から結果を予測できる
	㊸25. 出来事の原因や結果を推理できる

5.1.4 社会言語学的知識

社会言語学的知識は、「発話または文およびテキストが、言語使用の設定の特性にどのように関係づけられるか」であり、方言や変種、言語使用域、自然な表現や慣用的な表現、文化的指示および比喩的表現の4つの知識からなる (Bachman & Palmer 1996 邦訳 p.79)。日留試の聴解テスト・アイテムでは、東京共通語を使用しているため、方言や変種は適用しない。それ以外を表 5.5 (次頁) にまとめた。

言語使用域は、言語をどこで、どのような目的で使用するかによって特徴づけられる。言語使用域には、コミュニケーションをとる2人の間の人間関係が表現方法に影響することも含む (Bachman, 1990 邦訳 pp.109-112)。日留試の言語使用域は、留学生の大学生活場面が主であるため、「日本の大学」という場の理解」と考えた (㊸15. ㊸17.)。同様に場面と人間関係の理解が必要なものとして「場面や立場、目的、媒体などにより文体が異なる、位相を含めた言語形態の違いの理解」(㊸19. ㊸14.) を分類した。

自然な表現や慣用的な表現とは〈母語話者らしい〉談話をさす (Bachman, 1990 邦訳 p.112)。〈母語話

³ ㊸18.は場面、参加者が教授と学生に特化しているだけで㊸20.と同じ能力とした。

者らしい」とは、言葉や表現の機能と密接している。その表現でどのような気持ちを伝えたいのか、その表現にどのような含みを持たせるのかということである。例えば、

A:どこ行くの?

B:ちよつと。

この「ちよつと」は、「ちよつと駅まで行ってくる」であれば辞書の意味では「その行動が軽い気持ちで行われるさま」(大辞泉)であるが、上記の会話に続いて、

A:どこ行くのよ。

B:ちよつと。

となった場合は、「詳しいことは聞かないでほしい」という嘆願か、「いちいちうるさい」という文句にもなる。これは辞書には載っておらず、「ちよつと」という表現が日本語ではどのように使用されているのかを知らなければならない。さらに Bachman (1990 邦訳 pp.112-114)は、文化的指示および比喩的表現とは、特定の意味を持つ語彙が存在することであり、その意味を知らなければ理解できないと説明している。例えば、「A:哲学ってなんかよくわかんないんだよ。 B:富士山の樹海みたいなもんだね。」という会話における「富士山の樹海」が「迷い込んだら出られない」という意味で使用されていることは、日本人の成人であればほぼ理解できる。しかし、「富士山の樹海」が、場所を示す語彙という認識しかなければこの会話を理解できない。このような文化背景がなければわからない指示は日留試では出題されないと判断し、今回は適用しない。一方、「バケツをひっくり返したような雨」は語彙そのものに特殊な意味は含まれないため、比喩的表現として適用し、「慣用的表現の理解」として分類した(㊦26.)。

表 5.5 スクリプト分析におけるリスニング・スキルの枠組み④社会言語学的知識

社会言語学的知識の記述	Richard の下位スキル
「日本の大学」という場の理解	㊦15. 関連事項がわかる能力(冗談・余談・漫談)
	㊦17. 教室内での約束事を知っていること(ターンテイキング・質問を明確になど)
場面や立場、目的、媒体などにより文体が異なる、位相を含めた言語形態の違いの理解	㊦19. 文法型・センテンスタイプの違いで意味の違いを表現することがわかる(特別な意味を表すには違う言い方がある)
慣用的表現の理解	㊦26. 文字どおりの意味なのか慣用句的な言い回しなのか見分けられる

5.1.5 方略的能力

方略的能力は、メタ認知的構成要素であり、認知的活動だけではなく言語使用における認知的管理機能も含む高度な過程である。その使用領域は、目標設定(何を行うかを決定)、アセスメント(何が必要か、何を処理するのか、うまくやれたか)、計画(持っているものをどう使うか)である(Bachman & Palmer 1996 邦訳 pp.81-82)。また、Bachman (1990 邦訳 pp.118)は方略的能力の機能を、談話の言語情報と、予測や経験、既知の知識などを利用した状況判断の結果とが一致するかどうかの検討をして言語能力を最大限有効に使えるものとしている。

方略的能力には「言語知識を補う能力」として2つ(㊦22. ㊦10.)「戦略的聞き方」として(㊦32.)を分類した(表 5.6 次頁)。

表 5.6 スクリプト分析におけるリスニング・スキルの枠組み⑤方略的能力

方略的能力の記述	Richard の下位スキル
言語知識を補う能力	②22. 目的、ゴール、設定、進行具合などがわかるために現実の知識や経験を使うことができる ⑦10. 話題に対しての話者の態度を見極める能力
戦略的聞き方	②32. 聞き手の目的やゴールによって違う聞き方をすることができる

5.2 分析対象と分析方法

本研究のスクリプト分析の対象は、2002 年度から 2006 年度までに公開された聴解テスト・アイテム 200 である。初年度から 5 年間のテスト・アイテムを研究対象とした理由は、日留試が掲げたアカデミック・ジャパニーズが一番反映されていると考えたからである。

4 章でも述べたが、再度、図 4.1 (p.52) を用いて日留試の聴解テスト・アイテムの構成を説明する。④「本文」の前に、③「設問」がある。設問には 2 つの機能があり、談話の場の説明の①「場面・状況説明」と、本文を聞く目的である②「プリタスク (pre-task)」である。⑤が「ポスト・クエスチョン (post question)」、⑥が「選択肢」である。

大学の職員が、入学式のあと次の日の予定について話しています。・・・ ①場面・状況説明 留学生は、午前のオリエンテーションが終わったらどうしますか。・・・ ②プリタスク	} ③設問
それでは、明日の予定ですが、午前中の全体オリエンテーションは、10時30分から12時までで、この講堂で、全員に共通することについて説明します。えー、午後は1時半から、学科ごとにオリエンテーションを行います。この講堂ではなく、1号館で行いますから、気をつけてください。なお、留学生は、午前中のオリエンテーション終了後に生活ガイダンスを行います。健康保険の手続きに必要な書類も渡しますので、午前の部が終わり次第、留学生センターへ行ってください。時間は15分くらいです。そのあと、1時半からの予定は全員同じです。	
留学生は、午前のオリエンテーションが終わったらどうしますか。・・・ ⑤ポスト・クエスチョン 1. 学科のオリエンテーションを受けに講堂へ行く。 2. 生活ガイダンスを受けに留学生センターへ行く。 3. 学科のオリエンテーションの時間に生活ガイダンスを受ける。 4. 健康保険の手続きのために1号館へ行く。	} ⑥選択肢

図 4.1 日留試聴解テスト・アイテムの構成 (2006 年度 (平成 18 年度) 2 回目聴解テスト 1 番)

分析手順は、①日留試の聴解テスト・アイテムのスクリプトに、解答に関係がある情報だと筆者が判断した部分をマーキングする ②マーキングした部分の下位スキルを特定する ③解答を記述する ④記述した解答と選択肢を照らし合わせ、正答を選ぶ の4段階である⁴。

まず、①のマーキングについて説明する。スクリプトを読みながら、「談話内での特定人物」「解答部分に直接関係がある情報」「解答かもしれない情報」「解答と関係ない情報」「直接の解答情報ではないが、解答を得るための補佐的な情報」「ディスコース・マーカ―」の6つに分け、色で選別していった(図5.1)。「ディスコース・マーカ―」をマーキングする理由は、スクリプト内の接続標識としてテキストの一貫性を高める働きがあり(JACET, 2005, p.122)、聴解をする際、結束性の高いテキスト構造が内容理解をしやすくしているからである(藤森 1998, p.62)。

大学の職員が、入学式のあと次の日の予定について話しています。留学生は、午前のオリエンテーションが終わったらどうしますか。

それでは、明日の予定ですが、午前中の全体オリエンテーションは、10時30分から12時までで、この講堂で、全員に共通することについて説明します。えー、午後は1時半から、学科ごとにオリエンテーションを行います。この講堂ではなく、1号館で行いますから、気をつけてください。なお、留学生は、午前中のオリエンテーション終了後に生活ガイダンスを行います。健康保険の手続きに必要な書類も渡しますので、午前の部が終わり次第、留学生センターへ行ってください。時間は15分くらいです。そのあと、1時半からの予定は全員同じです。

	解答部分に直接関係ある情報
	解答かもしれない情報
	解答と関係ない情報
	直接の解答ではないが、解答を得るための補佐的な情報
	ディスコース・マーカ―
	特定人物

図5.1 スクリプト・マーキング例(2006年度(平成18年度)2回目聴解テスト1番)

次に、②の下位スキル特定について説明する。本文のスクリプト・マーキング後(図5.1の例)、マーキングした部分を、Richards のリスニング・スキル(表5.2~表5.6 pp.96-100)で分析した。Bachman & Palmer が挙げた方略的能力は、判断に関わる能力であるため、解答と判断した際の決め手となったキーワードの種類を分析することにした。

実際の日留試聴解テストでは、4つの選択肢から解答を選ぶのだが、本研究では、③解答を記述した。

⁴ 筆者が行ったスクリプト・マーキングの信頼性をみるために、全200アイテムから、その20%である40アイテムを無作為に抽出し、7人の評価者に同様にマーキングをしてもらった。その後、スピアマン・ブラウン予言公式(Brown 1996, pp.231-234)にて評価者間信頼係数を求めた結果、0.81であったため、筆者が行ったスクリプト・マーキングの信頼性は確保されたと判断した。

⁵ 2人の談話の場合、ブリタスクからどちらの意見・提案・考えなどについて聞くのが判明する。解答を言うであろう人物の発話により注意をはらうと考え、「特定人物」とした場合には、マーキングをした。独話の場合は、誰のための情報を聞かなければならないかがブリタスクで判明するため、その人を「特定人物」とした。

以下に理由を述べる。McNamara (2000 邦訳 p.41) は、テストで要求された基準に達しなかった受験者が、実際の場合では違う結果となりうる可能性があるとし、テストの真正性と実用性の間に、回避できない妥協があると述べ、その1つに多肢選択法をあげている。多肢選択法のテストでは、解答がわからなくとも選択肢から正答を選ぶ可能性があり、逆に解答が得られていても、選択肢から選べないことも起こりうる。多肢選択法の聴解テストでは、本文内容の意味理解ができなくとも、音の記憶で正答を選ぶことを避けるため、選択肢の文言を変えている場合が多い。このため、選択肢の理解も求められているのが現状である。本研究では、課題達成能力として、選択肢から正答を選ぶ能力は省き、設問は状況・場面説明と課題提示であるため、本文にのみ含まれる言語要素の分析をした。

スクリプト・マーキングをした部分をどのように分析したかを図5.1「スクリプト・マーキング例 (2006年度 (平成18年度) 2回目聴解テスト1番)」(前頁) を用いて説明する。マーキングした部分の下位スキルが Richards の下位スキルの何にあたるかを選び、解答だと判断した場合には「解答」とした。談話を最後まで聞かなければ「解答」とわからないため、保留とした「解答」にはクエスチョン・マークをつけた。「⑩18. 教授や学生の話の機能を認識する能力 (警告・提案・推薦・アドバイス・説明など)」は、言語学的機能も付加した。(表5.7 次頁)

表 5.7 下位スキル分析例 (2006 年度 (平成 18 年度) 2 回目聴解テスト 1 番)

㊦ : カンパセーショナル・スキル ㊧ : アカデミック・スキル

数字 : Richard の下位スキル番号 (表 5.2~表 5.6 pp.96-100)

マーキング部分	文法的知識	テキスト知識	機能的知識	社会言語学的知識	方略的能力
大学の職員が、入学式のあと次の日の予定について話しています。留学生は、午前のオリエンテーションが終わったらどうしますか。					
それでは	㊧4				
明日の予定ですが		㊧2			
午前中の全体オリエンテーションは、10時30分から12時までで、この講堂で、全員に共通することについて説明します			㊦20 ㊧18 行動		
午後は1時半から、学科ごとにオリエンテーションを行います			㊦20 ㊧18 行動		解答?
この講堂ではなく		㊧8			
1号館で行いますから、気をつけてください			㊦20 ㊧18 注意		
なお	㊧4				
留学生は午前中のオリエンテーション終了後に生活ガイダンスを行います			㊦20 ㊧18 行動		解答と判断
午前の部が終わり次第、留学生センターへ行ってください			㊦20 ㊧18 命令		
そのあと	㊧4				
1時半からの予定は全員同じです		㊧8			

5.3 結果の分析と考察⁶

聴解テスト・アイテムのスキプト分析の結果、スキプトに含まれていたリスニング・スキルは表5.8である。以下、4つの知識と方略的能力の5項目別に分析・考察する。

表 5.8 聴解テスト・アイテム本文のスキプト・マーキング部分にみられたリスニング下位スキル (2002-2006)

*B & P : Bachman & Palmer の言語能力 *「数」はアイテム数
 *カ : Richards カンバセーションナル・リスニングの下位スキル
 ア : Richards アカデミック・リスニングの下位スキル

B & P	言語能力の記述	リスニングの下位スキル	数	%
文法的知識	発話または文の構成要素の理解	ア4. 講義談話の特徴的なディスコース・マーカ―がわかる能力	164	82
テキスト知識	前後の文章理解が必要な文法的結束性の指示詞「こ・そ・あ」や、形式名詞などの理解	カ24. 出来事同士の関係やリンクを推測できる	0	0
		ア8. 結束性のマーカ―を見つけることができる能力	111	55.5
	それぞれの文章の役割と関係をテキスト全体から見つける能力	カ27. 複数の話者たちで会話されている内容を聞きながらトピックや構造上の一貫性を再現したり認識したりできる	0	0
		カ28. 談話上の一貫性（文法指示ではなく、推論によってわかる）がわかり、主題、裏付け、与えられた情報、新しい情報、通念、具体例の関係を見つけれられる	15	7.5
		ア3. ディスコースの中の複数の構成要素の関係がわかる能力（主題・通念・仮説・補佐的アイデア・例など）	39	19.5
	トピックが確認でき、展開についていける能力	ア5. 関係を推測する能力（理由・結果・まとめなど）	42	21
		カ11. キーワードが見つけれられる（トピックや主題が何であるかわかるため）	21	10.5
		ア1. 講義の目的や領域がわかる能力	2	1
		ア2. 講義のトピックがわかり、トピックの展開についていける能力	164	82
	未知語彙の説明やヒントが、テキストのどの部分で与えられたかを推測できる能力	ア6. トピックに関するキーワードを見つめることができる能力	12	6
カ12. 文脈から言葉の意味が推測できる		1	0.5	
機能的知識	場面、登場人物に関わる機能の理解	ア7. 文脈から言葉の意味を推測する能力	1	0.5
		カ20. 場面、参加者、ゴールにより発話のコミュニケーション機能がわかる ア18. 教授や学生の話の機能を認識する能力（警告・提案・推薦・アドバイス・説明など）	200	100
	談話の機能に自らの経験による既知概念をプラスして予測や推理をする能力	カ21. 場面、ゴール、参加者、進行具合などの再現あるいは推測ができる	0	0
		カ23. 出来事の描写から結果を予測できる カ25. 出来事の原因や結果を推理できる	1	0.5
社会言語学的知識	「日本の大学」という場の理解	ア15. 関連事項がわかる能力（冗談・余談・漫談）	0	0
		ア17. 教室内での約束事を知っていること（ターンテイキング・質問を明確になど）	0	0
	場面や立場、目的、媒体などにより文体が異なる、位相を含めた言語形態の違いの理解	カ19. 文法型・センテンスタイプの違いで意味の違いを表現することがわかる（特別な意味を表すには違う言い方がある）	0	0
方略的能力	慣用的表現の理解	カ26. 文字どおりの意味なのか慣用的な言い回しなのか見分けられる	31	15.5
		カ22. 目的、ゴール、設定、進行具合などがわかるために現実の知識や経験を使うことができる	0	0
	戦略的聞き方	ア10. 話題に対しての話者の態度を見極める能力 カ32. 聞き手の目的やゴールによって違う聞き方をする事ができる	0	0

⁶ この結果の分析と考察は、足立（2008a, pp.153-156）を参考に分析の見直しと新たな考察を加えたものである。

5.3.1 文法的知識

表 5.8.1 聴解テスト・アイテム本文のスク립ト・マーキング部分にみられたリスニング下位スキル
(2002-2006) 文法的知識

*B & P : Bachman & Palmer の言語能力 *「数」はアイテム数

*㊦ : Richards カンバセーションナル・リスニングの下位スキル

㊦ : Richards アカデミック・リスニングの下位スキル

B & P	言語能力の記述	リスニングの下位スキル	数	%
文法的知識	発話または文の構成要素の理解	㊦4. 講義談話の特徴的なディスコース・マーカーがわかる能力	164	82

「㊦4. 講義談話の特徴的なディスコース・マーカーがわかる能力」は、聴解テスト・アイテムの82%に見られた。ディスコース・マーカーは、次にくる情報の予測をたてることができ、理解を助ける。ほとんどの談話にディスコース・マーカーが入っていることから、ディスコース・マーカーの重要性がうかがえる。

日留試は統合的なテストであるため、文法的知識を個別に測る形式のテスト・アイテムはない。しかし、最低限生活ができるという場面ではないため、ある程度の語彙・統語・音韻などの知識を受験者に期待していると言ってよいだろう。日留試に対し、語彙や文法の能力を個別に測っていない、あるいは基準がないことを問題視している研究者たちもいるので、以下に研究者名とその考えをあげる。日留試発足当時から、日留試の研究を続けている「アカデミック・ジャパニーズ」研究会の報告書から、日留試の文法的知識について述べている部分である（門倉編 2003b, 2005）。

嶋田（2003, p.23）：学習活動の基盤となる語彙は重要である。

鈴木（2003, p.43）：1つの語彙が既知か未知かで正否が問われる場合もあるので語彙の基準が必要ではないか。

村上（2003, p.57）：語彙の不足を多くの留学生が指摘しているが、日留試が語彙を個別に直接問わない影響はあるのだろうか。

堀井（2003, p.119）：TOEFL と比較し、日留試にないものを改善点として漢字・語彙の知識もある程度問うべきではないか。

門倉（2003a, p.125）：漢字・語彙の理解力を見る問題がない。

言語能力を支える一要因として、文法的知識が重要であるという認識が、日本語教育の現場にはある。また、このような意見が出る背景には、旧日能試が「日本語能力試験出題基準」（国際交流基金・財団法人日本国際教育協会 1994）を公開しており、その中に、各レベルで必要とされる語彙や文法リストが記載されていたことも一因として考えられる。旧日能試を利用する受験者や関係者は、1級でどのような語彙、表現、文法が期待されているのかを知ることができた。しかし、日留試は具体的な項目や基準を示していないため、教育現場では、どの程度の日本語能力が期待されているかの具体的なイメージがしにくい。日留試は、目標基準準拠テストではないため、具体的な目標を提示する必要はないという考え方もあるが、日本語教育に関わる教師および学習者から、文法や語彙リストのニーズは高いと言ってよいだろう。日留試に必要な語彙や文法は、日本の大学で使用されている語彙・漢字・表現・文法などの調査をもとに決められていると考えられる。しかし、現在公表されていないため、国立国語研究所の話し言葉のコーパス開発「KOTONOHA」⁷や、筑波大学の『筑波ウェブコーパス』（Tsukuba Web Corpus:

⁷ 2014年3月8日 http://www.ninjal.ac.jp/corpus_center/kotonoha.html

TWC)⁸のようなコーパス研究を参考に独自にリスト化することが必要である。既に、公開された日留試を対象に語彙研究も行われており（山本 2005, pp.110-126）、日留試で測定対象となっている語彙が、リスト化される日も近いだろう。

5.3.2 テキスト知識

表 5.8.2 聴解テスト・アイテム本文のスク립ト・マーキング部分にみられたリスニング下位スキル
(2002-2006) テキスト知識

*B & P : Bachman & Palmer の言語能力 *「数」はアイテム数

*㊦ : Richards カンパセーションナル・リスニングの下位スキル

㊧ : Richards アカデミック・リスニングの下位スキル

B & P	言語能力の記述	リスニングの下位スキル	数	%
テキスト 能力知識	前後の文章理解が必要な文法的結束性の指示詞「こ・そ・あ」や、形式名詞などの理解	㊦24. 出来事同士の関係やリンクを推測できる	0	0
		㊦8. 結束性のマーカーを見つけることができる能力	111	55.5
	それぞれの文章の役割と関係をテキスト全体から見つける能力	㊦27. 複数の話者たちで会話されている内容を聞きながらトピックや構造上の一貫性を再現したり認識したりできる	0	0
		㊦28. 談話上の一貫性（文法指示ではなく、推論によってわかる）がわかり、主題、裏付け、与えられた情報、新しい情報、通念、具体例の関係を見つけられる	15	7.5
		㊦3. ディスコースの中の複数の構成要素の関係がわかる能力（主題・通念・仮説・補佐的アイデア・例など）	39	19.5
		㊦5. 関係を推測する能力（理由・結果・まとめなど）	42	21
	トピックが確認でき、展開についていける能力	㊦11. キーワードが見つけれられる（トピックや主題が何であるかわかるため）	21	10.5
		㊦1. 講義の目的や領域がわかる能力	2	1
		㊦2. 講義のトピックがわかり、トピックの展開についていける能力	164	82
		㊦6. トピックに関するキーワードを見つけることができる能力	12	6
	未知語彙の説明やヒントが、テキストのどの部分で与えられたかを推測できる能力	㊦12. 文脈から言葉の意味が推測できる	1	0.5
		㊦7. 文脈から言葉の意味を推測する能力	1	0.5

テキスト知識では、「トピックが確認でき、展開についていける能力」として「㊦2. 講義のトピックがわかり、トピックの展開についていける能力」が82%と多く見られた。本研究では、「講義のトピックがわかる」と「トピックの展開についていける」能力のどちらか一方がスク립トに見られた場合も

⁸ 2014年3月8日 <http://corpus.tsukuba.ac.jp/>

数えている。他に「㊦11. キーワードが見つけれられる（トピックや主題が何であるかわかるため）」は10.5%、「㊦6. トピックに関するキーワードを見つけることができる能力」は6%だった。カンバセーションル場面の㊦11.は2002年と2003年のキャンパス場面とライブ場面で多く出題されているが、それ以降はほとんどない。㊦6.は、ほとんどが講義場面であり、2002年以降、全20アイテム中17番目以降で定期的に出題されている。「㊦1. 講義の目的や領域がわかる能力」は、1%と少なかったが、講義の目的や領域は重要であるがゆえ、実際の談話内で明確に言語化することも多いと考えられテスト・アイテムとしにくいのだろう。「領域」に関しては専門的な語彙知識が必要となる可能性もあり、受験者が、入学後に進む領域により有利不利が発生することも考えられる。そして、かなりまとまった長さの講義内容を聞く必要があり、現時点での日留試では測定しにくい能力なのかもしれない。

トピック関連の理解能力はスクリプトの本文内だけではなく設問部分にも見られたために、トピック理解がどの時点で求められているのかを以下に例を挙げ、考察を試みる。

【例1 2006年2回目 5番】

男子学生と女子学生がニュースについて話をしています。
この女子学生がこのニュースについておかしいと思っていることは、どんなことですか。

例1の場面・状況説明では、トピックがまだわからないある1つのニュースか、放送形態の1つである「ニュース」がトピックであるともとれる。プリタスクで「このニュースについて」と言っていることから、何か1つのニュース内容を取り上げていることが、ここでわかる。しかし、トピックは不明であり、トピック提示は本文内でされる。

【例2 2006年2回目 7番】

コンピュータ室で、男子学生が係の人と話しています。この男子学生は、これから何をしますか。

例2は、トピックが示されていないが、「コンピュータ」に関わる事柄がトピックになるのではないかと推測される。

【例3 2006年2回目 11番】

先生が、授業で、動物園と農家が行っている協力関係について話しています。
先生は、この動物園の動物は、どのような野菜を食べていると言っていますか。

例3では、状況説明でトピックについて述べている。場面・状況説明を聞いた時点で、「動物園と農家間の協力関係」という構図がイメージでき、いくつかのキーワードが予測できる。プリタスクから、野菜を提供するという協力関係があることがわかる。例2では、設問時にトピック理解を求められている。

【例4 2003年1回目 16番】

女性が講演会で環境問題について話しています。この女性はどのような点を変えていかなければならないと言っていますか。

最近では環境問題が新聞で取り上げられることも多く、このことについて人々の関心も高まっています。・・・(省略)・・・

例4は、場面・状況説明で既に提示されたトピックが本文の冒頭でも提示されている(グレー部分)。

以上のように、トピックの提示位置と内容には一貫性が見られない。これは日留試のシラバス(2章「2.4.3 日本留学試験」p.31)にある測定言語下位技能

- ・言語下位技能
 1. 情報の全体の流れをとらえる
 2. 情報の全体を、ある判断や評価をしながらとらえる
 3. 特定の情報を抽出してとらえる
 4. 推測しながら情報をとらえる
 5. 予測しながら情報をとらえる
 6. その他

が、本文だけではなく設問にも求められているということである。設問に課題を提示する機能を持たせているため、設問時に使用する技能も測定せざるをえないということだろう。また、1アイテムの時間も限られており、設問でトピックを提示し、本文内ではトピック展開のみを扱うということも考えられる。

次に、多く見られたのは「前後の文章理解が必要な文法的結束性の指示詞「こ・そ・あ」や、形式名詞などの理解」の「⑦8. 結束性のマーカーを見つけることができる能力」で、55.5%だった。「@24. 出来事同士の関係やリンクを推測できる」は、今回はなかったが今より長い談話になれば見られるだろう。指示詞「こ・そ・あ」は前後の文脈の理解が問われ、旧日能試の読解でよく出題されていたのは旧知のことである。聴解は、読解のように戻って読むことができないため、短期記憶の保持が必要であり、指示詞の位置が、指示する言葉と離れているとわかりにくくなる。そして、指示詞を多用すると、全体の内容が曖昧になり注意が必要である。特に、指示語が現場指示に終始してしまうと、何を指しているのかがわからなくなる可能性が高い。聴解テストでは視覚情報がないため、受験者がどのような場面と登場人物の位置を想像するかはわからない。アイテム・ライターがアイテム作成時に想像した場面と異なった場면을想像した受験者が不利にならないように注意が必要である。この問題を解決するには、テスト用紙に写真を記載する、コンピュータ画面による場面の視覚情報の提示などが考えられるが、視覚情報は余剰情報も多く含まれ、一概に視覚情報を与えれば問題が解決するとも言えない。アイテム作成時に現場指示をできる限り避ける、文脈指示も多用しないという対応が現実的であろう。ここに、実際には見えていない場面の聴解をしなければならない問題点が浮上するが、テストは現実とは違い擬似的なものであるということをアイテム作成時に忘れてはいけない。

3番目に多かったのは「それぞれの文章の役割と関係をテキスト全体から見つける能力」の「⑦5. 関係を推測する能力(理由・結果・まとめなど)」で、21%である。「⑦3. ディスコースの中の複数の構成要素の関係がわかる能力(主題・通念・仮説・補佐的アイデア・例など)」は19.5%だった。講義を聞く

上で必要なリスニング・スキルであるが、日留試の聴解テスト・アイテムのスキ립トの長さは300～400字程度であり、談話内容が大きく発展展開する余裕はない。「関係の推測」や「複数の構成要素」の理解のためには、より長い談話構成が必要である。「㊦28. 談話上の一貫性がわかり、主題、裏付け、与えられた情報、新しい情報、通念、具体例の関係を見つけられる」は7.5%であった。比較的困難度が低いと予想されるテスト・アイテム配列の前半6番目ぐらいまでに、定期的に出題されている。談話構造の理解は、基本的スキルということだろう。「㊦27. 複数の話者たちで会話されている内容を聞きながらトピックや構造上の一貫性を再現したり認識したりできる」は、「3人以上の会話において、自分以外の人と話しているトピックを聞きながら自分が会話に参加するタイミングを図る」、また、「構造上の一貫性を認識し、それを自分の発話に取り込んで会話をつなげていく」という能力である。日留試で3人の会話が出題されたのは、2002年から2005年までの200アイテム中、3アイテムのみである。声音だけで2人以上の話者を特定することが難しく、出題しにくいのかも知れない。複数参加の会話場面は録画を見せることで解決できると思われるが、機器の使用など実施上の課題となる。聴解は聞いて理解するだけでなく、自ら会話に参加するための能力でもある。一方的な情報を聞くアイテムだけではなく、双方向、多方向の会話アイテムが増えることが望ましいが、会話テストの範疇での出題となるのだろうか。

今回対象とした聴解テスト・アイテムには、ほとんど見られなかった「未知語彙の説明やヒントが、テキストのどの部分で与えられたかを推測できる能力」について考察する。「㊦12. 文脈から言葉の意味が推測できる」「㊦7. 文脈から言葉の意味を推測する能力」は、それぞれ0.5%であった。実際の聴解では、文脈を理解する知識と方略的能力を併用し、聴解を補助していると思われる。

テキスト知識は、長い談話、構造化された談話から、必要な情報やキーワードを選び出す能力である。現在の日留試のテスト形式は、トピック提示が多い、一問一答である、長さが一定であるなど画一化されており、テキスト知識の測定には限界があると考えられる。

5.3.3 機能的知識

表 5.8.3 聴解テスト・アイテム本文のスキ립ト・マーキング部分にみられたリスニング下位スキル
(2002-2006) 機能的知識

*B & P : Bachman & Palmer の言語能力 *「数」はアイテム数

*㊦ : Richards カンバセーションナル・リスニングの下位スキル

㊦ : Richards アカデミック・リスニングの下位スキル

B & P	言語能力の記述	リスニングの下位スキル	数	%
機能的知識	場面、登場人物に関わる機能の理解	㊦20. 場面、参加者、ゴールにより発話のコミュニケーション機能がわかる	200	100
		㊦18. 教授や学生の話の機能を認識する能力（警告・提案・推薦・アドバイス・説明など）		
	談話の機能に自らの経験による既知概念をプラスして予測や推理をする能力	㊦21. 場面、ゴール、参加者、進行具合などの再現あるいは推測ができる	0	0
		㊦23. 出来事の描写から結果を予測できる	1	0.5
	㊦25. 出来事の原因や結果を推理できる	1	0.5	

「場面、登場人物に関わる機能の理解」の「㊦20. 場面、参加者、ゴールにより発話のコミュニケーション機能がわかる」「㊦18. 教授や学生の話の機能を認識する能力（警告・提案・推薦・アドバイス・説明

など)」は、すべてのアイテムに見られた。それは解答に必要な情報内に必ず、機能を示す表現や機能語が入っていたためである(例5)。

【例5 2006年2回目 1番】

大学の職員が、入学式のあと次の日の予定について話しています。
留学生は、午前のオリエンテーションが終わったらどうしますか。

・・・(省略)・・・1号館で行いますから、気をつけてください。【注意】・・・(省略)・・・
・・・(省略)・・・午前の部が終わり次第、留学生センターへ行ってください。【命令】・・・
(省略)・・・

旧日能試の『日本語能力試験出題基準』(1994,p.147)には、「高度な機能語の類いを数多く習得して正確に豊かに使うことが、語彙や漢字を増やすこととともに、1・2級レベルの学習者にとってはきわめて重要な課題である」と、機能語の重要性についての記述がある。日留試では機能語という表現知識ではなく、発話のコミュニケーション機能が認識できる能力を測定すると謳っているが、そのためには「～てください」「～なさい」が、「命令・指示」の機能を持つ言語表現形であるという、機能語の知識が必要である。

同様の能力と分類した「㊦21. 場面、ゴール、参加者、進行具合などの再現あるいは推測ができる」は0%だった。実際の会話では、最後まで聞かずともわかることも多い。それはトピック、文脈、会話相手と共有している情報、非言語情報などから推測できるからである。しかし、テストという性質上、解答に対してフォローができないため、言語化されていない部分を推測して答えよ、という形式は出題できないと考えられ、今回の結果となったのだろう。

「談話の機能に自らの経験による既知概念をプラスして予測や推理をする能力」の「㊦23. 出来事の描写から結果を予測できる」と、「㊦25. 出来事の原因や結果を推理できる」は、0.5%だった。実際の会話では、最後まで聞かずともわかることも多い。それはトピック、文脈、会話相手と共有している情報、非言語情報などから推測できるからである。しかし、テストという性質上、特に大規模テストでは解答に対してフォローができないため、言語化されていない部分を推測して答えよ、という形式は出題できない。

5.3.4 社会言語学的知識

表 5.8.4 聴解テスト・アイテム本文のスク립ト・マーキング部分にみられたリスニング下位スキル
(2002-2006) 社会言語学的知識

*B & P: Bachman & Palmer の言語能力 *「数」はアイテム数

*㊦: Richards カンパセーションナル・リスニングの下位スキル

㊧: Richards アカデミック・リスニングの下位スキル

B & P	言語能力の記述	リスニングの下位スキル	数	%
社会言語学的知識	「日本の大学」という場の理解	㊦15. 関連事項がわかる能力 (冗談・余談・漫談)	0	0
		㊦17. 教室内での約束事を知っていること (ターンテイキング・質問を明確になど)	0	0
	場面や立場、目的、媒体などにより文体が異なる、位相を含めた言語形態の違いの理解	㊧19. 文法型・センテンスタイプの違いで意味の違いを表現することがわかる (特別な意味を表すには違う言い方がある)	0	0
	慣用的表現の理解	㊧26. 文字どおりの意味なのか慣用句的な言い回しなのか見分けられる	31	15.5

「慣用的表現の理解」の「㊧26. 文字どおりの意味なのか慣用句的な言い回しなのか見分けられる」は15.5%であり、そのほとんどがキャンパス場面とライブ場面だった。講義では不特定多数に向けて話すため、なるべく明瞭に話すということから明確な言語化を心がけているからかもしれない。今回は見られなかった「場面や立場、目的、媒体などにより文体が異なる、位相を含めた言語形態の違いの理解」の「㊧19. 文法型・センテンスタイプの違いで意味の違いを表現することがわかる (特別な意味を表すには違う言い方がある)」にも同様のことが言える。

「㊦15. 関連事項がわかる能力 (冗談・余談・漫談)」「㊦17. 教室内での約束事を知っていること (ターンテイキング (発言の順番)・質問を明確になど)」は、アカデミック・リスニングの分類であったため、「日本の大学」という場の理解」とした。大学で実際の講義を聞く際には、必要な能力であると考えられるが、1～2分の談話の長さでは、余談などを入れる余裕がない。「ターンテイキング」は談話に参加している人の立場や立ち位置が関係する。社会的なルールとしての順番もあるが、それだけではなく、人間関係も影響を及ぼす。聴解理解の助けにはなるが、テストとして扱うスキルとしては、会話のほうが適切である。

今回、ほぼすべてのテキストにマーキングしたアイテムもあった。このようにアイテム内の情報の密度が濃い場合、受験者は集中力を切らすことができない。テストとはいえ、終始集中力を持続することは負担が大きい。テストに限らず外国語の聴解は、常に緊張の連続だと思われがちだが、実際にはどこに注目すべきかを、場面・状況・トピック・ニーズから判断している。例えば、初級レベルの人は、コンビニのレジで、店員の最後の言葉に注目する。そこで値段を言うと、推測しているからである。予想外の言語行動があった場合も、集中する。例えば、店員が購入した冷凍食品を手になにかを言った場合、「氷が必要か」と聞いているのではないかと推測し、「氷」が聞こえないかと集中する。実際の会話では、言語のみでなく、非言語も情報を与えており、レジの店員の言う文言すべてを聞き逃すまいと集中しているのではない。「袋は必要か」「支払いはカードか」などを聞き逃しても、買い物という課題は達成できるからである。聞くべき部分を聞き、あとはさほど注意せず聞き流すという緩急のある聞き方

は、ネイティブも通常行なっている。このような聞き方をするには、一連の流れのある談話が必要である。TOEFL iBT では、2006 年まで CBT にあった短い会話はずし、ある程度の長さのある談話に変更された。これは、留学生の言語能力に、何が必要とされているかを示唆している。日留試の聴解テストでも、講義の一部として話の切り出しから結論までを通し、さらに冗談などを交えながら話すという流れを聞かせ、そこから何が聞き出せるかというテスト形式の検討が必要である。

5.3.5 方略的能力

表 5.8.5 聴解テスト・アイテム本文のスク립ト・マーキング部分にみられたリスニング下位スキル
(2002-2006) 方略的能力

*B & P : Bachman & Palmer の言語能力 *「数」はアイテム数

*㊦ : Richards カンバセーション・リスニングの下位スキル

㊧ : Richards アカデミック・リスニングの下位スキル

B & P	言語能力の記述	リスニングの下位スキル	数	%
方略的能力	言語知識を補う能力	㊦22. 目的、ゴール、設定、進行具合などがわかるために現実の知識や経験を使うことができる	0	0
		㊧10. 話題に対しての話者の態度を見極める能力	0	0
	戦略的聞き方	㊦32. 聞き手の目的やゴールによって違う聞き方をする事ができる	0	0

今回のスク립ト分析の結果から、日留試では方略的能力を測定対象としていないことがわかった。「言語知識を補う能力」に「㊦22. 目的、ゴール、設定、進行具合などがわかるために現実の知識や経験を使うことができる」「㊧10. 話題に対しての話者の態度を見極める能力」を分類した。㊦22.は、テスト・アイテム作成時には言語能力以外の能力を測定すると考えられ、受験者個々の既知の情報量や経験に左右されないように注意するが、実際の聴解時にはネイティブも活用する能力である。リスニング・スキルとしては重要であるが、テストでは測定対象とはならない。㊧10. は、円滑なコミュニケーションを行うために、大切な能力である。特に、上級レベルの日本語を活用している留学生に対して、日本人は無意識に、社会言語学的な判断ができる、あるいは態度から真意が読み取れると判断しがちである。言語表現で明確に現わさなくとも理解する能力を持っていて当然であるという「以心伝心」の通念が社会の中で確立している。しかし、このような理解・判断力は相手との関係、場の状況、非言語的要素などが複合的に必要となるため、聴解テストでの測定は困難であると考えられる。「㊦32. 聞き手の目的やゴールによって違う聞き方をする事ができる」は、受験者が「戦略的に」使用する能力である。日留試にはプリタスクがあり、聞く目的やゴールが与えられるため、受験者はこのスキルを使用している可能性は高い。しかし、テスト・スク립トからはその能力を狙って測定するという事は見えなかった。方略的能力を補助的に使用するか否かには個人差があり、言語能力の測定値に大きく影響を与える測定項目としては扱いにくい。方略的能力の使用については、6章と7章で触れる。

5.4 選択肢

表 5.9 本文から得た解答と選択肢の関係 (2002-2006)

解答と選択肢の関係		数	%
直接	本文から得た情報と正答がほぼ同じ	75	37.5
本文中語彙	本文中にある語彙に言い換えられている	16	8
語彙	本文中にない語彙に言い換えられている	50	25
統語	統語的言い換えが行われている	10	5
混合	語彙、統語的な言い換えが混合している	21	10.5
方略	直接解答が得られず推測・判断が必要	28	14
計		200	100

本研究のスクリプト分析では、スクリプトの情報から自由解答し⁹、その後、それを選択肢と照らし合わせて正答を選んだ。自由解答と選択肢にはある関係が見られた(表 5.9)。自由解答とほぼ同じ内容の選択肢は4割弱であり、選択肢に語彙、統語等の言い換えが 48.5%あった。聴解テストの場合、意味がわからずとも音の記憶だけで選択肢を選んでしまう場合があり、それを避けるために言い換えがされている。しかし、言い換えの語彙や表現がわからなかった場合、本文で解答を得ていても答えられないということが起こってしまう。これは、本文を聞く際に様々な能力を活用して解答を得ても、選択肢で構造的知識を問われ、それがわかなければ答えられないこととなる。また、選択肢を言い換えることは、正答選択肢と、本文から得た解答が微妙に違う、解釈によっては誤答選択肢も正答となるなどの弊害もある。言語の聴解テストで、多肢選択法を使用する際の課題である。

選択肢の分析の結果、日留試の聴解テストの選択肢は、ある程度の長さを持っていることがわかった。名詞のみで作成されている選択肢(「名詞の名詞」も含む)は、全体の 14%であり、残りの 86%が文章レベルであった。日留試は、4つの選択肢すべてを、聞いて答えなければならない。1つ1つを聞いた時点で、誤答と判断できたものは記憶から削除できるが、解答になりそうだと判断した選択肢は、最後の選択肢を聞き終えるまで記憶にとどめておかなければならない。Hughes (2003 邦訳 p.175)は、聴解テストの選択肢は、記憶のテストにならないよう短くしなければならないと述べている。選択肢の記憶力測定を避けるには選択肢を記述しておく方法がある。留学生に求められる日本語能力は、日留試の日本語シラバスにあるように統合的な能力であり、聞くスキルのみ、読むスキルのみと厳密に独立して測定する必要はない。TOEFL や新日能試の聴解テストは、選択肢を文字情報で提示している¹⁰。日留試の聴解テストで何を測定するのか、測定する目的は何かを見据え、選択肢の記述についても検討をすべきではないだろうか。島田 (2006, p.23) は、聴解テストの選択肢を音声提示と文字提示とで与え、その結果を報告している。

1. 文字提示より音声提示のほうが難易度(困難度)は高い
2. 選択肢の長さによる違いはない
3. 漢字圏(中国、韓国)と非漢字圏による違いはない
4. 音声提示と文字提示の間に識別力の差はない

この結果から、日留試の聴解選択肢が文字提示であっても問題は無いと考えられる。「2.長さによる違い

⁹ 全 200 アイテムの 20%である 40 アイテムを無作為に抽出し(スクリプト・マーキングの信頼性のために行なったものと同一)、7人の評価者に自由解答をしてもらい、筆者の解答と比較し、一致度を求めた。その結果、78%であったため、筆者の解答の信頼性は確保されたと判断した。

¹⁰ 課題理解、ポイント理解、統合理解の一部の選択肢は文字表記されている。

はない」という結果についてであるが、島田 (2006, p.22) が使用した聴解アイテムは 16 アイテムであり、内 2 アイテムが 30 文字以下 (4 選択肢の提示文字数¹¹)、最長で 50 文字であった。日留試の選択肢は、例えば、2011 年 2 回目の 15 の聴解テスト・アイテムを見てみると、最低が 16 文字、次が 37 文字、40-60 文字が 4 アイテム、61-80 文字が 5 アイテム、81-91 文字が 4 アイテムであった。島田の使用したアイテムと比較すると、かなり長いことがわかる。福田 (2003) は、作動記憶と目標言語の聴解テストとの関係を、学習期間が長ければ長いほど個人差が影響してくるため、初級者には認知の負担を減らすよう短かい文に押さえるべきであり、上級者でも特別な訓練が必要であると述べている。日留試の聴解テストでは、講義内容の概要をまとめるようなこともあるだろう。おのずと、選択肢が長くなる傾向は避けられない。日留試のテスト形式のような練習問題を多くつんだ学習者、あるいは第一言語の作動記憶容量が高い学習者に有利にならないよう、選択肢の文字提示の検討が望まれる。

Hughes (2003 邦訳 p.83) は、多肢選択法のテスト形式で難しいのは、よい選択肢を書くことであると述べており、多肢選択のテスト・アイテムを作る時間と労力は、多大で熟練が必要なことから、教育機関の内部テストには適さないとしている。しかし、大規模テストでは、多肢選択法の形式が多く使用されている (3 章表 3.2 「聴解テスト・アイテムの設問、選択肢」 p.48 を参照)。多肢選択のテストは、採点が客観的にでき、採点の信頼性が高い。また、解答に時間がかからないため、テスト・アイテム数を多く出題できるという利点もある。出題数が多ければ、テストの信頼性は増す。このことから、大規模テストで多肢選択の形式をとることは妥当であろう。日留試では本文理解のみでなく、選択肢の聴解と理解も求められていることを常に念頭におき、テスト・アイテムを作成する際、課題と本文がいかによく作られていても、選択肢でアイテムの質を落としかねないということを忘れてはならない。

5.5 結論と今後の課題

5 章では、研究の目的 2 の、日本留学試験聴解テストが測定している聴解能力を、アカデミック・リスニングの視点から明らかにした。ここで、もう一度分析項目である Bachman & Palmer (1996, 邦訳 p.79) の言語知識の表 (表 5.1 p.93) を載せる。

表 5.1 Bachman & Palmer (1996) による言語能力の構成要素

<p>言語知識 (邦訳 p.79)</p> <p>構造的知識</p> <ul style="list-style-type: none"> ・ 文法的知識 (語彙、統語、音韻/書記体系) ・ テキストについての知識 (結束性、修辭的会話的構造) <p>語用論的知識</p> <ul style="list-style-type: none"> ・ 機能的知識 (概念的機能、操作的機能、学習的機能、想像的機能) ・ 社会言語学的知識 (方言や変種、言語使用域、自然な表現や慣用的な表現、文化的指示および比喩的表現) <p>方略的能力 (邦訳 p.82)</p> <p>(目標設定 アセスメント 計画)</p>
--

¹¹ 例えば、1.海 2.山 3.デパート 4.川 は提示文字数 7 である。

日留試のスク립ト分析から、日留試の聴解テストでは、語彙や文法を個別的に測定してはいないが、文法的知識のディスコース・マーカ、テキスト知識のキーワード、トピック展開、ディスコースの構成要素の関係という構造的知識が必要であることがわかった。日本語読解では、ディスコース・マーカに注目して読む重要性は既に知られているが、聴解ではまだ注目されていない。読解や作文（記述）は、文字テキスト内のディスコース・マーカにより、読み手に方向性やその後続くテキストの推測を促す。今後、聴解でも、ディスコース・マーカが注目されることを期待したい。

日本語における機能語は語彙・文法として学習するため、学習段階では文法的知識である。しかし、語彙や表現の機能的意味、あるいは多義性が、場面や文脈理解と連結した場合、語用論的知識が問われることになる。日本で留学生活をおくる際、語用論的知識をもっとも必要とするのが聴解ではないだろうか。語用論的知識をテストで測るためには、どのような形式が考えられるのかの議論が必要である。

社会言語知識の「方言や変種」を、日留試で測定することはない。それは、日本語教育が国内においても海外においても、東京共通語によって教材が作成され、使用されているからである。国内においては、地方の日本語教育機関で方言を使用した教材もあるが、まだ少数である。「言語使用域」は、留学生にとっては大学生と教員・職員で形成された集団での言語使用がほとんどである。今回分析したアイテムには、内容的に大学生には必要性が低い場面・状況もあり、どの程度「大学生」に特化した内容に絞られるのか、今後もテスト・アイテムの分析は必要である。「文化的指示および比喩的表現」は、今後出題される可能性はあるが、大学生活において使用頻度が高くないこと、あるいは留学後に習得すればよいとも考えられる。

「方略的能力」は、スク립ト分析からは判断できないため、6章と7章で論述する。

5章の結果から、日留試では「文法的知識」が必要であることがわかったが、日本語学習者はどのような点を重点的に学習しているのだろうか。尹（2001, p.64）は、日本語初級レベルの中国人大学生に、教室外の聴解練習について、アンケート調査した。その結果、「目標をたてる」「テープを繰り返して聴く」が多かった。その他に、「日本の映画を見る、ラジオや歌を聞く、日本人と話す」など、なるべく日本語ネイティブの日本語に触れるようにしていることがわかった。留学生が挙げた12項目のなかで、構造的知識を強化する項目は、「語彙や文型を覚える」の1つだけであった。学習者にとって、文法的知識は聴解能力の最重要項目ではないのである。三國・小森・近藤（2005）は、語彙知識が読解と聴解の内容理解に対してどの程度影響を与えるかを調べた。その結果、読解では47%だが聴解では30%であり、聴解では文法的知識以外の能力の重要性が高いと述べている。

選択肢の分析の結果、選択肢では語彙や表現の言い換えがされ、最終的に文法知識を問うテスト・アイテムが多いと言える。聴解テストの本文作成時に選択肢も考慮し、音の一致だけで正答が選ばれないよう、本文内の語彙や表現をそのまま使用した正答を避ける傾向がある。これが、選択肢で語彙の言い換えが多用される原因である。Heaton（1988）は、聴解テストを作成する場合、実際の話文字化することを勧めている。聞き手の記憶に残るのは、何を言っていたかであり、それをどのような単語でどのような構文で言っていたかではない（Heaton, 1988, pp.121-123）。大規模テストでは、多肢選択肢法に頼らざるを得ない現実ではあるが、適切な選択肢を作成することの重要性を感じる。菅井（2006a, pp.293-294）は、聴解テストの本文と選択肢の関係を、アイディア・ユニット（IU）を用いて分析した。その結果、魅力的な誤答肢は、「本文のIUと対応する数が多く、推測や伸張による対応が多い」とし、正答以外の誤答肢の作成方法を提案している。聴解テスト・アイテムの本文および選択肢の作成について、更なる研究がすすみ、よりよい選択肢が作られることを望む。

2章表 2.4 「聴解テスト・アイテムの比較：旧日能試、日留試、日能試」(p.33) でまとめたように、旧日能試と日留試の聴解テスト・アイテムの形式は、ほとんど変わらない。宮城(2006, p.57) は、旧日能試の聴解テストで測定対象となっている「聴解行動能力」は、ストラテジー能力が重視されていると考えた。ストラテジー能力を測定するには、短いテキストでパターン化・タイプ化された問題形式では難しく、無理が生じると述べている。日留試の聴解テスト・アイテムは、一問一答、固定化されたテスト形式(設問・本文・ポスト・クエスチョン・音声による選択肢)、本文の長さは1～2分とパターン化されている。例えば、一問一答を一問多答にすることで、1アイテムの長さが自由になり、アイテム数をより多く確保でき、真正性が高まる可能性もある。テストで測定できる能力は、テスト・アイテムの多種多様な形式により広がるのではないかと考えられる。福田(2005, pp.79-81) は、目標言語の学習者の作動記憶の処理効率、習熟度が高いほどよくなるため、語彙などの保持率が上がりテスト得点が高くなるとしている。また、習熟度の高い学習者の聴解過程では、文法理解、語の認知などのボトムアップ処理は自動化しており、トップダウン型の処理を用いる傾向が強いため、場面や文脈に依存した推測理解など、高次な処理が必要なテスト・アイテムが適していると述べている。この研究結果からも、日留試の聴解テストには、ある程度の長さのテキストで、一問多答の形式が有効だと考えられる。

さらに、テスト形式について、他の可能性も考えてみる。最初から課題¹²がわかっている聴解と、聞きながら自ら課題を発見していく聴解が、両方とも日常生活のなかに存在することから、場面と聞く目的によって、テスト・アイテムの形式を変えということも1つの方法である。

平尾(1999) は、留学生に講義を聞いてもらい、アンケート、インタビュー調査、講義理解評価を行った。その結果、講義聴解の理解を助けるのは、図表や文字資料の利用と、講義内容の背景知識と、語彙力であったと述べている。また、講義内容に対するテストでは、必要情報のメモと、講義内容理解度が影響していた。留学生は、わからない語として、「専門用語：講義内で使用された概念や知識を表す専門性の高い語」、「学問用語：講義の分野研究において中等教育レベルで紹介される概念や知識を表す語」、「一般用語：一般的な概念や知識を表す未習の語」の3種類を挙げた。この中で、講義ノートに書き留めるべき語は専門用語であり、それは聞き手が加工処理してノートに書いてはいけない語である。講義理解は、一般の会話聴解とは違い、講義内容について解説・考察・議論ができるような生産的な聞き方であるとしている。このようなテスト形式の例がIELTSにある。IELTSの聴解テスト・アイテムには、聴覚情報と同時に視覚情報を与えられる聴解テスト・アイテムがあり、解答方法も多肢選択と記述がある。記述は書きこんだりメモを取るなど、実際に聴解をしながらする行動を組み合わせている。

日留試の聴読解は、図表や文字資料が提示され、談話を聞くという点で、講義聴解に近い。現時点では聴読解のアイテムの中には、選択肢としての視覚情報という扱いのものもあり、講義理解のための補助としての役割が薄い場合もあるが、講義理解の形式には適していると思われる。視覚情報を見ながら聴解するという形式だけではなく、ある程度の長さの読解資料を読ませ、その後それについての講義を聴覚情報として聞かせるという形式にすると、学生が事前に本を読んでから講義に出席するという実際の場面に近づけることができる。聴解スクリプトを長くすることで、聴解スクリプト内のキーワード(専門用語)をメモをしておかないとその後の会話の理解がしにくい、メモに頼らないと解答が難しくなるという形式も実際の講義聴解に近い。解答方法として要約するなどの記述(全文記述や穴埋め記述)も考えられる。また、ある講義を聞き、それに対して自分の考えを書く、考察するなどの記述に繋げるこ

¹² 本論文では「プリタスク」と限定して使用する以外は「課題」という語彙を使用する。しかし、参考図書で「タスク」を使用している場合、その内容からの考察時には「タスク」を使用する。

ともできる。伊東他（2009, p40-41）は、読解テストにおける設問で、「語彙・構造」を問うものや答えが明確に文中にあるものなどは低次レベルであり、「論の把握」「主題・論旨の把握」「推論」などを高次レベルとし、低次レベルのような基礎的能力は、高次レベルの設問で内包されるため、アカデミック・リーディングでは高次レベルの設問をすべきであると述べている。アカデミック・リスニングにも、同様のことが言える。ある程度の長さを持った談話を用いて高次レベルの設問を使用し、推測や予測によって理解するという能力測定も必要である。以上のように、聴解テストでの解答形式だけではなく、複合スキルを測定することについても、今後、様々な試みがされ、より真正性の高いテスト形式の開発が望まれる。

本研究では、Richards のリスニング・スキルを参考にしたが、変わりつつある大学の授業形式に伴い、新たなリスニング・スキルが必要となる可能性もある。今後、日本語特有のリスニング・スキルも視野におき、留学生だけではなく、日本人大学生にとっても、講義理解を助ける研究として、リスニング・スキル研究の重要性は高い。

6章 テスト時の聴解行動（設問あり）

6章7章では、本研究の3つ目の目的「聴解テスト時、受験者はどのような聴解行動をとっているのかを明らかにする」。

6.1 テスト時の聴解行動の研究目的

日留試は、外国人留学生が日本の大学で学習をする際に必要な日本語のコミュニケーション能力を測定するテストである。コミュニケーション能力とは、実際の言語使用の場・状況・人間関係などを理解し、それに適した言語機能をもった文を使用する能力のことである（Richards, Platt & Weber 1985 邦訳 p.63）。しかし、テストという枠組みの中では、実際の使用場面と同じ状況で行うことは難しい。スキルとしての聴解ではなく、聴解活動としてとらえた Ur (1984, p.37) は、聴解活動の特徴は聞き手が何らかの目的を持って聞いている、としている。また、笠原他 (1994, p.38) は、聴解活動を、「聞き手がある目的を持って音声テキストから情報を引き出そうとする積極的活動である」と定義している。「ある目的を持って」とは、日留試の聴解テスト・アイテムではプリタスクにあたる。しかし、聴解アイテムのプリタスクは、実生活の聴解活動において聞き手自らが決めた目的とは異なる。

以上のことから、この章では実生活の自然な聴解活動と区別するため、テストにおける「聴解行動」とし、「聴解行動」は、「聞き手が自ら聞くという行為の方向性を決め、情報をどのようにして得たかの一連の行動をさす」¹と定義する。本章では、聴解テスト時の聴解行動の概念モデルを作成し可視化することで、聴解テスト時に受験者がどのような聴解行動を行っているのかを解明する。そして、留学生と日本人大学生の聴解行動を比較し、「アカデミック・リスニング」の特徴を分析・考察する。

6.2 研究の枠組み

3章「3.5 聴解の質的研究方法」(p.50)で述べたように、本研究ではM-GTAを分析方法とした。しかし、M-GTAは多くのデータを集める必要があるため（Ryan & Bernard, 2000 邦訳 p.179）、本研究のように1人の対象者が話した一言や、少数の対象者を対象とした研究に適していない。本研究では、大多数の人に見られる聴解行動を一般化してモデルを作ることを目的とはしていない。対象者の中の1名だけに見られた聴解行動も、1つの事象として捉え、多種多様な聴解行動をモデル化することが目的である。西條 (2007, p.4, 2008, p.93) は、M-GTAのデータを基礎とした研究法の長所を生かし、1人の対象者が話した一言や、少数の対象者を対象とした研究に適している構造構成的質的研究法（Structure-Construction Qualitative Research Method: SCQRM）というメタ研究方法を使用し、概念モデルを作成している。本研究では、この枠組みを使用し、少数データに適応させることにした。

構造構成的質的研究法（図6.1次頁）は、構造構成主義を超メタ理論とするメタ研究方法であり、関心相関性を中核概念としている。関心相関性は、「存在や意味や価値といったものは、すべて身体や欲望、関心、目的といったものと相関的に規定される」という原理である。例えば、「関心相関的構造構成法」はある「現象」に関心を抱き、その現象の中から「特定の事象」に着目する（関心の探索的明確化）。既存の知見に対して独自の観点から分析する場合は、構造探索過程の「発展的継承（関心相関的継承）」である。次に理論、方法論、枠組み、対象者などを研究目的にあわせて選び（関心相関的選択）、対象者に対

¹ 穂田 (2009, p.98) は、日本語では「きく」は「聞く hear」「聴く listen」「訊く ask/inquire」と3つの概念に分かれているとしている。本研究では、データ時に対象者がした行動が明確に「聴く」行動であると判断するのは難しいと考え、すべて「聞く」と表記する。

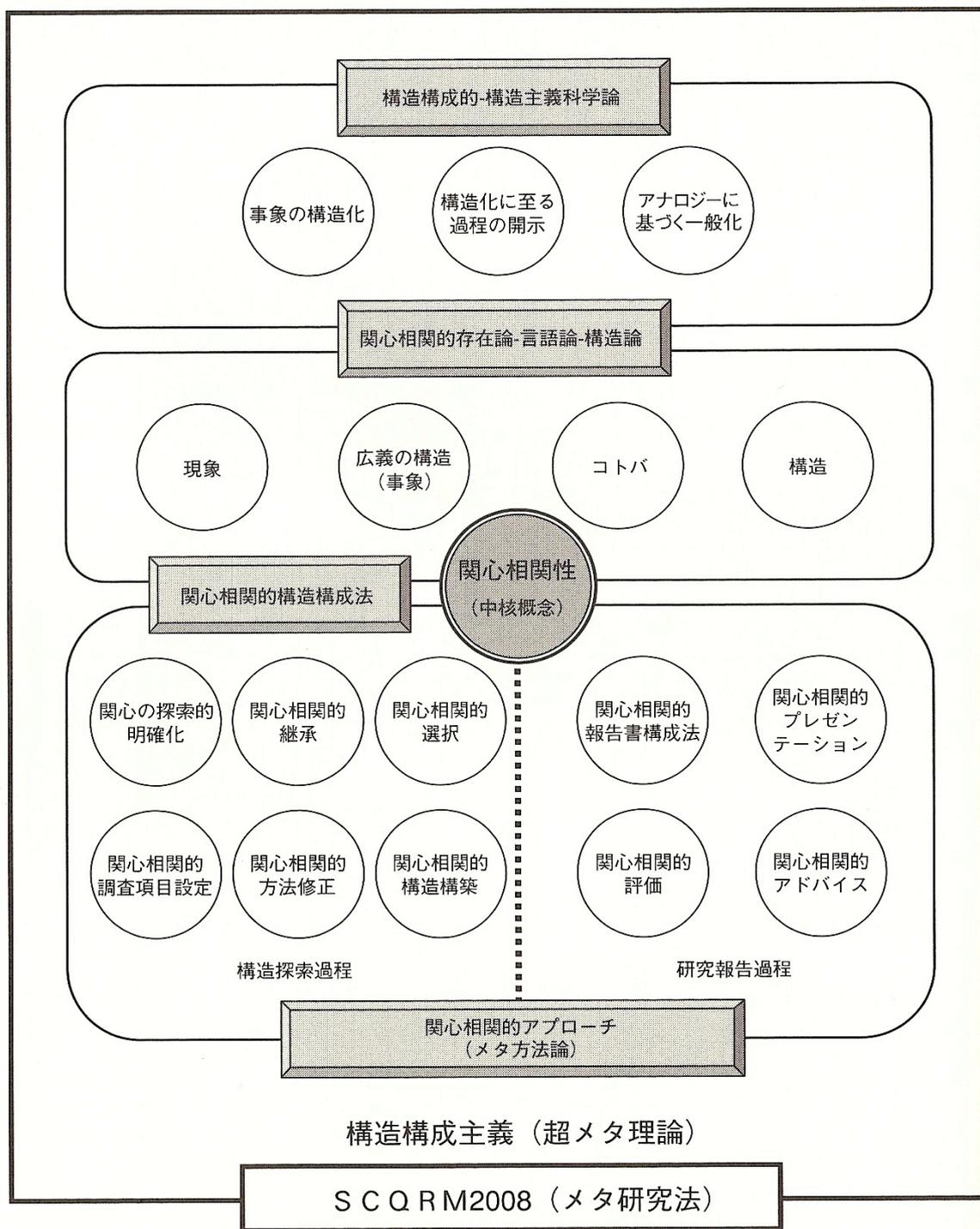


図 6.1 SCQR の構造モデル (西條 2008, p. 217)

しての質問項目を考える (関心相関的調査項目設定)。研究推進という目的にあわせ、方法概念を修正し (関心相関的方法修正法)、得られたデータから理論を構築する (関心相関的理論構築)。(西條 2008, pp.216-220)

構造構成的質的研究法の手法は、まず、得られた音声データから、ある事柄について話していると思われる部分を区切って、見出しをつける。このグループにまとめられたテキストのキーワードや見出し等に、概念として名前をつけ、構造化して全体像をみる。個々の概念の上位概念はカテゴリーとよび、包括的概念となる。最終的に、改善案を付け加え、動的な構造化をめざすのである（西條 2007）。

6.3 データ収集と分析の手順

6.3.1 対象者

本研究では、2グループの対象者からデータを収集した。第1グループは、日留試の受験対象者の外国人留学生（以下、留学生）である。第2グループは、アカデミック・リスニングのレベルを保持していると想定される日本人大学生である。留学生と日本人大学生の聴解行動を比較することで、「アカデミック・リスニング」におけるノンネイティブとネイティブの違いを分析・考察することにした。

青山他（2003, p.27）は、「非漢字圏と漢字圏では、因子構造が異なり、同じ語彙問題でも異なる能力、あるいはストラテジーを用いていることが示唆された」とし、聴解時において漢字圏、非漢字圏は語彙をどのように情報源として用いているのかを探った。三國他（2005, p.82）は、既知語彙率と内容理解の関係について調査した結果、既知語彙率が92%から93%の間で内容理解が高くなることが見られたが、これは韓国人に限っており、中国・台湾の漢字圏では相関が見られなかったと報告し、この結果から、漢字圏の人は語彙を音ではなく視覚情報として利用していると結論づけている。三國他の研究結果と、母語の言語的類似点が認知に影響を及ぼす可能性を考え、本研究では、漢字熟語、音読みの類似がある漢字圏、語彙の発音と統語的な類似がある韓国語圏、この2者以外を非漢字圏とした。

Ryan & Bachman (1992)は、インドヨーロッパ語族 (IE) と、そうでないグループ (NIE) に英語テストを実施し、その違いを見た。聴解ではIEのほうが高い得点を得たが、男女差はなかった。話題に関しては、受験者の学習背景が関係しており、NIE はアメリカ文化と、アカデミックで専門的な内容の結果がよかった。原因は、NIE の受験者は TOEFL の受験者でもあり、TOEFL は大学志向が強いからである。山下（2000）は、聴解理解に関わる因子を研究した結果、留学生には年齢、性別、出身国、学習歴、日本語能力などの影響があまり見られなかったと報告している。

以上のことから、本研究では年齢、性別、学習歴では対象者を大別しないことにした。出身国で3グループ（中国語圏、韓国語圏、非漢字圏）に分け、対象者の日本語レベルは、日留試を受験するレベルである中級修了以上とした。日本人大学生は、日留試が日本の大学に入学できる日本語能力の測定であること、また、大学生活に必要な知識とスキルを測定するということから、教養・共通科目（非専門科目）を多く学習する大学1年生、2年生を対象とした。

本研究の目的は、聴解テスト時の聴解行動そのものであり、広い枠組みから無作為抽出した対象者の一言一言を記述し、分析・考察することに意義がある。対象者の様々な特徴により分析する場合は、特徴の1つ1つが聴解行動の違いの要因であるかに焦点があり、聴解行動に影響を与える要因解明の研究となる。ある聴解行動が見られる対象群が、どのような特徴を持っているかという研究は、様々な聴解行動が明らかになった後、今後の課題とする。

本研究の留学生の対象者は、日本国内の大学日本語クラスで中級レベル²を修了した者から超級レベル³の学生であり、漢字圏6名・韓国語圏6名・非漢字圏9名の計21名である。発話思考法 (think aloud)

² 目安として旧日能試の2級レベルの文法的内容のクラスを修了した留学生。日常生活において日本語のみで十分生活できる日本語能力を有している。

³ 目安として自分の専門分野の大学の講義を日本語のみで受講できる日本語能力を有している。

と発問法 (question-asking method) の併用で話してもらった中から、言語的類似点の影響があると考えられた発話例には漢字圏・韓国語圏・非漢字圏を明記する。留学生 21 名中、旧日能試受験経験者は 17 名、日留試受験者は 0 である。日本人大学生は大学 1 年生 8 名・大学 2 年生 1 名、計 9 名である。

6.3.2 対象アイテム

使用アイテムは、公開されたテスト・フォームの中から、ライフ場面 1 アイテム、アカデミック場面の講義場面と授業場面から各 1 アイテムずつ計 3 アイテムを任意に選んだ⁴。本研究では、アカデミック場面が重要であると考え、アカデミック場面の「学術的」な「講義場面」と、「学問時」の知識や情報を授受する「授業場面」の 2 アイテムを選んだ。そして、日留試の 2002 年から 2006 年までの聴解テスト・アイテムでは、ライフ場面が一番多かった (全体の 42%) ため、ライフ場面からも 1 つ選んだ。3 場面のアイテムを使用して聴解行動の分析を行ったが、アイテムの場面が違っていても、キーワードを推測したり、不安をどう解決していくかなど、行動に共通性が見られたため、3 アイテムの聴解行動すべてを 1 つの概念図に記した。こうすることで、学習者の母語の違いと、テスト・アイテムの特性に影響されない包括的な概念図を作成することができると考えた。聴解アイテムの時間軸にそい、最終的に改善案を付け加え、動的な構造化をめざした。

実際に使用した 3 アイテムを以下に載せる。

・使用した 3 アイテム (設問あり)

1 番 (2006 年第 2 回実施 16 番 ライフ場面)

女子学生と男子学生が、ある町の、雨水を貯める取り組みについて話しています。この女子学生は、雨水を貯める一番の目的は何だと言っていますか。

女：最近、私が往んでいる町でね、雨水を貯めるための貯水槽を造ると、町が補助金を出してくれることになったんだって。

男：へえ でもその水って、トイレの水ぐらいにしか使えないんじゃない？

女：まあ、あとは草木の水やりとかね。

男：それじゃ、大して水の節約にはならないよ。

女：節水が目的というわけじゃないみたい。都会じゃ、短時間に大量の雨が降ったとき、道路が川のようになっちゃうことがあるでしょう？

男：うん。一度にたくさんの雨が下水道に流れ込んで、あふれちゃうんだよね。

女：そうそう。だから、雨水を一時的に貯めておけば、そういうことが防げるじゃない。

男：水道代とか節水の話かと思ったら、そういうことか。

女：そう。

この女子学生は、雨水を貯める一番の目的は何だと言っていますか。

1. 道路に水があふれないようにすること
2. 雨水をトイレの水に使うこと
3. 雨水で植木に水をやること

⁴ 3 アイテムとした理由を以下に述べる。筆者が行ったパイロット研究 (足立 2008b) から、留学生が聴解テストを行い、その後のインタビューに答えるのに適切な時間が 1 時間半から長くとも 2 時間以内であり、アイテム数は 6 アイテムが妥当であった。7 章で論じる「設問がないアイテム」も同じ対象者で行ったので、設問があるアイテム 3 と設問がないアイテム 3 とした。

4. 水道代を節約すること

2番 (2006年第2回実施 15番 アカデミック授業場面)

先生がディスカッションの進め方について話しています。先生は、この方法でディスカッションをすると、どんな利点があると言っていますか。

今日のディスカッションでは、ひとつ皆さんに心がけてもらいたいことがあります。相手の意見をよく聞くことはディスカッションの基本ですが、今日は、それを言葉で表現するようにしてください。つまり、相手の意見に対する自分の意見を述べる前に、まず、相手の意見を要約して述べてください。こうすると、「私はあなたの話をきちんと聞いていましたよ」ということが相手に伝わり、相手も「この人は私の意見を正しく理解してくれているな」と感じ、一種の信頼感が生まれます。これによって、もし相手の意見に反対しても、相手は素直な気持ちで聞いてくれるようになります。

先生は、この方法でディスカッションをすると、どんな利点があると言っていますか。

1. 相手の話をうまく要約できるようになる。
2. 反対意見を素直に受けてとめてもらえるようになる。
3. 相手が自分の意見に同意するようになる。
4. 話し方に自信がつくようになる。

3番 (2006年第2回実施 20番 アカデミック講義場面)

先生が、微生物について話しています。この先生は、実験から何がわかると言っていますか。

私たちの体の表面には、とても多くの微生物が活動しています。これらの微生物は、病原菌やカビなどから皮膚を守る働きをしています。しかし強い石鹸で皮膚を繰り返し洗うと、これらの微生物は体の汚れとともに皮膚から流れ落ちてしまいます。それでは、何回くらい洗うと微生物が排除されてしまうのでしょうか。手を石鹸で洗う実験をしてみると、1回目では、手を洗う前と比べて微生物は70%に減るだけです。2回目では半分になりました。3回目で10%まで減少し、4回目では全く検出されなくなりました。体を丁寧に洗うことは清潔でいいことのように思われていますが、この実験結果から、何度も洗うと、皮膚に悪影響を与えることがわかります。

この先生は、実験から何がわかると言っていますか。

1. 洗いすぎると、皮膚を守る力が落ちる。
2. 何度も洗うと、皮膚が丈夫になる。
3. 丁寧に洗うと、皮膚を清潔に保てる。
4. 洗う回数に関係なく、強い石鹸は皮膚に悪い。

6.3.3 データ収集と分析の手順

データ収集の手順は、以下のとおりである。対象者に選択肢の番号が書かれた解答用紙(参考資料7

「解答用紙」p.199)を渡し、1アイテムずつ聞いてもらった。メモは自由にとってもよいという指示⁵と、1アイテムの解答が終わった時点で、どのように解答を得たのかを話してもらうという説明をした。1番の選択肢を聞いた時点で正答だと確信した選択肢には「○」を、誤答だと確信した選択肢には「×」を、正答の可能性はあるが正答と確信まではできなかった選択肢には「△」を、選択肢を聞いてもわからなかった場合は「？」を書き込んでもらうように指示した。2番の選択肢、3番の選択肢、4番の選択肢を聞いたときも同様にしてもらった。これは解答後、なぜ正答と判断したのかをインタビューする際、対象者に参考にしてもらうためである。4つの選択肢を聞き終えた後に、正答を1つ決め、その番号に「○」をつけてもらった。解答後、「会話を聞いているときに、頭の中で考えていたことを話してください。」という問いかけをして、記憶していることを話してもらい、録音した。対象者は、聴解テスト・アイテムを聞きながら、消え去る音声の処理をフルスピードで行っているため、どのような行為をしているのかを、その行為と同時に話すことはできない。したがって、問題の解答を得たあとに、その過程を振り返ってもらい話す発話思考法 (think-aloud) の手法をとった。ただし、解答を得るまでにしている行為を適切に捉え、言語として表出できる対象者ばかりではないため、半構造化インタビュー⁶ (参考資料8「インタビュー項目」p.199)も並行して行った。

上記で得られた音声データ⁷を文字化し、聴解行動の具体例(発話内容)を抽出した。日留試の聴解アイテムは、①状況説明 ②プリタスク ③本文 ④ポスト・クエスチョン ⑤選択肢の時間的流れがあり、各部分がそれぞれ機能を持っているため、それぞれの部分で分析を行った。対象者の具体例で、同じ内容と判断したものを1つのヴァリエーションとし、上位聴解行動として整理する⁸。それぞれの聴解行動を説明する概念名をつけ、その定義をおこなった(図6.2)。この際、1人だけの行動であっても大多数の行動であっても、行動の種類としては1つと捉えた。なぜならば、この研究では、テスト時の一般的な聴解行動パターンを構築するのではなく、多種多様な聴解行動を可視化し、客観的にその行動を振り返ることができるようにするからである。聴解アイテムの時間軸にあわせ、上記の①から⑤までの間に行われた聴解行動をモデル化した。その後、時間的に、前にある行動が後の行動に影響していると考えられるものを線で結んだ(図6.3「留学生の聴解行動(設問ありのアイテムに対する行動)」p.125)。

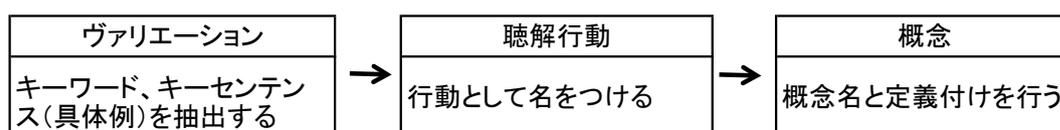


図 6.2 分析の手順

6.4 留学生の聴解行動の分析結果と考察(設問あり)

留学生の対象者全員を漢字圏・韓国語圏・非漢字圏の3類としたが、母語使用以外に3類の差があま

⁵ 日留試の聴解テストではメモを自由にとってもよいので、同じ条件とした。

⁶ 質問項目を固定した構造化インタビューと、研究の関心である質問を最初に尋ね、あとは対象者の話すままに任せる非構造化インタビューの間にある。今回は研究の関心に照らし合わせ、対象者によって、インタビュー項目の数や内容を変化させてインタビューした。

⁷ データの1例は参考資料9「留学生のデータ対象者 No.9」(p.200)を参照。

⁸ 例えば、「問題文を聞かなきゃって」と「質問がどんな利点があるかだったので、利点とメモして○をして、それを聞こうとした」は、同じ行動であり、ヴァリエーションの「解答部分を探す」に入り、プリタスクの「限定する」聴解行動として1つにまとめた(表6.3,p.129)。

り見られなかった。早川（1993, p.31）は、聴解理解において上級レベルでは漢字圏と非漢字圏の差はなくなるという結果を得ており、今回の研究でも差がなかった。そこで、留学生全員を1つの類とし、母語による違いが見られたものだけ母語を記載した。

図6.3「留学生の聴解行動（設問ありのアイテムに対する行動）」（次頁）は、留学生の聴解行動である。聴解テストの時間的な流れは上から始まり、最後の「解答」で終わる。また、右横の「複数スキル併用」と「複数言語利用」という概念は、テスト・アイテムが始まってから終わるまでの全体に関わる行動であるため、これを「全体」として分析・考察をする。□は、テスト・アイテムを聞きながらされた聴解行動であり、それらのグループの概念が■である。○は、ある聴解行動をした結果起こった現象である。これは、聞き手が自ら決めた聴解行動ではないため現象として区別した。「本文」と「ポスト・クエスチョン」の間にある概念「解答の言語化」は本文が終わった直後であり、ポスト・クエスチョンが始まる直前である。「本文」と「選択肢」には複数の行動が見られるが、時間的な違いはない。「→」はある行動を取ると決定した場合、次の行動に影響を与えると判断したものである。

留学生に特徴的に見られた現象として、状況設定時の「わからない」、プリタスク時の「聞き逃す」、本文時の「その後聞けない」、ポスト・クエスチョン時の「曖昧である」があり、ある要因により解答を得るための障害となっている⁹。これらの現象については各項目の分析と考察で後述するが、時間が進むにつれて行動に負の影響を及ぼしている。

以下、アイテムの各部分の聴解行動を詳細に分析・考察する。個々の対象者がコメントした発話内容は、ヴァリエーションの1例であるため、「例」と記述する¹⁰。

⁹ 「解答放棄」は日本人大学生にも見られたため、留学生だけの特徴としては取り上げなかった。

¹⁰ 例えば、対象者が発話思考法とインタビュー時に「メモを取るの、聞いているときに書いているほうが集中できるから」と言及した場合、『メモを取る理由として「書いているほうが集中できる」という例があった』と記述する。

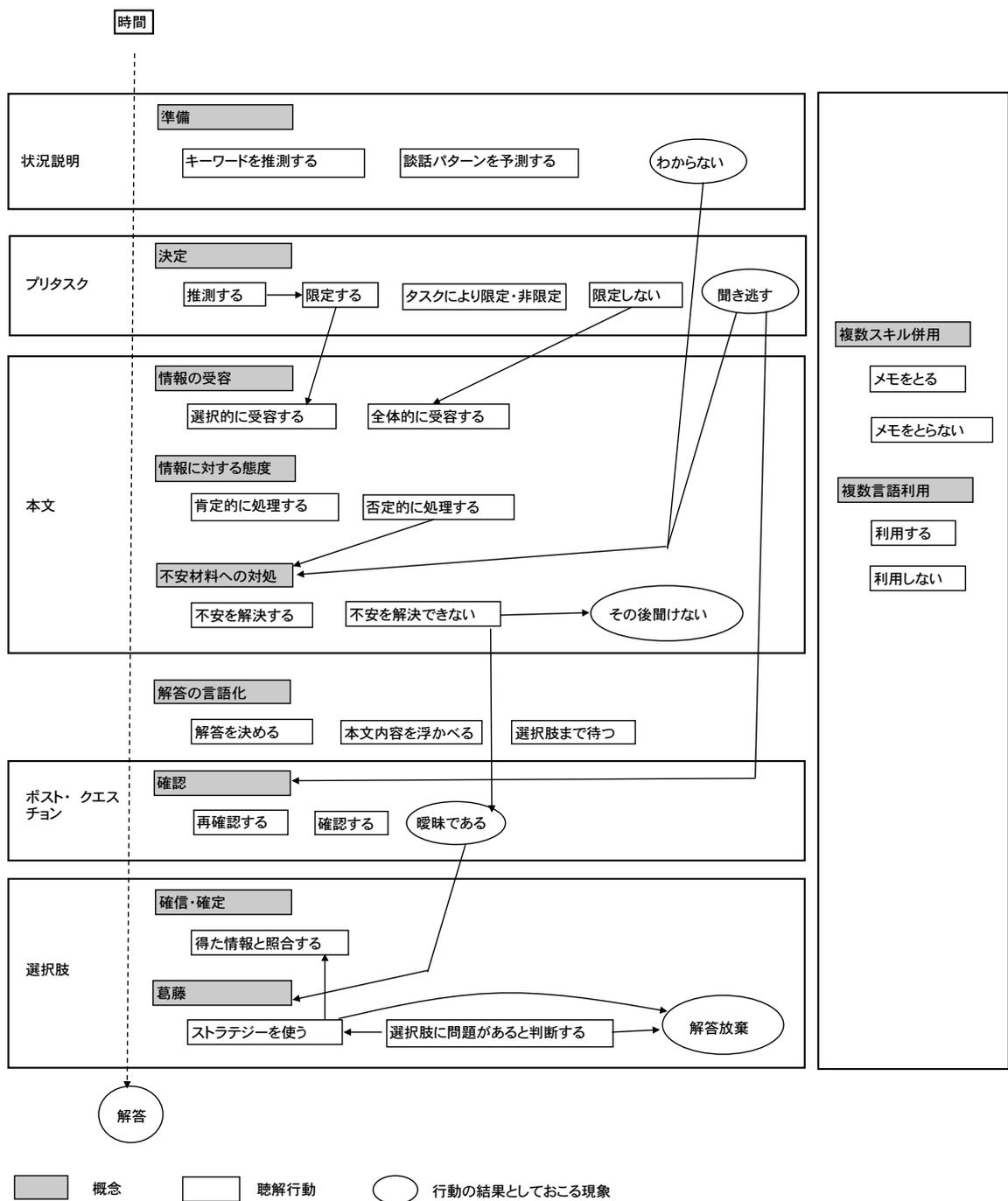


図 6.3 留学生の聴解行動（設問ありのアイテムに対する行動）

6.4.1 全体

表 6.1 アイテム全体の概念と聴解行動のヴァリエーション①（留学生設問あり）

7行ム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
全体	複数スキル併用	聞きながらメモをとる	メモをとる	①推測の手がかりにする（語彙） ②記憶のためにする ③整理する ④集中する	
			メモをとらない	①本文が聞けなくなる ②本文に集中できない	
全体	複数言語利用	理解する際に複数言語を利用する	複数言語を利用する	①母語のほうがメモをするスピードが速い ②何語でも1番速いものを使う（漢字圏） ③カタカナ語の理解のために英語での意味を考える ④訳している ⑤安心する	
			日本語のみを利用する		

表 6.1 は、留学生の設問があるアイテム全体に対して見られた行動とそのヴァリエーションである。1 アイテムの最初から最後まで、全体において見られた行動には「複数スキル併用」と「複数言語利用」があった。

まず、「複数スキル併用」を分析する。複数スキルは、聞く行動と書く行動を同時に行うことである。1つ以上のアイテムでメモをとる行動が見られたのは、対象者 21 名中 17 名であった。対象者によって、聴解テストのすべてのアイテムでとる、すべてのアイテムでとらない、アイテムによってとったりとらなかったりするという3つのタイプが見られた。

「メモをとる」のヴァリエーションは4つである。メモは、ほとんどが単語レベルであったが、文章も少し見られた。メモの内容は、プリタスクやキーワード、キーセンテンスと判断したものであった。「①推測の手がかりにする（語彙）」は、設問に出てきた語彙から本文の内容を推測する、本文内の語彙をヒントに正答を推測するなどである。「②記憶のためにする」は、内容理解に「必要である・解答のヒントとなる・解答だ」と感じたときにメモを取っておくという、消え去る音を記録する聴解スキルに特有な行動と言える。「③整理する」は、聞き取った情報を整理し、内容理解の補助とするためである。「④集中する」は、「書いているほうが集中して聞ける」「聴解テストは動きを伴わないためあえて体の一部を動かして集中する」という例があり、聴解テスト環境に対する行動である。「メモをとらない」ヴァリエーションは、2つの行動が同時にできない、書くほうが遅いので聞いていることと一致しなくなる、聞くことに集中したいため「①本文が聞けなくなる」、そして「②本文に集中できない」という例であった。

「メモをとる」行動は、通常の聴解活動の中で必ずするものではないが、授業中や会議中に記録しておかなければならない場合には行う。授業のタイトルや会議の重要部分を、記録するためのメモである。これらは、「②記憶のためにする」「③整理する」ためにしているものであり、テスト特有の行動ではないが、聴解スキルとして必要である。「①推測の手がかりにする（語彙）」は、テスト特有の行動ではあるが、留学生の実生活を考えると、わからない語彙を書いておき後で辞書で調べる、友人に訊くという行為につながる。「④集中する」は、身体的行動を伴うことで聞くスキルを向上させるストラテジーの範疇であると考えられる。足立（2010,p.127）は、被験者が聴解テスト時にメモをとる行動は、テスト特有の行動であり、これが留学生が大学の講義ノートをとる行動に将来発展するかには疑問を呈しながらも、タ

スクに対して必要情報を抽出してメモができる行動は、大学生に必要な行動であるとしている。このようにメモを取る行動が留学生の大学生活上、現実的であり必要だと考えられるため、聴解テスト時にメモをとる必然性が生じるテスト・アイテムが、必要ではないかと考えられる。

日留試の聴解テストでは、メモを自由にとることが可能であり、問題用紙にメモ用の余白が設けられている。記憶にかかる負荷を軽減させるための余白ではあるが、メモをする・しないは受験者の自由である。聞きながらメモができるという能力はノート・テイキングに置き換えることができ、テストで測定する形式も考えられる。日留試には、聴読解という聴解・読解の複合スキルの分野が既に存在する。IELTS の聴解テストでは、電話を聞きながらメモをしたり、休暇願いの書類の書き方を聞きながら書類に書き込むというような複合スキルを測定している。留学生の聴解テストで、①聴解スキルのみ、②聴解・読解の複合スキル、③聴解・メモ¹¹の複合スキルを測定することは、実際の聴解活動に近づき、真正性が高いと言ってよいだろう。

次に、「母語利用」に関してだが、母語について言及した対象者と、メモに母語使用が見られたのは21名中10名であった。メモの大部分を母語でとっている対象者は2名で、あとの8名はカタカナ語を英語で書く、わからない語彙を音としてメモするときに母語を利用する程度であった。なお、漢字圏の対象者に英語使用なども見られたため、「母語利用」とはせず、「複数言語利用」という概念とした。「①母語のほうがメモをするスピードが速い」は書き慣れているからである。「②何語でも1番速いものを使う」は、メモに適した言語を選択する能力があることを示している。例えば、漢字圏の対象者が、アルファベットや英語で日本語語彙をメモしている例や、漢字ではなくひらがなを使用している例が多く、書くスピードを重視していた。「③カタカナ語の理解のために英語での意味を考える」も、対象者の母語と無関係であった。今回、すべての対象者が英語既習者であり、日本語のカタカナ語が主に英語の単語に由来することを既知知識として持っていたからだと考えられる。カタカナ語に関しては、非漢字圏、特にゲルマン諸語・ロマンス諸語圏が有利というわけではなかった。日本語のカタカナ語の音が、本来の英語の音と少し違う、あるいは日本語音に置き換えられていると推測しにくく、英単語の知識が高い非漢字圏対象者は似た発音の語彙を多く知っているがゆえ間違った推測をしてしまう(例えば、「ディベート」を「divide」と推測)という例も見られた。韓国語圏の対象者は、母語でも英語の発音に近い語彙を使用しているため、推測はしやすかったと述べているが、漢字圏の対象者は、外来語を母語に訳しているため音からの推測には困難が伴うようであった。足立(2010, p.127)は、日本語のカタカナ語使用は今後も増えると予想し、大学生活における一般性を考慮してテストに出すべきだと述べている。例えば、今回使用したテスト・アイテムの中に「ディスカッション」という語彙があったが、日本人大学生からは「大学入学後に聞くようになった語彙である」という例があり、大学内での使用の頻度や、一般性についての調査が必要である。「④訳している」は、全文翻訳ではなく、単語レベルであった。「⑤安心する」は、母語利用により理解を確実にすることから派生していると思われる。「日本語のみを利用する」のヴァリエーションの欄には何も無いが、対象者が中上級から上級であったため、日本語のみの使用が自然・当然と考えており、特に語らなかつた結果である。

¹¹ メモではなく記述(自由解答)のほうが現実に即しているが、採点時間および採点の信頼性を考慮すると、メモのほうがテストの実施性が高い。

6.4.2 状況説明

表 6.2 状況説明時の概念と聴解行動のヴァリエーション②（留学生設問あり）

7行目	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
状況説明	準備	本文を聞く前に推測する	キーワードを推測する	①大切な語彙を推測する ②常識的な解答になると推測する	語彙がわからない
			談話パターンを予測する	①ダイアログかモノログか	

状況説明では、本文を聞く前の「準備」をする。

「①大切な語彙を推測する」は、状況説明時で本文のトピックがわかるため、プリタスクや本文でどのような語彙が使用されるのかを推測する。「②常識的な解答になると推測する」は、既有知識から解答を予測している。例えば、状況説明で、「ディスカッションの進め方について話している」と聞いた場合、本文ではディスカッションをする際の順番が話されるのではないかと予測をし、本文の最初と最後がキーワードになるのではないかと推測するのである。行動の結果として起こる現象に、「語彙がわからない」とあるが、状況説明の段階で使用されている語彙がわからない場合は、本文を聞く際の不安材料となる。何について聞くかを知らないで聞く状況が、現実には起こりにくいことから聞く動機付けが必要である（加藤 1992, p.35）。その動機付けが、状況説明にもあるのだが、本文で問われる日本語能力より簡単に理解できるものでなければならない。状況説明の1文は、極力簡単な文章におさえ、本文で本筋のトピックに自然と流れていくような談話内容にするなどの工夫が必要だと考えられる。

「②ダイアログかモノログか」は、本文の談話パターンを推測することである。ダイアログの場合は男女の会話であるため、「どちらかが質問をしてもう1人が説明するのでその中に答えがある」「1つ聞き逃しても次の返事でだいたい想像できる」など、会話のパターンを知っており、解答を得る準備をしている。「2人だと感情が声のトーンにあらわれる」「2人だと性格などがヒントになる」「会話のほう聞いていておもしろい」など、会話に対して肯定的に捉えている例があった。一方、「会話は2つの意見が出る可能性があり緊張する」「会話があちこちにいく（意見や内容が一貫していない）」と、否定的な例もあった。モノログの場合、「構成がはっきりしている」「流れがある」と、談話構造による聞きやすさに対して肯定的な例がある反面、「展開が速い」「イントネーションに特徴がなくキーポイントがわかりにくい」という否定的な例もあった。日留試の聴解テスト全体の構成をみると、モノログが後半に多く見られる。これは、アカデミック場面が多く、語彙などが専門的であり、困難度が高いという理由であると考えられる。しかし、「留学生にとって大学生活より日常生活の語彙のほうが使用頻度が低く難しい」という例から、日本国内の生活者や海外の日本語学習者にはない、留学生特有の語彙の困難度があるに違いないと推測する。日留試では、「日本での留学生活を送る上で、日本語によるコミュニケーション能力があるかどうか」を測るのであるが、留学生の生活の大半は大学内にあると考えるべきである。大学内で使用頻度や重要度が高い語彙の研究、既出のアイテム分析から留学生にとっての困難度の特徴などの研究が待たれる。

6.4.3 プリタスク

表 6.3 プリタスク時の概念と聴解行動のヴァリエーション③（留学生設問あり）

7行4	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
プリタスク	決定	解答になる部分に焦点をあてて聞くか、全体を聞くかを決める	推測する	①キーワードを推測する ②本文構造を推測してどこに大切な部分があるか予測する ③常識的解答になると推測する	語彙がわからない
			限定する	①解答部分を探す ②会話の場合プリタスクで限定されたほうに集中する ③全部聞こうとすると聞き逃す	
			プリタスクにより限定するまたは限定しない	①プリタスクによる ②大切な部分の比重は重くするが全体を聞く	
			限定しない	①1語彙で内容が変わることがあるから全部に注目する ②談話内に解答が散在している ③質問の答えばかりに注目すると本文内容把握がおろそかになる ④とりあえず全部聞く	
			なし	①設問がわからなくても本文でわかる ②簡単すぎて油断 ③形式に慣れていない	聞き逃す

プリタスクでは、本文で聞くべき内容がわかるため、本文をどのように聞くかを「決定」する。

プリタスクでは、状況説明にあったトピックがより詳細になる。例えば、状況説明では「微生物について話しています」だが、プリタスクでは「実験で何がわかりますか」であるため、微生物全般ではなく、実験結果についての話だとわかる。そして、何を聞かなければならないかが明確になる。上記の例では、実験方法でもなく、実験時の注意でもなく、実験結果を聞かなければならないことがわかり、状況説明時の推測より絞った「①キーワードを推測する」ことができる。「②本文構造を推測してどこに大切な部分があるか予測する」は、例えば、「実験」の談話構造は最初に実験方法や手順があり、結果があり、結論がくるだろうと推測される。よって、談話の最後のほうに解答部分が話されるのではないかと予測するのである。ここでの「推測する」行動は、本文を聞く際に解答部分に集中して聞く「限定する」行動の前段階と考えられる。

「限定する」は、本文内の情報を解答に必要なか不必要かと、選択的に聞くことである。「①解答部分を探す」は、本文を聞く際に常にプリタスクを意識しながら聞く行動である。そのため、プリタスクをメモしている対象者が多かった。「②会話の場合プリタスクで限定されたほうに集中する」は、例えば、プリタスクで「女子学生は～と言っていますか」という場合、女性の発話に解答があると予測し、女性の発話により集中して聞くことである。この2つの行動は、積極的に解答部分のみの情報に絞って聞こうとしている行動である。「③全部聞こうとすると聞き逃す」は、今までの聴解テストの経験から、全体を聞くより解答部分に絞って聞いたほうが解答できると判断した例である。

「プリタスクにより限定するまたは限定しない」の「①プリタスクによる」は、プリタスクを聞き、何を解答するのが明らかになった時点で、解答を得るために全体を聞いた方がよいのか、部分でよいのかを決定する。「②大切な部分の比重は重くするが全体を聞く」は、「限定する」聞き方と「限定しない」聞き方の両方を行うことで、聞き逃しのリスクを避けた例である。

「限定しない」は、本文全体を偏りなく聞く行動である。「①1語彙で内容が変わることがあるから全部に注目する」「②談話内に解答が散在している」は、今までの聴解テストの経験をもとに、談話構造の展開を最後まで聞かないといけないと判断している。「③質問の答えばかりに注目すると本文内容把握がおろそかになる」「④とりあえず全部聞く」は、部分理解ではなく全体理解を重視している。

我々が日常生活のなかで行う聴解は、部分理解で、自分の得たい情報のみに注意して聞く場合が多い。しかし、講義のような場合は、自分にとって必要な情報が前もって限定できないため、全体を聞き、重要だと思われる部分のみを抽出して記録する。つまり、何をどのような目的で聞くかにより、聞き方は変わる。聴解テストで設問がある場合は、何を聞くべきかがわかるため部分理解の聞き方でよいと思われるが、全体理解の聞き方をする対象者は、自分の聞くスタイルを持っており、それに従って聞いたほうが、理解しやすいと判断したのではないだろうか。

意図的な聴解行動ではないが、結果的に生じた行動として「聞き逃す」があった。例えば、「①設問がわからなくても本文でわかる」と、設問を重要視していない例があった。古川他(1992, p.175)の研究結果にも聴解テストの設問に対して「設問があると解答となるポイントしか聞かなくなるため望ましくない」と述べている日本語学習者がおり、また、上記のように、設問があっても全体を聞く対象者はいる。聴解テストであるから、設問があるから、という理由で、日常と違う聞き方を強いられる受験者がいるのであれば、設問の有無の検討が必要である。「②簡単すぎて油断」と「③形式に慣れていない」は、設問の必要性はわかっていながら、聞き逃した例である。この場合、本文を聞く際の不安材料になると思われたが、今回の対象者は解答に窮した様子ではなかった。「②簡単すぎて油断」の例の対象者は、日本語に限らず外国語の聴解テストを受けた経験が多く、設問がない聴解テスト・アイテムがあることも知っており、さほど不安を感じていなかった。反して、「③形式に慣れていない」対象者は、このような聴解テストを受けた経験が全くなく、どのような構成でアイテムができているのかを知らなかった。アイテムを聞く前に一度説明はしていたが、イメージできなかつたようである。この2名とも、本文内でプリタスクの聞き逃しを解決し、ポスト・クエスチョンで確認していた。

ここで、「限定して聞く」場合の、限定する強さについて分析する。解答を確信した後は、本文内容を聞かず、実際は聞き流した部分に解答があったため選択肢から解答を選ばなかった例や、本来は詳細な正答の情報を得なければならなかったのだが、自分が考えた解答に満足してしまい、選択肢から選ぶ際に適切なものが見つけれなかった例があった。これらは限定を強くしすぎたために、解答に必要な情報を取りこぼしてしまったためである。逆に、限定が弱く様々な情報を抱えてしまい、正答と錯乱肢で迷った例もあった。限定して聞いた場合、本文が終わってから選択肢を聞くため、受験者が本文を聞きながら集めた情報との間に差が生じてしまう。限定せずに全体内容把握に決定した場合、必要な情報の取りこぼしは少なくなるが、一貫して集中力を保たなければならない。どちらの聞き方をするにせよ、受験者は聴解テストの形式に慣れていた方がよいと言える。

以上のことから、プリタスクにおける聴解行動は、「限定して聞く」か「限定せずに聞く」かの2つになる。この行動は、どちらが効果的であるということではない。行動決定の要因の1つは、解答部分さえわかればよいと考えるか、全体の内容把握をしないと不安だと考えるかの個人差である。もう1つは、自らの受験経験をもとにした判断であり、これも個人の経験値の差である。

6.4.4 本文

本文では、プリタスク時に決定した聞き方にそって情報を収集していく。

本文の聴解行動は、「情報の受容」「情報に対する態度」「不安材料」の3つの概念に分け、分析を行う(表6.4次頁)。

表 6.4 本文時の概念と聴解行動のヴァリエーション④（留学生設問あり）

7/74	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例		
本文	情報の受容	解答のための情報を得る	選択的に受容する	①解答のための情報と認知する ②不必要な情報は捨てる ③解答を得た後はあまり集中しない			
			全体的に受容する	①内容把握する ②全体内容を要約する ③最後が変わることがある ④解答は最後にくることが多い ⑤常識と違うことを言う可能性がある			
	情報に対する態度	難易に関わる要因から情報処理をする	肯定的に処理する	言語形式・談話構造 ①音声の特徴がヒント ②スピードがよい ③男女の会話なのでどちらが言うかわかる ④語彙が簡単 ⑤統語的機能がわかる ⑥ディスコース・マーカ―に注意する ⑦構成が明確だ ⑧論理的な構造だ ⑨解答が明確に提示される テキスト外 ①常識から判断 ②常識と違う ③背景知識から判断 ④話すときの習慣から判断する ⑤話し手の性格を推測する ⑥惑わし選択肢、罣を推測する ⑦テスト・テクニックを使う キーワード ①トピックから推測する ②プリタスクから推測する ③頻度が高い ④内容理解から推測する			
			否定的に処理する	言語形式・談話構造 ①スピードが速い ②発音が不明瞭だ ③母語と違うため（非漢字圏は統語的相違）・語彙がわからない・カタカナ語（漢字圏）が難しい ④統語機能がわからない ⑤談話構造に対するビリーフと違う ⑥トピック展開が速い ⑦解答が不明確だ ⑧量が多い ⑨情報量が少なすぎる テキスト外 ①背景知識がない ②常識と違う ③間違った常識を使う ④トピックに興味がない、苦手だ、専門外だ、遭遇頻度が少ない ⑤メモに集中しすぎる ⑥解答を確信した後聞かない ⑦集中力が低下する ⑧緊張 ⑨課題が予想外 キーワード ①キーワードの勘違い（音から判断するため）			
			不安材料 ¹²	わからない語彙や表現がある	不安を解決する	①文脈から推測する ②母語から推測する ③カタカナ語は英語から推測する ④語彙分析から推測する ⑤トピックの背景知識を使う ⑥回避する	
					不安を解決できない	①集中力が分散する ②その後の本文を聞けない ③思考が止まる	その後聞けない・パニックになる→次のア

¹² 7章で聴解テスト・アイテムの設問がない場合の分析を行うが、「不安材料」に関して、設問がある場合とない場合に少し違いが見られたので、設問がある場合を「不安材料1」とし、設問がない場合を「不安材料2」とした。

					アイテム時にも 悪影響
--	--	--	--	--	----------------

本文の聴解行動は、「情報の受容」「情報に対する態度」「不安材料」の3つの概念に分け、分析を行う。

まず、「情報の受容」についてである。「プリタスク」時の決定が、限定した聞き方であれば、「選択的に受容」し、限定しない聞き方であれば、「全体的に受容」する。「選択的に受容」する場合、「①解答のための情報と認知する」と、その部分を記憶あるいは記録し、「②不必要な情報は捨てる」。「③解答を得た後はあまり集中しない」という例があったが、日常生活の聴解においても、必要な情報を得たあとは聞き流すため、自然な行動である。「全体的に受容」する場合は、「①内容把握する」「②全体内容を要約する」のように、会話内容全体を理解し、そこから解答を得ようとしている。「③最後が変わることがある」「④解答は最後に来ることが多い」「⑤常識と違うことを言う可能性がある」は、今までの受験経験からの例だが、これは通常の聴解とテスト聴解の展開が違うことを示唆している。日留試の聴解アイテムは、1分～2分程度の談話で構成されており、その短い中に情報を集約させ、明確な解答を得させなければならない。また、錯乱肢のための情報も入れ込むため、不自然さを感じる談話もある。自然な談話構成では、結論などは最後に来ることが多いが、その知識のみで解答が得られることを避けるために、談話構成の組み替えなどの工夫をしている。

日留試は、項目応答理論 (item response theory:IRT) を使用し等化¹³を行っている。項目応答理論の中には複数の分析方法があるが、よく使用されているのが1パラメーター・ロジスティック・モデルのラッシュ分析 (Rasch analysis) である (Hughes 2003 邦訳 p.243, Bachman 1990 邦訳 p.232)。ラッシュ分析は、アイテムを困難度順に並べることが可能であり、受験者が一定の能力レベルを保持しているという前提で成り立つ。あるアイテムがラッシュモデルに適合しなかった場合、そのアイテムに欠陥がある可能性を示唆しており、そのアイテムは次回から使用できないことになる。もし、1アイテムを長くし、設問を増やしたアイテムに (例えば、1アイテムで3つの設問)、ラッシュモデル不適合の設問が2つ出たとする。適合した設問には欠陥がなくとも、このアイテムは再利用できない。これでは、アイテム作成の労力とアイテム数確保に対してリスクが大きい。あくまでも推測の範疇ではあるが、日留試の聴解テスト・アイテムが、一問一答であり、アイテムの長さが短い理由の1つではないかと想像する。しかし、日留試の目的を考えると、消極的に短いアイテムで1つの解答をさせるのではなく、積極的に長い講義を聞かせ、講義内容の理解能力、必要な部分の抽出能力、課題設定能力という、アカデミック・リスニングに必要な能力を測定すべきではないだろうか。

次に、「情報に対する態度¹⁴」を分析する。「肯定的に処理する」は、聞き取った情報から解答が得られる、あるいは簡単だと感じており、「否定的に処理する」は、解答が得られない、難しいと感じたときである。「情報に対する態度」では、「何から情報を得たのか」に注目し、言語形式・談話構造、テキスト外、キーワードの3つに分けた。

肯定的に処理した「①音声的特徴がヒントになる」「②スピードがよい」と、否定的に処理した「①スピードが速い」「②発音が不明瞭だ」は、音についての処理で、聴解の特徴と言える。音を語彙や文法表現として認知し、語彙の意味や統語機能がわかることから、音に対して肯定的に処理ができるか否かは、

¹³ 「テストの異なるフォームから得られた測定結果を相互に比較可能にするため共通尺度上で表すための手続きである」(石川他 2011, p.45)。

¹⁴ 「態度」とは language attitudes のことである。ある言語に対する肯定的・否定的感情は、難易度 (困難度)、重要性、社会的価値などに対する印象に反映することがある (Richards 他『ロングマン応用言語学用語辞典』1985, p.200)。

理解度に影響すると考えられる。また、「③男女の会話なのでどちらが言うかわかる」も、声に注目している。プリタスクで、男性か女性か、どちらの意見が解答となるかを指定された場合は、声で判断できるとした例である。例えば、プリタスクが「男性はどう思っていますか」の場合、「限定して聞く」対象者は、情報を言うであろう男性の声により注目して聞く。「より注目する」とは、女性の会話の中にも補助的な情報があることを知っており、また、女性が解答情報を述べて男性が肯定するという談話展開も予測している。「限定しない」で聞く対象者も、「1語彙で内容が変わる場合がある」と言っており、これは男女どちらが重要なことを言うか、両方に注目しているということである。テスト聴解の2人の会話が、必ず男女の会話であることの理由がここにある。テストでは聞き返しができないため、音の点では滑舌がよく、スピード調整とノイズの排除がされる。反面、このような調整をすることで、自然な発話でなくスクリプトを読んでいる不自然さが出てしまう（笠原他 1994, p.40）。テストという時間的にも空間的にも制限された中で、談話の音に関して、特にコミュニケーション・アプローチが主流となっている昨今において、ポーズ、フィラー、緩急のあるスピード、ノイズ、言いよどみ、言い間違えと訂正、話者の癖など、どのように扱うのか。聴解の目的と照らし合わせ、どのような音の条件が妥当かの検討が、日本語聴解テストのみならず、聴解練習にも求められる。

音に対しての処理と同様、語彙や統語、談話構造に関しても、肯定的な処理と否定的な処理は対をなしているものが多い。言語の知識量が多ければ内容理解がしやすいと考えられるが、それだけではなく、聞こえた音と、自らが持っている言語知識とを一致させられるかどうかが重要である。ただし、知識量が多ければ、曖昧な音でも推測がしやすいということは言える。本研究では、対象者の日本語知識量の違いによる処理のしかたについては検討しなかったため、明確な記述は控えるが、知識量が多い対象者がより肯定的な処理をしたとは、一概に言えない。談話構造に関しては、肯定的処理の「⑦構成が明確だ ⑧論理的な構造だ ⑨解答が明確に提示される」と、否定的処理の「⑤談話構造に対するビリーフと違う ⑥トピック展開が速い ⑦解答が不明確だ」から、談話展開のパターンを多く知っているほうが理解しやすい反面、聞いている談話が自分の既有知識と違う場合は、聞き方にブロックがかかってしまうようである。否定的処理に見られた「⑧量が多い」は、情報処理のスピードに関係している。否定的処理の「⑨情報量が少なすぎる」は、日留試の1アイテムの長さの上限も関係あるのではないだろうか。

「否定的に処理する」例として多かったのは、「語彙がわからない」であった。田中他（1986）は、日本語ネイティブに比べ、ノンネイティブは、わかる言葉とわからない言葉の間にぼんやりとわかる言葉が存在し、それを語彙のリファレンス辞書から推測していると述べている。このことから、ノンネイティブは、語彙の推測を頻繁におこなわなければならないだろうことが推測される。水田（1996, p.53）は、日本語学習者が聞き取れない原因を、6つに下位分類し、調査を行った結果、「文節化（語の認知ができない）」と「辞書的意味（語句の文字通りの意味がわからない）」の2つで79.6%と、圧倒的に語彙の知識不足から来ていると結論づけている。これらの研究結果は、語彙の重要性を示唆しており、本研究の語彙に関する否定的処理が多かった結果と合致した。山本（2005, p.125）は、対談、講演、講義、テレビドラマの音声言語を分析し、大学学部留学生に必要な語彙は、旧日能試の1級や級外の語彙ではなく、2級以下の基本的な語彙や文法が重要であると述べている。この章で分析している設問のある3アイテム（pp.121-122）のキーワードは、1番「雨水、貯水槽」、2番「ディスカッション」、3番「微生物」であった。すべて、旧日能試では級外の語彙である。しかし、「雨水」は音韻変化の知識（2級レベル）があれば認知でき、「貯水槽」は「雨水を貯める（2級）ための」と修飾されているので、「貯水槽」が何であるかはわかる。「ディスカッション」は、本文内の「言葉（2級）で表現する（2級）、意見（2級）」

を言う、相手（2級）は聞いてくれる」の文から意味の推測は可能である。「微生物」は、「生物（2級）」の派生語の知識があればある程度の意味はつかめる。（今回の3アイテムに限ってであるが）、上述した山本（2005,p.125）の分析を裏付けている。

「情報に対する態度」のヴァリエーションの「テキスト外」には、常識や背景知識など、テキスト外の知識をどう使うかが挙げられた。肯定的な処理では、常識や背景知識使用が解答を得るためにプラスとなり、否定的な処理ではマイナスとなっている。双方で挙げられた「②常識と違う」例は、肯定的な処理では「テストであるから、常識で答えられるものではないだろう」という予測であり、否定的な処理では「常識と違うため混乱してしまった」という例である。前者は聴解テスト・アイテムで常識を問うことはしない¹⁵ことを知っており、後者はそれに慣れていないと考えられる。肯定的処理に、「⑥惑わし選択肢、畏を推測する ⑦テスト・テクニックを使う」という例があるが、テスト聴解対策の意識が強いことがうかがえる。「否定的に処理する」では、トピックに言及している例「興味がない、苦手だ、専門外だ、頻度が少ない」があり、これは個人差のため、トピックが広く網羅されることを期待するしかない。しかし、「遭遇頻度が少ない」という例に、「留学生は1日の時間で授業の占める割合が多いため日常的な雑談などはあまりしないことから、友人との会話展開の方が難しく感じる」という対象者がいた。また、今回使用したアイテムの1番「水の貯水槽」（p.121）のような話題で、友人とおしゃべりしたことはほとんどないという意見も挙げられた。これらのことから、対象者がテスト・アイテムを聞きながら、自分の聴解環境に置き換えているということがわかった。日留試の聴解テストの場面は、回を重ねるごとにアカデミック場面が増加しているため（4章図 4.5「場面の割合の推移」p.60 参照）、留学生にとってより身近になりつつあると言える。今後、友人との会話や、ライフ場面のトピックや状況が、留学生にどの程度親近感があり、日常的に接するかの検証が必要ではないかと考えられる。否定的処理では、「⑤メモに集中しすぎる ⑥解答を確信した後聞かない」といった聞き方の例や、「⑦集中力が低下する ⑧緊張 ⑨課題が予想外」など心理面の例も挙げられた。

「キーワード」は、談話内容理解を助けたり、解答を得るためのヒントとなるものである。肯定的処理では、「①トピックから推測する ②プリタスクから推測する」のように本文が始まる前から準備している例や、機械的に判断する「③頻度が高い」語彙に注目した例、「④内容理解から推測する」といった全体理解の中からの抽出例が見られた。一方、否定的処理では、「①キーワードの勘違い」のみであった。これは「ディベート」を音が近い「divide」と考えたもの、「アスファルト」を最近ニュースで見た「アスベスト」と勘違いしてしまったものである。両者とも、その語彙の意味に自信がなかったが（表 6.4の「不安材料」p.131）、本文内容に大きな齟齬を感じなかったため（表 6.4の「不安を解決する」p.131）、最後まで聞いていた。実際の会話場面であれば、聞き返しや質問が可能だが、テストの場合それができないため、誤答を選んでしまった。

最後に、「不安材料」について分析する。不安という感情は、わからない語彙や表現があったときに生じる。テスト受験者は、本文を聞きながら不安材料を並行的に処理しなければならない。不安が生じたときにとる行動は、「不安を解決する」と「不安を解決できない」の2つである。

「不安を解決する」例として特徴的なものは、①から④にみられる「推測する」という行動である。②の「母語から」と③の「英語から」は、漢字圏や韓国語圏は漢字の音読みの音が似ていること、日本語のカタカナ語が英語由来のものが多いことから推測している。英語以外のロマン語圏の対象者は、母語に近い音の語彙を、英語ではなく母語から推測していた。「④語彙分析から推測する」は、「非～」の

¹⁵ 常識で解答できる聴解テスト・アイテムの場合、常識という世界知識を測定しており、聴解能力を測定していない。

ような接頭語から「～ない」という意味となるだろうと推測することである。「⑤トピックの背景知識を使う」は、例えば、石けんで手を何回も洗うのはよくないことを知っていたので、本文を聞かなくても正答を推測した例である。「⑥回避する」は、わからなくても気にしない、あるいはその後本文内で説明などがあるかもしれないので、その場では回避するという一時的な行動をとった例である。

「不安を解決できない」例として、「①集中力が分散する ②その後の本文を聞けない ③思考が止まる」の3つがある。これらは本文を聞かなければならないと理解していても、わからない語彙などがあると、思考力を不安解決のほうへ向けてしまう例である。①では本文を聞く行動と不安解決行動を並行して行っているための弊害であり、②は不安解決にのみ意識が行ってしまっている。③ではわからないという状態から抜け出せず、思考が止まってしまっている。横山（2005a, p.52）は、③の行動を「放置」とし、問題解決方法の1つとしているが、テストの場合、質問や確認ができないため問題解決とはならない。不安が解決できない場合、その後の本文、ポスト・クエスチョン、選択肢が聞けなくなり、さらに不安が増大し、次のアイテム時にも不安を抱えたままの状態（負のスパイラル）が続く対象者がいた。

6.4.5 本文終了時

表 6.5 本文終了時の概念と聴解行動のヴァリエーション⑤（留学生設問あり）

714	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
本文 終了 時	解答の 言語化	解答を明確 に決める	解答を決める	①自分の言葉で決める ②確信しているとき ③早く正解がわかれば次の問題の準備ができる	
			選択肢まで待つ	①だいたい答えがわかっても選択肢を聞いて決める ②解答に自信がないとき ③選択肢を聞きながら考える ④本文は内容把握のみする	

表 6.5 は、本文終了時からポスト・クエスチョンまでの数秒間における聴解行動である。「解答の言語化」は、選択肢を聞く前に解答を明確に決めるかどうかである。「解答を決める」の、「①自分の言葉で決める」は、聞いたテスト・アイテム内にある文章のままではなく、要約した文章や簡潔な文章などをさす。「②確信しているとき」は、確信しているときのみには解答を決め、そうでないときは決めないという流動的な行動であり、「③早く正解がわかれば次の問題の準備ができる」は、聴解テスト形式を熟知した計画的な行動である。「選択肢まで待つ」の、「①だいたい答えがわかっても選択肢を聞いて考える」「②解答に自信がないとき」は、漠然と解答は形作られているが、最終決定は選択肢を聞きながら決める例である。これに対し、「③選択肢を聞きながら考える」は、選択肢を聞く前は、解答が明確になっておらず、選択肢を聞いて最良のものを選択肢から選ぶ。③には、選択肢と自分の解答のギャップを避けるため、選択肢を聞くまでは解答は漠然としておいたほうがよいと、計画的にしている対象者もいた。「④本文は内容把握のみにする」は、本文内容を理解していれば、選択肢から正答を選ぶことができるという考えである。対象者が本文を聞き終えた段階では、本文の文言のままの直接正答が多い。5章のスク립ト分析の結果、正答選択肢で本文と同じ表現や語彙を使用した直接正答は37.5%（5章表 5.9「本文から得た解答と選択肢の関係」p.113を参照）であった。自分の言葉で解答を決めていても、選択肢で表現や語彙などが言い換えられる、あるいは全体理解のアイテムの場合、正答の文章が推測しにくいことを、対象者は経験上知っており、その対策をしていると言える。

6.4.6 ポスト・クエスチョン

表 6.6 ポスト・クエスチョン時の概念と聴解行動のヴァリエーション⑥（留学生設問あり）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
ポスト トQ	確認 ¹⁶	もう1度聞く	再確認する 確認する		曖昧である

ポスト・クエスチョンは、プリタスクに比べ重要性を感じている対象者は少なかった。結果としてポスト・クエスチョンは、一度設問時に聞いているため、再確認する程度であった。設問時に課題がわからなかった対象者は、ここで課題を確認する。ただし、プリタスク時に曖昧だった対象者の中には、もう一度聞いても曖昧であり、課題を2度聞くことで理解度があがらなかった者もいた。プリタスクもポスト・クエスチョンも重要性が低いと感じている対象者に、「課題がわからなくても選択肢を聞いたときに、意味をなさない選択肢があるので簡単に削除できる」という例があった。多肢選択肢の作成が難しいことを物語っている。

6.4.7 選択肢

本研究は、解答を得るまでの行動に注目した研究であるため、選択肢から正答が選ばれたか否かではなく、正答と確信・確定したか否かという2つの概念で分析した。選択肢を聞いている段階で、本文終了時に得た解答と照らし合わせて一致した場合は「確信」に変わり、曖昧だった場合やぼんやりとした解答イメージが選択肢の1つと一致できた場合は「確定」とした。（表 6.7 次頁）

日留試の聴解テストの選択肢は短くとも形式名詞であり、1語彙のみのものはない¹⁷。よって、対象者が解答と考えているものと100%一致する選択肢は少ないため、「①自分が決めた答えから推理し」、「確信・確定」に至った行動である。「②錯乱肢だとわかる」は、正答選択肢を確信・確定する前後に関わらず、明らかに誤答と「確信・確定」している例である。

¹⁶ 7章で聴解テスト・アイテムの設問がない場合の分析を行うが、「確認」に関して、設問がある場合とない場合に違いが見られたので、設問がある場合を「確認1」とし、設問がない場合を「確認2」とした。

¹⁷ 今回使用したテスト・アイテムの1番（2006年第2回16番p.121参照）の正答選択肢は、本文内の語彙と表現とほぼ同様の直接解答。2番（2006年第2回15番p.122参照）の正答選択肢の語彙は、ほぼ本文と同様の直接正答。3番（2006年第2回20番p.122参照）の正答選択肢は、統語的な言い換えがされている。

表 6.7 選択肢時の概念と聴解行動のヴァリエーション⑦（留学生設問あり）

7/74	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
選択肢	確信・確定	得た情報との一致度が高い	得た情報と照合する	①自分が決めた答えから推理する ②錯乱肢だとわかる	
	葛藤	解答を選ばなければならない	ストラテジーを使う 選択肢に問題があると判断する	①背景知識や常識から推測する ②常識だから誤答と考える ③消去法 ④唯一選択をする（唯一わからない選択肢、1つしかわからない選択肢、意味のない選択肢） ⑤本文の語彙と一致、類似、相違から判断する ⑥わからなかった語彙を音から漢字に置き換える（非漢字圏） ⑦統語機能から判断する ⑧本文内容から推理する ⑨選択肢から本文内容を推測する ⑩わからない語彙を回避する ⑪テスト・テクニックを使う ⑫勘、当て推量、運 ①選択肢同士が似ている ②曖昧だ ③言い換えがされている ④適当な答えがない ⑤正答が複数ある	解答放棄 解答放棄

次に、「葛藤」だが、「ストラテジーを使って」解答を導き出した行動と、「選択肢に問題があると判断して」解答を導き出せなかった行動があった。「ストラテジーを使う」は、「①背景知識や常識から推測する」場合と「②常識だから誤答と考える」場合の既有知識利用の相反した例である。「③消去法」は、複数の選択肢から正答を選ぶテスト形式に見られるストラテジーであり、選択肢1つ1つを聞きながら正答ではないと判断したものを消去し1つに絞って行く方法である。正答選択肢に絶対的な自信はなく、「解答ではないだろうか」「解答ではないかもしれない」という様々な心理状態であると考え、「葛藤」とした。「④唯一選択をする（唯一わからない選択肢、1つしかわからない選択肢、意味のない選択肢）」の「唯一わからない選択肢」は、4つのうち3つの選択肢に対して、消去するほどの情報を持っていないことから、本文理解が不十分であったと考え、残った1つの選択肢が理解できない＝正答であると判断した。「1つしかわからない選択肢」は、正答の確信はないのだが、他の3つの選択肢を消去できる方法がないため選んでいる。「意味のない選択肢」は、すべての選択肢に対して正答であるかどうかの判断がつかないため、1つだけ異質な選択肢が正答ではないかというストラテジーである。「⑤本文の語彙と一致、類似、相違から判断する」は、本文内の語彙と選択肢の語彙が一致あるいは類似していた場合は、正答の可能性が高く、相違の場合は誤答であると判断している。「⑥わからなかった語彙を音から漢字に置き換える」であるが、言及したのは非漢字圏の対象者であった。しかし、本文を聞いているときに母語に置き換えている対象者がいることから、漢字圏の対象者も同様のことをしていると考えられる。本文内で使用された語彙が選択肢で言い換えられる場合、選択肢のほうが簡単な語彙でなければならない¹⁸。⑥の行動をとった対象者は、本文を聞いている際に、わからなかった語彙を音で記憶あるいはメモをしておき、選択肢で提示された語彙を聞いて、自分が持っているリファレンス辞書（田中他 1986, p.133を参照）から同義語・類義語を探し出し、わからなかった語彙と一致するものを見つけ、漢字にして初

¹⁸ 聴解テストは、本文理解を問うものであり、選択肢理解を問うものではない。よって、選択肢の語彙・表現は本文のものより簡単でなければならない。

めて意味を知ったのである。「⑦統語機能から判断する」は、例えば、フリタスクで「〇〇の意見は何か」と問われ、選択肢に「～だと思ふ」という文章があった場合、機能としては意見なので課題の解答となる可能性が高いと考えた。「⑧本文内容から推理する」は、自分が考えた解答と選択肢が合致しなかったため、本文内容を思い出し、近い選択肢を選んだ。「⑨選択肢から本文内容を推測する」は、選択肢の情報から本文内容理解を補強し、解答を選んでいる。「⑩わからない語彙を回避する」は、本文のように文脈からは推測できず、考えている時間的な余裕もないため、わからない語彙を回避してしまう。「⑪テスト・テクニックを使う」は、「残りの3つの選択肢を聞かなくなるので最初の選択肢は正答になりにくい」、「本文内に節約があり選択肢にも節約が出て来たのでひっかけだとすぐわかる」のような例があり、今までの受験で得た知識とテスト作成側の考えを考慮している。今回、⑪の例から、テスト受験者とテスト作成者の思惑のかけひきをしているような印象を受けた。大規模テストでの多肢選択法は、採点時間の短縮や採点ミスの軽減のメリットがあり、今後も使用されるであろう。しかし、テスト作成者は、テストで測定する真の能力は、選択肢から正答を選ぶ能力ではないことを忘れてはならない。「⑫勘、当て推量、運」に頼る受験者がいることは、テストの限界の1例である。

今回の対象者は、正答が明確に絞れない場合「選択肢に問題があると判断」した。対象者は解答だと思った言葉や文章を簡単にメモしているか、ぼんやりと記憶しているにすぎない。そのため、自分が確信しているあるいは不明瞭なイメージの解答と選択肢との間で、葛藤が起こる。「①選択肢同士が似ている ②曖昧だ ③言い換えがされている ④適当な答えがない ⑤正答が複数ある」は、すべて本文から得た解答と選択肢にずれがあるために起こっている。その原因として、正答選択肢が完全なパラフレーズになっていない場合と、抽象度をあげすぎた言い換えの2つが挙げられる。選択肢から選ぶことができない場合は、解答を放棄してしまう対象者もいた。解答放棄をした対象者の中には、ストラテジーを使用しても正答が明確に得られなかった対象者と、選択肢に問題があると判断して解答放棄をした対象者の2タイプが見られた。

5章のスク립ト分析でも述べたが、よい選択肢を作成することは難しい。しかし、大規模テストで多肢選択法が利用される限り、よい選択肢を作り続けなければならない。今回、協力をしてくれた留学生（対象者）が、選択肢から正答を選ぶ際に、様々な心理的葛藤と戦っている事実が明らかになったことが、今後のテスト・アイテム作成に役立つのではないかと期待している。

6.5 日本人大学生の聴解行動の分析結果と考察（設問あり）

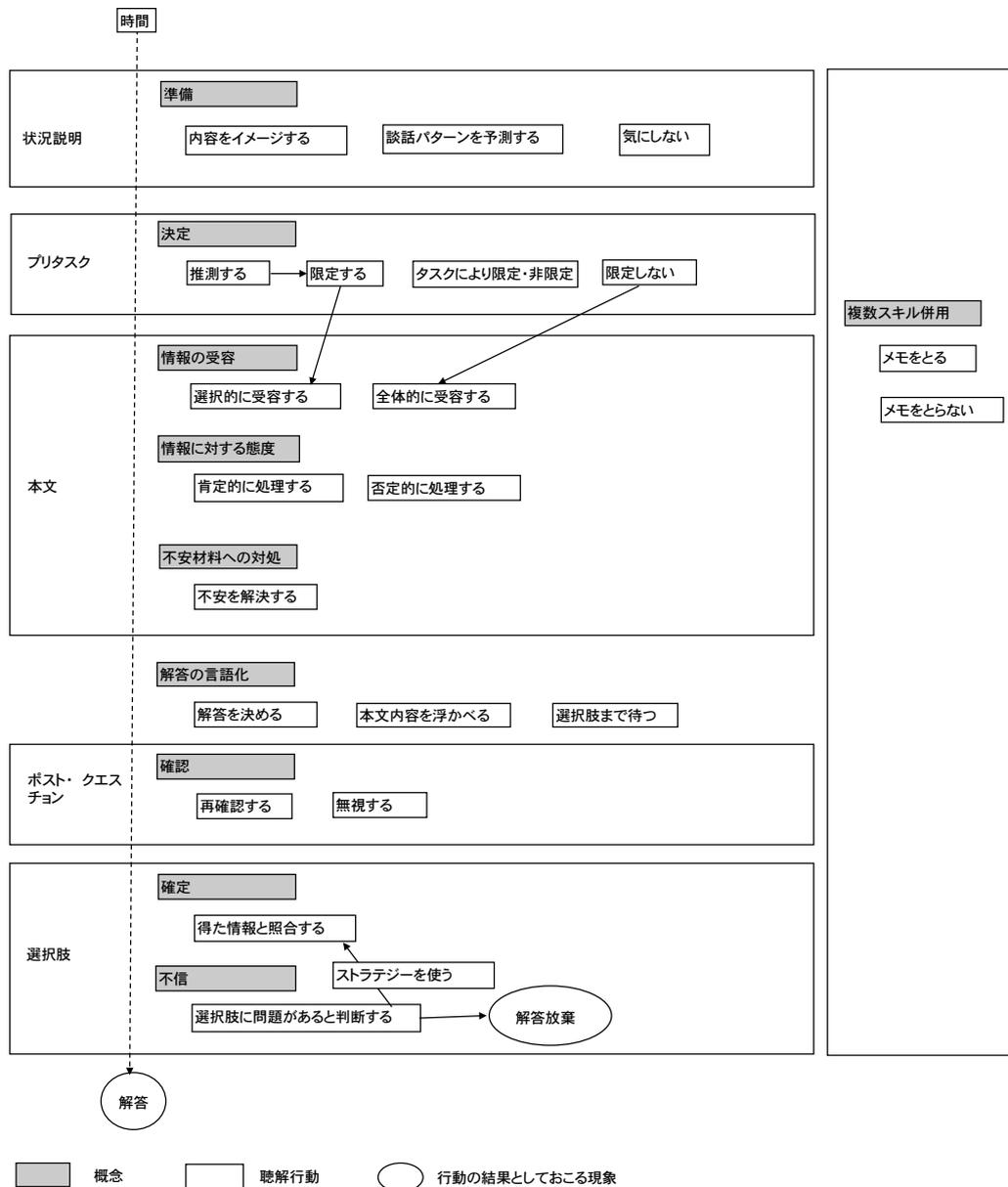


図 6.4 日本人大学生の聴解行動（設問ありのアイテムに対する行動）

図の時間的な流れおよび使用している記号は、留学生の聴解行動（図 6.3 p.125）と同様である。

留学生の図 6.3 と比較すると、テスト・アイテムの各部分（状況説明・プリタスク・本文・ポスト・クエスチョン・選択肢）での行動間に次の行動に繋がる「→」が少ない。これは留学生に見られた「わからない」や「不安を解決できない」などがなかったためである。日本人大学生は心理的に負の材料を持たず本文を聞けることが留学生と異なる。

以下、アイテムの各部分の聴解行動を上述した留学生と比較しつつ詳細に分析・考察する。

6.5.1 全体

表 6.8 は、日本人大学生の設問があるアイテム全体に対して見られた行動と、そのヴァリエーションである。

表 6.8 アイテム全体の概念と聴解行動のヴァリエーション①（日本人大学生設問あり）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
全体	複数スキル併用	聞きながらメモをとる	メモをとる	①記憶のためにする ②課題やキーワードを書く ③整理する ④聴解テストの経験から	
			メモをとらない	①記憶にある ②会話についていけなくなる ③集中しにくい ④解答しにくくなる	

1 アイテムの最初から最後まで、全体において見られた行動は、「メモをとる」のみである。9名中7名がメモをとる行動をしている。留学生は、すべてのアイテムでメモをとる、メモをとらない、アイテムによってとったりとらなかったりする、と3パターンが見られたが、日本人大学生は1人が最初のアイテムのみメモをとらなかったが、残りの2アイテムはメモをとり、それ以外の7名はすべてのアイテムでメモをとっていた。

ヴァリエーションは、4つである。「①記憶のためにする」は、重要だと判断したり、気になった語彙を書き留めることで、記憶にかかる負荷を減少させ、継続して聞くことができる。「②課題やキーワードを書く」は設問から解答と思われる、あるいは関連語彙を書き留めておく。「③整理する」は、本文の内容を要約し、整理して、理解の補助としている。「④聴解テストの経験から」は、英語の聴解テストでメモが解答する際に役立ったという経験を持っていた。

「メモをとらない」対象者は2名だけであったが、「①記憶にある ②会話についていけなくなる ③集中しにくい ④解答しにくくなる」の4つのヴァリエーションがあった。①は記憶できなければ、メモをとる行動が変わる。日留試の聴解アイテムの長さは、1分～2分程度の短い談話のため、記憶可能な長さと言ってよい。ヴァリエーション②③④は、メモをとることが聴解の妨げになるためであるが、今回は、聴解テストという状況下、課題を聞いて解答を得るための聞き方であり、集中して聞くための行動であった。実際の講義や授業では、聴解の妨げになるのでノートをとらない行動を選択することはないだろう。また、実際の講義は、冗長性もあり、余談もあり、ノートをとる時間がある。

大学生生活では、ノートをとりながら講義を聞く聞き方と、おしゃべりやテレビから流れてくるニュースを単に聞くだけの聞き方がある。ある対象者は、「友人とのおしゃべりのアイテムはさほどメモをとらなかったが、先生の講義の内容のアイテムのときは、自然とメモをとらなければならないと感じた。普段の授業の影響かもしれない」と話している。留学生は、聴解テストを強く意識し、解答を得るための聞き方に終始している例だけであったが、日本人大学生は、普段の生活と同じ聞き方が、テストでも自然とできていた。聴解テスト・アイテムを利用した聴解練習で、講義場面ではメモをとる、おしゃべり場面ではメモをとらないと、普段の生活の聴解に近い聞き方の練習をすることが、留学生に有効ではないかと考えられる。

留学生と日本人大学生は「複数スキル併用」とその聴解行動が同じため、両者の比較をしながら分析・考察を行う（表 6.8.1 次頁）。両者とも聴解行動のヴァリエーションに大きな違いはないが、留学生の「メ

メモをとる」ヴァリエーションに【留学生】「①推測の手がかりにする」があるが、日本人大学生にはない。日本人大学生は、今回使用した3アイテム内にわからない言葉がなかったからである。

表 6.8.1 アイテム全体の聴解行動のヴァリエーション①（留学生と日本人大学生設問ありの比較）

聴解行動	ヴァリエーション（留学生設問あり）	ヴァリエーション（日本人大学生設問あり）
メモをとる	①推測の手がかりにする（語彙） ②記憶のためにする ③整理する ④集中する	①記憶のためにする ②課題やキーワードを書く ③整理する ④聴解テストの経験から
メモをとらない	①本文が開けなくなる ②本文に集中できない	①記憶にある ②会話についていけなくなる ③集中しにくい ④解答しにくくなる

6.5.2 状況説明

表 6.9 状況説明時の概念と聴解行動のヴァリエーション②（日本人大学生設問あり）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
状況説明	準備	本文を聞く前に推測する	内容をイメージする	①談話内容をイメージする ②内容からプリタスクの方向を予想する	
			談話パターンを予測する	①ダイアログかモノログか	
			気にしない	①プリタスクのほうが大切 ②意識的に何も考えない	

留学生と同様、状況説明時の行動概念は「準備」であるが、聴解行動は留学生と少し異なっている。留学生と日本人大学生を比較した表が表 6.9.1 である。

表 6.9.1 状況説明時の聴解行動のヴァリエーション②（留学生と日本人大学生設問ありの比較）

聴解行動	ヴァリエーション（留学生設問あり）	聴解行動	ヴァリエーション（日本人大学生設問あり）
キーワードを推測する	①大切な語彙を推測する ②常識的な解答になると推測する	内容をイメージする	①談話内容をイメージする ②内容からプリタスクの方向を予想する
談話パターンを予測する	①ダイアログかモノログか	談話パターンを予測する	①ダイアログかモノログか
		気にしない	①プリタスクのほうが大切 ②意識的に何も考えない

まず、「内容をイメージする」で、【日本人大学生】「①談話内容をイメージする ②内容からプリタスクの方向を予想する」が挙げられた。「①談話内容をイメージする」は、今回使用した3アイテムの状況説明が「誰が何について話している」だったため、本文内容を事前にイメージすることができたようである。留学生は、この段階で、本文内容まではイメージしておらず、キーワードや解答などの推測を行っていた。【日本人大学生】「②内容からプリタスクの方向を予想する」の例に、「雨水の取り組みだから、ためる方法とかを聞かれる」と述べた対象者がおり、テスト・アイテムが始まった段階でプリタスクと本文内容までもイメージできると理解の助けになるだろう。留学生に聴解テスト・アイテムの状況説明

を聞かせ、キーワードを3つ挙げる、会話を作ってもらうなど、これから話される内容の予測をする練習が日本語学習に役立つ可能性がある。

「談話パターンを予測する」は、「①ダイアログかモノログか」であり、留学生と同様である。「男女の会話の場合は、1人の会話に誰かが入ってくると予測できるので、突然違う声が入ってくるのではと不安に思いながら聞く必要がない」という例があった。しかし、「先生が話している」という状況説明から、「男性が話すと思っていたら（男性教授が多いため）、女性で、しかもかわいらしい声だったので学生が話していると勘違いし、混乱した」という例もあった。大規模テストは、不特定多数の人が受験するため、男女差や地位、貧富の差などが談話内に影響を及ぼさないよう極力排除される。しかし、実際、日本では女性の社会進出が遅れており、大学講師は男性の割合が多い。日本人大学生だからこそ、「先生」を男性と思い込んだのだろう。女性の上司、男性の保育士のような役割で多くのテスト・アイテムを作題することは性の公平性を保つためにはよいが、実際の日本社会とのずれがあるため、過剰にすべきではない。全体的に会話に対して「同じ言葉を何回も使うためわかりやすい」「確認しあう」「1人が質問してもう1人が答える」など、わかりやすいと肯定的な意見が多かった。モノログは、「感情がなくどこが大切かわかりにくい」「全体が流れてしまい最後にならないと大切な部分がわからない」と否定的であった。一方、会話に対しても、「話し方が違うということに気をとられて集中できない」という例があり、必ずしも、日本人にとって男女の会話が聞きやすいわけではないようだ。留学生から「日本の大学生活では、友人とのおしゃべりなどの日常的な場面のほうが少なく授業関連のほうが多い」ということが挙げられたが、日本人大学生は大学生だからといって授業場面の割合が多いというわけではないようである。留学生は、日本人の学生に比べ、日本語の勉強をしていたり、課題をするにも時間がかかり友人との時間が少ないと考えられる。また、長い休みには自分の国に帰国する学生もいる。留学生は日本人大学生の日常生活と全く同じ過ごし方ではないことを考慮し、日留試で測定する生活スキルはどのような場面が適切か、学習スキルは日本人大学生と同様でよいのかなどの検討が望まれる。

「気にしない」という概念は留学生には見られなかった。ヴァリエーションとして【日本人大学生】「①プリタスクのほうが大切 ②意識的に何も考えない」とあり、日本人大学生にとって、状況説明の重要性は低い。

留学生は、状況説明時に、聴解テストを強く意識するが、日本人大学生は談話内容、談話の流れをある程度推測する程度である。日留試の聴解テストは日本人にとっては簡単であるという意識もあるからだろう。

6.5.3 プリタスク

プリタスク時の行動としては、留学生と同様、「推測する」「限定する」「プリタスクにより限定するまたは限定しない」「限定しない」の4つが見られた（表6.10次頁）。

まず、「推測する」を分析する。「①本文内容を推測する」は、例えば、状況説明で「ある町の雨水をためる取り組みについて」というトピックだけだが、プリタスクでは、「雨水をためる一番の目的は何か」であり、本文で「雨水をためる目的について話す」と推測できる。「②どちらの話者が解答を言うか推測する」は、例えば「女子学生は、・・・だと言っていますか」というプリタスクから、女性が解答を言うだろうと推測できる。「③解答とならないものを推測する」は、例えば、「実験から何がわかるか」が課題であれば、通常解答は「実験値」と考えられるが、テストであるため直接実験値を問うことはないと考え、それ以外何を問われるのかと推測する。

推測したあとには、限定するか限定しないかを決めて聞く。「限定する」の「①解答部分を探す ②会話の場合プリタスクで限定されたほうに集中する」は、留学生と同様の行動であり、解答に集中する聞

き方である。

表 6.10 プリタスク時の概念と聴解行動のヴァリエーション③（日本人大学生設問あり）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
プリタスク	決定	解答になる部分に焦点をあてて聞くか、全体を聞くかを定める	推測する	①本文内容を推測する ②どちらの話者が解答を言うか推測する ③解答とならないものを推測する	
			限定する	①解答部分を探す ②会話の場合プリタスクで限定されたほうに集中する	
			プリタスクにより限定するまたは限定しない	①プリタスクによる	
			限定しない	①経験上全部を聞く（英語試験） ②（会話）両方の話者を均等に聞く ③内容全体を聞く	

「プリタスクにより限定するまたは限定しない」も留学生と同様、「プリタスクにより」部分を聞いたほうがよいのか、全体内容を把握したほうがよいのかを決定する。

「限定しない」は「①経験上全部を聞く（英語試験） ②（会話）両方の話者を均等に聞く ③内容全体を聞く」の3つのヴァリエーションである。アイテム全体に見られた「メモをとる」という行動も「今までの聴解テストの経験から」というヴァリエーションがあったが、日本人大学生は言語テストであることから、自らが受けてきた外国語テストでとった行動を行っていることがわかる。

以下は、留学生と日本人大学生の聴解行動の比較をしながら分析・考察を行う（表 6.10.1）。

表 6.10.1 プリタスク時の聴解行動のヴァリエーション③（留学生と日本人大学生設問ありの比較）

聴解行動	ヴァリエーション（留学生設問あり）	ヴァリエーション（日本人大学生設問あり）
推測する	①キーワードを推測する ②本文構造を推測してどこに大切な部分があるか予測する ③常識的解答になると推測する	①本文内容を推測する ②どちらの話者が解答を言うか推測する ③解答とならないものを推測する
限定する	①解答部分を探す ②会話の場合プリタスクで限定されたほうに集中する ③全部聞こうとすると聞き逃す	①解答部分を探す ②会話の場合プリタスクで限定されたほうに集中する
プリタスクにより限定するまたは限定しない	①プリタスクによる ②大切な部分の比重は重くするが全体を聞く	①プリタスクによる
限定しない	①1語彙で内容が変わることがあるから全部に注目する ②談話内に解答が散在している ③質問の答えばかりに注目すると本文内容把握がおろそかになる ④とりあえず全部聞く	①経験上全部を聞く（英語試験） ②（会話）両方の話者を均等に聞く ③内容全体を聞く
なし	①設問がわからなくても本文でわかる ②簡単すぎて油断 ③形式に慣れていない	

プリタスクにおける聴解行動は、留学生も日本人大学生も同じである。留学生はここで「聞き逃す」という行動が生じたが、日本人大学生は母語という有利さと、プリタスクが大切だという意識も強かったため、聞き逃すことはなかった。留学生の分析の際、「限定して聞く」と「全体を聞く」弊害について述べたが(6.4.3「プリタスク」p.129を参照)、日本人大学生には見られなかった。「推測する」場合、留学生が「キーワード、談話構造」を推測したのに対し、日本人大学生は「本文内容、話者、解答」を推測した。留学生は、キーワードという部分情報と談話の骨組みを推測し、肉付けを本文を聞きながらしていく形である。日本人大学生は全体像と解答の担い手と不必要情報を推測し、本文を聞きながら拾い出す形である。言い換えれば、留学生はボトムアップ的、日本人大学生はトップダウン的とも言える。

6.5.4 本文

表 6.11 本文時の概念と聴解行動のヴァリエーション④(日本人大学生設問あり)

7行目	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
本文	情報の受容	解答のための情報を得る	選択的に受容する	①解答のための情報として認知する ②事前推測を訂正する ③不必要な情報は捨てる ④解答を得たあとはあまり集中しない	
			全体的に受容する	①内容把握する ②解答があった後も正確な情報を得るために最後まで聞く ③内容を追う	
	情報に対する態度	難易に関わる要因から情報処理をする	肯定的に処理する	言語形式・談話構造 ①話し方が明確だ ②統語的機能がわかる ③ディスコース・マーカをヒントにする ④構成が明確だ ⑤論理的な構造だ ⑥解答が明確に提示される テキスト外 ①知識は使わない ②常識は使わない ③テストの談話構造と判断する ④複数解答候補から絞る ⑤解答は確定せず漠然としておく ⑥トピックから解答を推測する ⑦トピックが身近だ ⑧情報が多い キーワード ①プリタスクから推測する ②全体把握のために拾う	
			否定的に処理する	言語形式・談話構造 ①発音が不明瞭 ②音声予測の間違い ③解答確信後に逆接で混乱する ④情報の時間差が大きい ⑤解答が不明確だ テキスト外 ①常識と違う	
不安材料 ²	記憶に残りにくい	不安を解決する	①メモする		

表 6.11 は日本人大学生が本文時にした聴解行動の概念とヴァリエーションである。「情報の受容」「情報に対する態度」「不安材料」の3つの概念があるが、留学生の「不安材料1」の定義は「わからない語彙や表現がある」であり、日本人大学生に見られた「不安材料2」の定義は「記憶に残りにくい」と異なっている。以下では、3つの概念ごとにわけ、留学生と日本人大学生を比較しながら、分析・考察を行う。

表 6.11.1 本文時の聴解行動のヴァリエーション④（留学生と日本人大学生設問ありの比較）

「情報の受容」

聴解行動	ヴァリエーション（留学生設問あり）	ヴァリエーション（日本人大学生設問あり）
選択的に受容する	①解答のための情報と認知する ②不必要な情報は捨てる ③解答を得た後はあまり集中しない	①解答のための情報として認知する ②事前推測を訂正する ③不必要な情報は捨てる ④解答を得たあとはあまり集中しない
全体的に受容する	①内容把握する ②全体内容を要約する ③最後に変わることがある ④解答は最後にくることが多い ⑤常識と違うことを言う可能性がある	①内容把握する ②解答があった後も正確な情報を得るために最後まで聞く ③内容を追う

まず、「情報の受容」は、「選択的に受容する」か「全体的に受容する」かで、留学生と同様である（表 6.11.1）。「選択的に受容する」のヴァリエーションは、「①解答のための情報として認知する ②事前推測を訂正する ③不必要な情報は捨てる ④解答を得たあとはあまり集中しない」の4つである。①③④は留学生と同じ行動だが、「②事前推測を訂正する」は日本大学生のみに見られた。日本人大学生は、プリタスク時に本文内容を推測していたため、本文を聞きながらその推測を正しく訂正している。「全体的に受容する」場合は、ヴァリエーションとして「①内容把握する ②解答があった後も正確な情報を得るために最後まで聞く ③内容を追う」の3つである。留学生が「③最後に変わることがある ④解答は最後にくることが多い ⑤常識と違うことを言う可能性がある」というテスト・アイテム対策の聞き方をしているのに対し、日本人大学生は、テスト・アイテムによく見られる情報提示の特徴は意識せず、全文を聞くことで正確な情報が得られるとコメントしている。

次に、「情報に対する態度」は、「肯定的に処理する」「否定的に処理する」で、これも留学生と同じである（表 6.11.2 次頁）。言語形式・談話構造の、音に関するものでは、肯定的処理の「①話し方が明確だ」、否定的処理の「①発音が不明瞭 ②音声予測の間違い」が挙げられた。「①話し方が明確だ」は、「不自然に明確である（実際の会話はこれほど明確ではないので）」「通常の話し方であれば、考えながら話すため文章の中に違う文章が入ったりするが、そういうものがないので聞きやすかった」と、肯定的に捉えてはいるものの、テスト・アイテムが、日常的な聴解より聞きやすく加工されていることを指摘している。IELTS では、英語の発音の違い（地域差、位相）の聞き取りを評価項目として挙げている（British Council, 2003, *IELTS handbook*）が、日留試の聴解テストでは、フィラー以外、言い間違いや訛、方言、位相などは見られない。英語は、イギリス英語、アメリカ英語、オーストラリア英語、インド英語など、多種多様な発音で使用され、発音が違うことも世界的に認知されている。日本語使用の大部分は、日本国内の日本語ネイティブに占められており、日本語発音の地域差はあるが、メディアでは、ほぼ東京共通語で統一されていると言ってよい。このような現状を考えると、訛や方言をテストに取り入れることは現実に即していない。ある程度長い談話での言い間違いは自然であるが、短い談話の場合、言い間違いが測定項目であるかのような印象を与えてしまう可能性もある。そうであれば、スクリプトを明確な発音で読んでいる現状が、妥当であると言えるだろう。しかし、日本人が聞いたときに、明らかに不自然であると指摘されたことは、聴解テスト・アイテムの課題としなければならない。否定的処理の「①発音が不明瞭」は、「微生物」が「未生物」に聞こえ、何のことも最初はわからなかったが、本文を聞いたら「微生物」のことを話しているとわかった例であった。豊富な語彙知識を持っているネイティブでも、聞こえた音を微調整しなければ、語彙の認知ができない。スピーカーを通した音の不明瞭さにどのように対処していくのか、解決すべき問題である。

表 6.11.2 本文時の聴解行動のヴァリエーション④（留学生と日本人大学生設問ありの比較）
「情報に対する態度」

聴解行動	ヴァリエーション（留学生設問あり）	ヴァリエーション（日本人大学生設問あり）
肯定的に処理する	<p>言語形式・談話構造</p> <p>①音声の特徴がヒント ②スピードがよい ③男女の会話なのでどちらが言うかわかる ④語彙が簡単 ⑤統語的機能がわかる ⑥ディスコース・マーカ―に注意する ⑦構成が明確だ ⑧論理的な構造だ ⑨解答が明確に提示される</p> <p>テキスト外</p> <p>①常識から判断 ②常識と違う ③背景知識から判断 ④話すときの習慣から判断する ⑤話し手の性格を推測する ⑥惑わし選択肢、罠を推測する ⑦テスト・テクニックを使う</p> <p>キーワード</p> <p>①トピックから推測する ②プリタスクから推測する ③頻度が高い ④内容理解から推測する</p>	<p>言語形式・談話構造</p> <p>①話し方が明確だ ②統語的機能がわかる ③ディスコース・マーカ―をヒントにする ④構成が明確だ ⑤論理的な構造だ ⑥解答が明確に提示される</p> <p>テキスト外</p> <p>①知識は使わない ②常識は使わない ③テストの談話構造と判断する ④複数解答候補から絞る ⑤解答は確定せず漠然としておく ⑥トピックから解答を推測する ⑦トピックが身近だ ⑧情報が多い</p> <p>キーワード</p> <p>①プリタスクから推測する ②全体把握のために拾う</p>
否定的に処理する	<p>言語形式・談話構造</p> <p>①スピードが速い ②発音が不明瞭だ ③母語と違うため（非漢字圏は統語的相違）・語彙がわからない・カタカナ語（漢字圏）が難しい ④統語機能がわからない ⑤談話構造に対するピリ―フと違う ⑥トピック展開が速い ⑦解答が不明確だ ⑧量が多い ⑨情報量が少なすぎる</p> <p>テキスト外</p> <p>①背景知識がない ②常識と違う ③間違った常識を使う ④トピックに興味がない、苦手だ、専門外だ、遭遇頻度が少ない ⑤メモに集中しすぎる ⑥解答を確信した後聞かない ⑦集中力が低下する ⑧緊張 ⑨課題が予想外</p> <p>キーワード</p> <p>①キーワードの勘違い（音から判断するため）</p>	<p>言語形式・談話構造</p> <p>①発音が不明瞭 ②音声予測の間違い ③解答確信後に逆接で混乱する ④情報の時間差が大きい ⑤解答が不明確だ</p> <p>テキスト外</p> <p>①常識と違う</p>

「肯定的に処理する」では留学生に比べ日本人大学生は「言語形式・談話構造」を情報とすることが少ない。「テキスト外」では留学生が「常識や背景知識を利用する」ことに対し日本人大学生は、テストであるため常識や背景知識の使用をせず、あくまでも談話内で話されている情報からのみで解答すると応えた。聞いた内容を、すべて完璧に理解できるという裏付けがあつての行動である。【日本人大学生】「③テストの談話構造と判断する」とは、惑わし情報が必ず入っているので、最後まで聞き、選択肢と照らし合わせるという例である。【日本人大学生】「②統語的機能がわかる」は、肯定否定、納得、驚きなどの表現から、談話理解がしやすいということである。これらは、接続詞のようなディスコース・マーカ―と同様に解答のヒントになる。留学生も、接続詞や助詞などのディスコース・マーカ―に注目していたが、日本人大学生は機能動詞にも注目しているようである。日本語は文末の動詞によって、アスペクト、ヴォイス、ムードなどを表現するため、動詞の機能は重要であり、日本語ネイティブが、普段

から動詞の機能に注目して聞いていることがわかった。談話構造の【日本人大学生】「③構成が明確だ ④論理的な構造だ」は、例えば、実験後結果が得られた後に意見を言うという流れや、会話では肯定否定が明確に言語化されているなどの例である。また、論理的構造だとスクリプト (script) 知識が使用できる¹⁹。

「否定的に処理する」の【日本人大学生】「②音声予測の間違い」は状況説明時、「先生が話している」と言っていたので、男性の声を予測していたが、女性の若い声だったため学生が話していると勘違いし、少し混乱したという例である。2人の会話では、男女が登場し、声から誰が発話したかをすぐ判断できるようにしているが、男性か女性かを状況説明で言う場合と言わない場合がある。職業名 (先生、医者、警察官など) の場合、男女は言わないようである。対象者でこのことに言及したのは1人だけだが、談話参加者が男女であることを設問時に述べる時間は1秒程度であることから、すべてのアイテムに性別を入れるか否かの検討があつてよいと思われる。職種や立場、性別や年齢などにより、言葉遣いが変わる日本語では、必要な情報であると考え。「女性の教授」と設問に入れるだけでよいので検討してもらいたい。【日本人大学生】「③解答確信後に逆接で混乱する ④情報の時間差が大きい」は、解答と思われた情報が何回も変わることに不満、必要情報から次の必要情報までに不必要情報が入りすぎるといふ例である。留学生も談話構造に対して「談話構造に対するピルーフと違う」という例があつたが、これらはテスト・アイテムとして人工的に作られ、談話を最後まで聞くような工夫や、錯乱肢になる情報を不自然に入れていたことが原因である。

留学生と日本人大学生両者が述べている「解答が明確・不明確」について考察する。本文内で、解答を明確に言っているか否かの問題もあるが、5章の選択肢の分析 (5.5「選択肢」 p.113 を参照) でも述べたように、本文を聞いた段階で得た解答と、正答選択肢の内容がずれている場合があることも一因である。解答があまりに明確であるとテストとしての難易度が下がる可能性が高く、どの程度で収めるかはアイテム・ライター課題である。「否定的に処理する」の「テキスト外」は留学生は様々なヴァリエーションをあげているが、日本人大学生は「①常識と違う」のみである。これは正答がわからなかったのではなく、「常識と違うことをテスト内容で使ってよいのか」という否定的な意見であつた。ネイティブは常識と違うことがわかるため問題はないが、「留学生が受験するテストに、常識と違う内容を出題することは、負の波及効果となるのではないか」という心配であつた。

表 6.11.3 本文時の聴解行動のヴァリエーション④ (留学生と日本人大学生設問ありの比較)

「不安材料」

聴解行動	ヴァリエーション (留学生設問あり)	ヴァリエーション (日本人大学生設問あり)
不安を解決する	①文脈から推測する ②母語から推測する ③カタカナ語は英語から推測する ④語彙分析から推測する ⑤トピックの背景知識を使う ⑥回避する	①メモする
不安を解決できない	①集中力が分散する ②その後の本文を聞けない ③思考が止まる	

最後に「不安材料」の比較をみる (表 6.11.3)。留学生の「不安材料 1」の定義は「わからない語彙や

¹⁹ 人は談話の流れにそつて変化していくことを、背景知識として持っている。この知識をスクリプトという。例えば、レストランに入る、席に座る、注文する、食事が運ばれる、食べる、会計をする、という流れである。

表現がある」であり、その不安を解決するために様々な情報を元に推測し、回避するというも行っている。しかし、日本人大学生に見られた「不安材料2」の定義は「記憶に残りにくい」ということで、解決するためには「①メモする」だけであった。日本人大学生にとって本文理解の不安材料は「記憶」だけであり、日留試の聴解テスト・アイテム談話は簡単に理解できる内容であったためである。

6.5.5 本文終了時

表 6.12 本文終了時の概念と聴解行動のヴァリエーション⑤（日本人大学生設問あり）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
本文終了時	解答の言語化	解答を明確に決める	解答を決める	①自分の言葉で決める	
			本文内容を浮かべる	①選択肢で迷わないようにする	
			選択肢まで待つ	①タスクと本文の情報が曖昧なので待つ	

表 6.12 は本文終了時に見られた日本人大学生の聴解行動の概念とヴァリエーションである。留学生の概念は日本人大学生の「解答の言語化」と同様だが、聴解行動が少し異なっているため、ヴァリエーションを比較しながら分析・考察する（表 6.12.1）。

表 6.12.1 本文終了時の聴解行動のヴァリエーション⑤（留学生と日本人大学生設問ありの比較）

聴解行動	ヴァリエーション（留学生設問あり）	ヴァリエーション（日本人大学生設問あり）
解答を決める	①自分の言葉で決める ②確信しているとき ③早く正解がわかれば次の問題の準備ができる	①自分の言葉で決める
本文内容を浮かべる		①選択肢で迷わないようにする
選択肢まで待つ	①だいたい答えがわかっても選択肢を聞いて決める ②解答に自信がないとき ③選択肢を聞きながら考える ④本文は内容把握のみする	①タスクと本文の情報が曖昧なので待つ

留学生と違う行動に、「本文内容を思い浮かべる」がある。これは、解答を自らの言葉で1文にしてしまうと、選択肢と違う表現になっていた場合は迷うので、概要をイメージしておき選択肢と照らし合わせたほうがよいという判断であった。「選択肢まで待つ」は、プリタスクの解答として明確な解答が得られなかったとき、このテスト・アイテムで期待される答えが何かを探すために選択肢を待つということだった。留学生には、常に選択肢を聞いてから正答を決めるスタイルの対象者もいたが、日本人大学生は個人の聞き方のスタイルではなく、本文内容の曖昧さからとった行動である。プリタスク時に、全体を聞くように決定した対象者の中には、本文終了時に明確な解答が得られていれば、最初に決めた内容把握のみに固執せず解答を言語化する場合もあった。留学生に比べ、本文終了時に解答を得ている確率は高く、本文内容がわからない、解答が全く得られていないということはない。

6.5.6 ポスト・クエスチョン

表 6.13 ポスト・クエスチョン時の概念と聴解行動のヴァリエーション⑥（日本人大学生設問あり）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
ポストQ	確認1	もう1度聞く	再確認する	①プリタスクの情報の再確認 ②解答に自信がなかった	
			無視する	①覚えていない ②聞いたが要らない	

「ポスト・クエスチョンはなくてもいいが、あればもう一度確認できる、解答らしきものが明確になかったように思うのでポスト・クエスチョンで確認した、あればテストらしい」という例が挙げられた。「無視する」は、不必要だという意見である。本文で全体内容把握をしながら聞いている対象者も、解答はほぼ絞れているので、ポスト・クエスチョンがなくても答えられるということで、全員「なくても大丈夫である」という意見であった。留学生も、ポスト・クエスチョンに対する重要性は感じていなかったが、日本人大学生はそれよりさらに必要性を感じていないという結果となった。ポスト・クエスチョンに関しては、留学生、日本人大学生とも必要性をあまり感じておらず、特別な聴解行動も見られなかったため、比較をしておいた分析は行わない。

6.5.7 選択肢

表 6.14 選択肢時の概念と聴解行動のヴァリエーション⑦（日本人大学生設問あり）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
選択肢	確定	得た情報との一致度が高い	得た情報と照合する	①自分が決めた答えと一致している ②言い換えと照合する ③消去法	
	不信	自分の解答との一致度が低い	選択肢に問題があると判断する	①選択肢が本文の要点ではない ②自分の解答と違う ③正確な答えがない ④正答が複数ある	解答放棄
			ストラテジーを使う	①自分の解答の先に何があるかを考え選択する ②作成者の立場で考える	

選択肢時の聴解行動には、留学生と日本人大学生には違いが見られたので、表 6.14.1 と表 6.14.2（次頁）で両者の比較をしながら分析・考察する。

表 6.14.1 選択時の聴解行動のヴァリエーション⑦（留学生と日本人大学生設問ありの比較）

「確信・確定」と「確定」

聴解行動	ヴァリエーション（留学生設問あり）	ヴァリエーション（日本人大学生設問あり）
得た情報と照合する	①自分が決めた答えから推理する ②錯乱肢だとわかる	①自分が決めた答えと一致している ②言い換えと照合する ③消去法

留学生には、選択肢を聞いた時点で、「確信・確定」と2つの概念が見られた。これは解答が絞られておらず、選択肢を聞いた段階で解答を得た場合は「確信」であり、既に解答が得られていた場合は「確定」としたからである。日本人大学生は全員、本文終了時に解答は得ていたため、「確定」のみとした。

聴解行動は「得た情報と照合する」で両者とも同様である（表 6.14.1 前頁）。日本人大学生のヴァリエーションは、「①自分が決めた答えと一致している ②言い換えと照合する ③消去法」の3つである。今回使用したアイテムの選択肢の語彙や表現などの言い換えがされていたため、②や③のような例が挙げられた。

表 6.14.2 選択時の聴解行動のヴァリエーション⑦（留学生と日本人大学生設問ありの比較）
「葛藤」と「不信」

聴解行動	ヴァリエーション（留学生設問あり）	聴解行動	ヴァリエーション（日本人大学生設問あり）
ストラテジーを使う	①背景知識や常識から推測する ②常識だから誤答と考える ③消去法 ④唯一選択をする（唯一わからない選択肢、1つしかわからない選択肢、意味のない選択肢） ⑤本文の語彙と一致、類似、相違から判断する ⑥わからなかった語彙を音から漢字に置き換える（非漢字圏） ⑦統語機能から判断する ⑧本文内容から推理する ⑨選択肢から本文内容を推測する ⑩わからない語彙を回避する ⑪テスト・テクニックを使う ⑫勘、当て推量、運	選択肢に問題があると判断する	①選択肢が本文の要点ではない ②自分の解答と違う ③正確な答えがない ④正答が複数ある
選択肢に問題があると判断する	①選択肢同士が似ている ②曖昧だ ③言い換えがされている ④適当な答えがない ⑤正答が複数ある	ストラテジーを使う	①自分の解答の先に何があるかを考え選択する ②作成者の立場で考える

留学生は選択肢を聞いた際、解答を「確信・確定」できないと、「葛藤（定義：解答を選ばなければならない）」と考え、「ストラテジー」を駆使して解答を選ぶ。解答が選べない場合は「選択肢に問題がある」と判断した。日本人大学生は、解答を「確定」できないと、「不信（定義：自分の解答との一致度が低い）」を抱き、まず「選択肢に問題があると判断する」。その後、テストであるため、解答を選ぶために「ストラテジーを使う」（表 6.14.2）。日本人大学生は解答を得ていない対象者がおらず、自分の解答に自信があるため、選択肢に解答がない場合は選択肢に問題があるとした。留学生は、選択肢に解答が見つからなかった場合、自分の理解力が不足している可能性を考えるのに対し、日本人大学生は選択肢の問題を一番に考えた。留学生は、自分の内側で「葛藤」が起きるのだが、日本人大学生はテスト・アイテムに対して「不信」を抱く。留学生が解答を得るためにストラテジーを使用するのに対し、日本人大学生は既に解答を持っているため、自分の解答と選択肢の整合性を得るためにストラテジーを使用する。留学生の中には、自分の日本語能力の不足で解答放棄をした人と、テスト・アイテムの問題として解答放棄をした2タイプが見られたが、日本人大学生は、後者の理由のみであった。

正答を選んだ理由に、「それ以外は違うので」という例もあり、選択肢の正確さ、妥当性に疑問が残る。課題の捉え方も、例えば、設問で「ディスカッションの利点」を問われた際、話し手の利点、聞き手の利点のように立場を変えた場合は違ってしまいうこともあり、プリタスクの曖昧さも挙げられた。「ディスカッション」という語彙について、高校までは「話し合い、議論」などを使用していたため、テスト内で使用されていた「ディスカッション」の使い方に少し違和感があったという例があった。その対象者は、大学に入ってから「ディスカッション」という語彙を聞くようになり、正確な意味についてはまだ

曖昧であると話していた。カタカナ語に関しては、大学の専門分野で使用頻度、使用カタカナ語が違ふと思われる。また、個々により語意観の違いがあることは否めない。テスト使用、特にキーワードに使用する際は、慎重さを要する。

6.6 設問があるアイテムにおける2者の総括と考察

留学生の聴解行動で特徴的だったのは、不安が及ぼす影響である。設問で抱えた不安は、本文聴解に負の影響を与え、本文で不安を抱えると選択肢の聴解に負の影響を与える。本文で解答を得たと考えた留学生も、選択肢で迷いが生じ、不安になる。その不安が、次のテスト・アイテムの聴解にまで負の影響を及ぼす。不安を抱えたままの聴解は、負のスパイラルに陥る危険性があることがわかった²⁰。

日本人大学生にとっての不安は、「記憶に残らない」だけであり、これはメモをとれば解決できた。そして、選択肢に対しては、解答がない場合は自分の理解不足ではなく、テスト・アイテムの問題であると、自身の内部に不安を抱えない。心理的な負担をどのように、どこへ処理するのかという点で、留学生は内側に抱え、日本人大学生は外部に転嫁した。聴解テストは、受験者がコントロールすることができず、次々と音が流れてくる。そのような状況下では、できる限り集中し、平常心を保ちたい。

Bachman & Palmer (1996 邦訳 pp.78-86)は、言語知識を、言語使用における談話の解釈をする際のメタ認知的ストラテジーにおいて利用できる情報領域と定義している。言語知識は、文法的知識とテキスト知識からなる構造的知識と、機能的知識と社会言語学的知識からなる語用論的知識であり、メタ認知的ストラテジーの領域は目標設定、アセスメント、計画としている。留学生と日本人大学生の言語知識を比較すると、文法的知識においては、日本人大学生が高く、留学生は日本人大学生と同等程度の人から低い人までと幅がある。テキスト知識は談話構成の知識であり、留学生と日本人大学生は年齢や経験から母語においてはほぼ同等であると考えられる。機能的知識は、日本語の使用下においては日本人大学生は高く、留学生は高い方から低い方まで幅があるが、基本的な機能的知識²¹は留学生と日本人大学生に差はない。社会言語学的知識も日本語使用下では、日本人大学生が高く、留学生は日本人大学生と同等程度の人から低い人まで幅がある。対象者が属する大学内の知識は、日本人大学生と留学生とほぼ同程度ある。留学生に不足しているのは、日本語の文法的知識と日本語の言語的特徴を要する語用論知識である²²。留学生は不十分な知識であっても、他の知識で補う、あるいは構成要素からの推測などストラテジーを多用していた。また、留学生は方略的能力（目標設定・アセスメント・計画）の使用が、日本人大学生に比べ頻度が高く、戦略的に使用していた。

嶋田 (2006, pp.56-57) は、大学で期待される基礎日本語能力を以下の5つに分類している。

1. 基本的思考能力
2. 能力操作能力（課題分析能力、能力行動能力、モニター能力）
3. 学業日本語運用能力（語彙的知識運用能力、文脈構成能力）
4. 自立能力（補い能力、学習ストラテジー能力、社会適応能力）
5. 学園情報獲得能力（履修情報能力、学業支援サービス情報獲得能力、学生生活情報獲得能力）

²⁰ アイテムの特性による困難度の違いがあるため一概に正答率の評価はできないが、本研究で行なったテストの正答率は留学生：78%、日本人大学生：90%であった。

²¹ Bachman & Palmer (1996 邦訳 pp.80-81)は、機能的知識を、概念的機能知識（経験から意味を表明し、解釈する）、操作的機能知識（相手に影響を与えるために言語使用をする）、学習的機能知識（知識を広げるために言語使用をする）、想像的機能知識（ユーモアや美的目的で世界を広げるために言語使用をする）の4つに分けて説明している。

²² 今回の留学生対象者は、日本語の中級レベル修了者から超級レベルの人がおり、一概には言えないが、不足している人も一部いたという意味である。

嶋田の分類に従うと、本研究で見られた行動は、プリタスクを理解する「課題分析能力」、テスト・アイテムをどのぐらい理解できているか、問題点は何かを探る「モニター能力」、構造的知識である「語彙的知識運用能力」、構造的知識の足りない部分を補う「文脈構成能力、補い能力、学習ストラテジー能力」である。留学生も日本人大学生も聴解テストの経験は豊富であるが、テスト・テクニックの使用は、留学生に限って見られ、言語知識の不足を補う形で使われている。日本人大学生は、日本語の言語知識の不足がなかったため、逆にテスト・テクニックを避け、純粹に聞いた情報のみの解析で解答していた。課題分析能力およびモニター能力においては、留学生と日本人大学生に差はないと考えられる。

設問の有無による聴解行動の提案と今後の課題は、7章で記述する。

7章 テスト時の聴解行動（設問なし）

7章では、6章と同様の研究枠組み、対象者、データの収集と分析の手順を使用し、日留試験聴解テスト・アイテムから設問を外した3アイテムを分析・考察する。設問を外した理由は、設問という課題がない場合、聞き手は自ら課題発見をしなければならない。これは、堀井（2006,p.116）のいう問題発見解決能力で、アカデミック・ジャパニーズに必要とされている。課題がない場合のアカデミック・リスニングを、留学生と日本人大学生で分析・考察することで、6章での分析・考察をさらに深める。設問のプリタスクだけでなく、状況説明も外した経緯は、足立（2009a）の研究から、状況説明の情報はあってもよいが、本文から状況がわかるため必要性はさほど高くないという結果が得られたため、設問の影響がなにもないほうがよいと判断した。実際に使用した3アイテムを以下に載せる。

・使用した3アイテム（設問なし）

《 》内の文章は設問である。対象者は設問を聞いていない。

4番 （2006年第1回13番 ライフ場面）

《女性と男性が、道路のアスファルトについて話しています。この道路の温度が高くないのは、なぜですか。》

女性：見て、この表示。道路を涼しいアスファルトに変えました、だって。

男性：ああ、それは新しいタイプのアスファルトのことだよ。水分を吸収して長い間保てる素材が混ぜてあるんだ。

女性：それが、どうして、涼しいアスファルトになるの。

男性：例えば、雨が降るとするね。すると、その素材が水分を吸収する。そして、その水分が蒸発する時にアスファルトの熱を奪う。それで、道路の表面温度を下げられるってわけさ。

女性：じゃあ、土と同じようなものね。長い間って、どのくらい？

男性：一週間たっても6度から8度くらいは温度を下げた状態にしておけるらしいよ。

女性：へえー、すごいねえ。

この道路の温度が高くないのは、なぜですか。

1. アスファルトが水で直接冷やされるから
2. アスファルトに特別な水が混ぜてあるから
3. アスファルトの中で水が循環しているから
4. アスファルトに水分を吸収する素材が入れてあるから

5番 （2006年第1回15番 アカデミック授業場面）

《先生が、授業で行ったディベートの感想文について話をしています。次回は、ディベートのやり方をどのように変えますか。》

一回目のディベートについて、皆さんの感想文を拝見しました。ディベートでは、決められた論題について、自分の意見とは無関係に賛成か反対かどちらかの側に立ってその立場の正しさを証明しなければなりません。多くの人がこれには抵抗感があつたようですね。でも、これはディベートの根本ルールですから、変えようがありません。次に多かつたのは、自分が意見を述べる時、対戦相手ではなく、審判員

たちに顔を向けて話さなければならないのは、やりにくいという声でした。ディベートは相手を言い負かすゲームではなく、聞いている審判員が勝敗を決めるからなのだとことを、皆さんよくわかってはいるようですが、やはりこの点はまだ難しいようです。そこで、次回は、相手方を見て話すことにしましょう。ただし、ディベートの目的は、審判員を説得することだということは忘れないようにしてください。

次回は、ディベートのやり方をどのように変えますか。

1. 自分の意見と関係のあることを述べる。
2. 賛成か反対かどちらかの側に立って話す。
3. 対戦相手の方を見て話す。
4. 審判員を説得する。

6番 (2005年第2回20番 アカデミック講義場面)

《先生が「都市政策」すなわち、都市に対する政策について話しています。この先生は、今後の都市政策にはどうすることが必要だと言っていますか。》

従来日本で行なわれてきた都市政策には、経済効果ばかりを重視し、住民のことをあまり考えていない、という問題点がありました。例えば自動車道路の建設では、住民の生活の快適さよりも、交通の効率化が優先されていました。これからの都市政策は、人間が中心です。住民が、自分のまちに誇りと愛着を持って暮らすことができるようなまちづくりが求められています。そうしたまちづくりの主役となるのは、実は行政ではなく、ほかならぬ住民である、と私は思います。そのまちに住む住民自身が、都市政策を行政任せにせず、政策を立てる段階から積極的に関わっていく。そういうことが、今後重要になってくると思っています。

この先生は、今後の都市政策にはどうすることが必要だと言っていますか。

1. 交通手段を整備し、他の都市との関係を密にすること
2. 住民自身も、都市政策を立てることに協力していくこと
3. 住民が都市政策を立てることに對して、行政が指導すること
4. 行政が都市政策から手を引き、住民に全てをまかせること

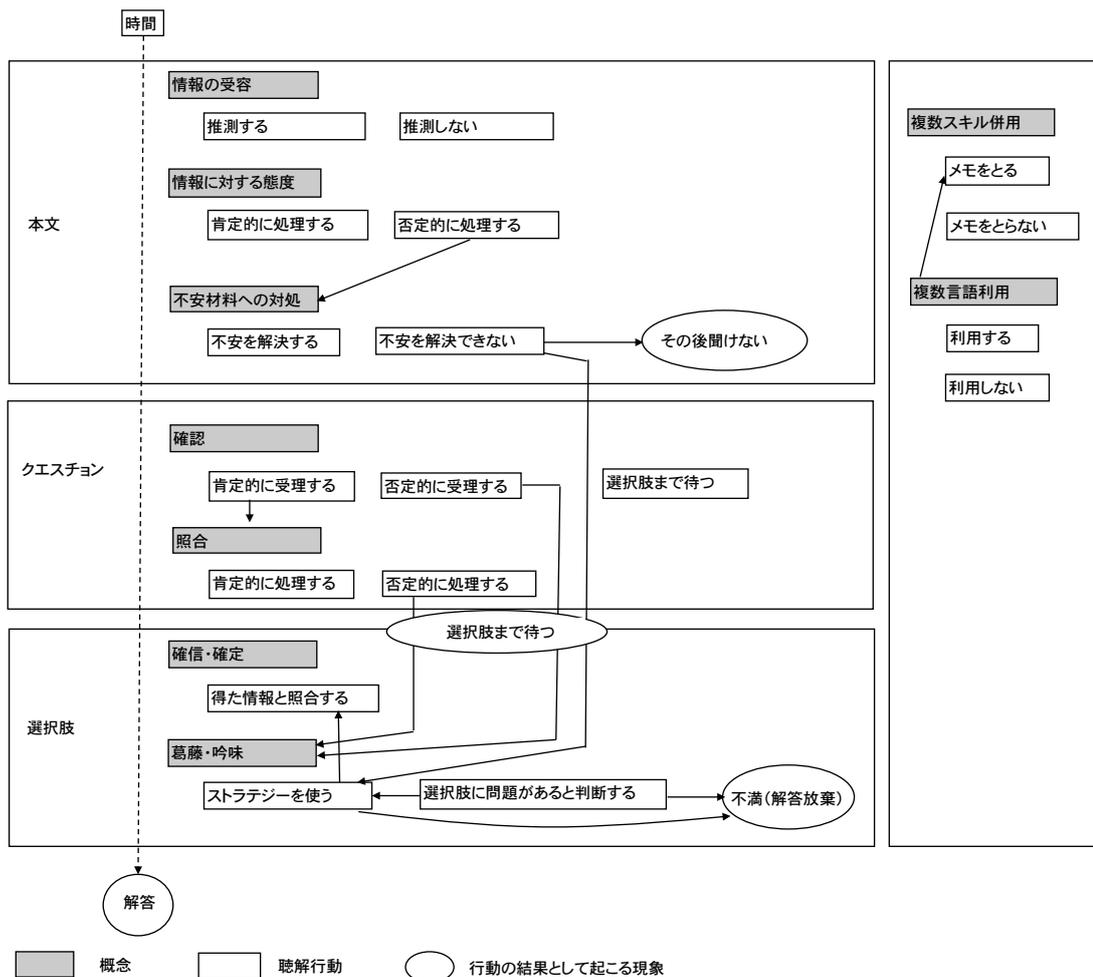


図 7.1 留学生の聴解行動（設問なしのアイテムに対する行動）

7.1 留学生の聴解行動の分析結果と考察（設問なし）

図 7.1 は留学生の聴解行動である。聴解テストの時間的な流れは上から始まり、最後の「解答」で終わる。また、右横の「複数スキル併用」と「複数言語利用」という概念は、テスト・アイテムが始まってから終わるまでの全体に関わる行動であるため、これを「全体」として分析・考察をする。□ はテスト・アイテムを聞きながらされた聴解行動であり、それらのグループを概念として表したものが ■ である。○ はある聴解行動をした結果、起こった現象である。これは、聞き手が自ら決めた聴解行動ではないため、現象として区別した。「本文」と「ポスト・クエスチョン」の間にある概念「解答の言語化」は本文が終わった直後であり、ポスト・クエスチョンが始まる直前である。「本文」と「選択肢」には複数の行動が見られるが、時間的な違いはない。「→」はある行動を取ると決定した場合、次の行動に影響を与えると判断したものである。

以下、アイテムの各部分の聴解行動を詳細に分析・考察する。個々の対象者がコメントした発話内容は、ヴァリエーションの 1 例であるため、「例」と記述する。

設問がないアイテムで、留学生に特徴的に見られた現象として、本文時の「その後聞けない」（設問があるアイテムと同様）と、選択肢のときの「わからない」である。これらの現象については各項目の分析と考察で詳述する。

7.1.1 全体

表 7.1 アイテム全体の概念と聴解行動のヴァリエーション①（留学生設問なし）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
全体	複数スキル併用	聞きながらメモをとる	メモをとる	①推測の手がかりにする（語彙） ②記憶のためにする ③集中する ④全部・できるだけ書いておく	
			メモをとらない	①本文が聞けなくなる ②集中できない	
	複数言語利用	理解する際に母語、英語などを使用する	複数言語を利用する	①書くスピードが速い ②カタカナ語理解のために英語での意味を考える ③訳している ④安心する	
			日本語のみを利用する		

表 7.1 は、留学生のアイテム全体にわたって見られた行動であり、6章で分析した設問がある場合と7章で分析した設問がない場合を比較した表である。

アイテム全体にわたって見られたのは、「複数スキル併用」と「複数言語利用」であった。聴解行動も設問があるアイテムと同様である。よって、聴解行動のヴァリエーションの違いを設問がある場合と設問がない場合で比較してみる（表 7.1.1）。

表 7.1.1 アイテム全体の聴解行動のヴァリエーション①（留学生設問ありと設問なしの比較）

聴解行動	ヴァリエーション（留学生設問あり）	ヴァリエーション（留学生設問なし）
メモをとる	①推測の手がかりにする（語彙） ②記憶のためにする ③集中する ④全部・できるだけ書いておく	①推測の手がかりにする（語彙） ②記憶のためにする ③集中する ④全部・できるだけ書いておく
メモをとらない	①本文が聞けなくなる ②集中できない	①本文が聞けなくなる ②集中できない
複数言語を利用する	①書くスピードが速い ②カタカナ語理解のために英語での意味を考える ③訳している ④安心する	①書くスピードが速い ②カタカナ語理解のために英語での意味を考える ③訳している ④安心する
日本語のみを利用する	なし	なし

併用スキルは、聞きながらメモをとるか否かである。メモをとる場合、設問がある場合は、プリタスクやキーワード、解答と思われる部分などのメモが主だったが、設問がない場合は、【設問なし】「②記憶のためにする」であり、全部あるいはできるだけ書いておくという対象者もいた。これは、何が解答になるか絞れないためである。メモの内容の違いはあるが、テスト全体にわたっての聴解行動は、設問の有無に関わらずほぼ同様であった。

7.1.2 本文

本文では、設問がある場合と同様、「情報の受容」「情報に対する態度」「不安材料」（表7.2）の3つが見られた。

表 7.2 本文時の概念と聴解行動のヴァリエーション②（留学生設問なし）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
本文	情報の受容	解答のための情報を得る	推測する (質問される可能性のある部分に注目する)	言語形式・談話構造 ①談話構造から ②談話のコミュニケーション機能から テキスト外 ①メイン・トピックを考える ②主張は何か考える キーワード ①トピック・キーワードを考える	
			推測しない (全体を聞く)	①全体理解が大切 ②内容把握に支障が出る ③余裕がない ④わからない語彙があり考えられない	
	情報に対する態度	難易に関わる要因から情報を処理する	肯定的に処理する	言語形式・談話構造 ①統語的機能がわかる ②ディスコース・マーカーをヒントにする ③構成が論理的 ④構成が対比 テキスト外 ①背景知識・常識・経験を使う ②トピックが得意(既知語彙) ③トピックがわかると談話構成が予測できる ④テスト・テクニックを使う ⑤先読みする キーワード ①頻度が高い ②全体を聞くとわかる	
			否定的に処理する	言語形式・談話構造 ①スピードが速い ②発音が不明瞭だ(活用語の音聞き間違える・特殊音) ③声・話し方に抑揚がない ④語彙が難しい(わからない・漢字熟語が多い) ⑤トピックが日常的ではない(専門用語) ⑥トピック展開が中途半端 ⑦談話構造が日常談話と違う ⑧情報量が多い テキスト外 ①背景知識・常識・経験がない ②トピックに興味がない・反発する ③設問がない形式に慣れていない キーワード ①キーワードの語彙がわからない	
	不安材料2	わからない語彙や表現、複数解答候補がある	不安を解決する	①推測(文脈から・母語から・英語から・発音から漢字・語彙分析から)する ②勘	
			不安が解決できない	①集中力がきれる ②思考が止まる	その後聞けない・選択肢まで待つ

設問がある場合と比較すると「情報の受容」「情報に対する態度」「不安材料2」の3つの概念があることは同じであった。この3つの概念を設問がある場合と比較して分析を行う(表7.2.1表7.2.2表7.2.3)。

表 7.2.1 本文時の聴解行動のヴァリエーション②（留学生設問ありと設問なしの比較）

「情報の受容」

聴解行動	ヴァリエーション（留学生設問あり）	聴解行動	ヴァリエーション（留学生設問なし）
選択的に受容する	①解答のための情報と認知する ②不必要な情報は捨てる ③解答を得た後はあまり集中しない	推測する（質問される可能性のある部分に注目する）	言語形式・談話構造 ①談話構造から ②談話のコミュニケーション機能から テキスト外 ①メイン・トピックを考える ②主張は何か考える キーワード ①トピック・キーワードを考える
全体的に受容する	①内容把握する ②全体内容を要約する ③最後が変わることがある ④解答は最後にくることが多い ⑤常識と違うことを言う可能性がある	推測しない（全体を聞く）	①全体理解が大切 ②内容把握に支障が出る ③余裕がない ④わからない語彙があり考えられない

まず、「情報の受容」から分析・考察する。概念は設問がある場合と同様だが、聴解行動に違いが見られたため、聴解行動とヴァリエーションの比較をする（表 7.2.1）。設問がある場合は、選択的か全体的かを決めて聞くが、設問がない場合は、本文を聞きながら課題を推測し絞り込んで聞くか、推測はせず全体把握をするかを決定する。「課題を推測する」と、「心理的にアイテムの困難度が下がる」と述べた対象者がいた。何を聞くべきかがわかっているほうが、記憶にかかる負荷が減らせ、情報を得たあとは聞き流せばよい。テスト・アイテムの困難度が下がるか否かは判断できないが、心理的な負担は減るためそのような行動をとった可能性が高い。【設問なし】言語形式・談話構造の「①談話構造から」は、例えば、「以前は～。今後は～。」という本文内容であれば、時系列のどちらかが課題になると推測する。「②談話のコミュニケーション機能から」は、例えば、本文中のテキスト内に説明部分と意見部分があった場合は意見部分が課題となる、テキスト内に質問と応えがある場合は質問の応えが課題になる、テキスト内に相談とアドバイスがある場合はアドバイスが課題となる、テキスト内容に評価がある場合は良い点あるいは悪い点が課題となるだろうと推測する。【設問なし】「テキスト外の「①メイン・トピックを考える ②主張は何か考える」は、本文全体から最も重要だと考えられる内容を推測し、【設問なし】「キーワードの「①トピック・キーワードを考える」は、トピックがわかった時点で考える。このように課題を推測して情報を得た場合、本文のあとに聞くクエスチョンと課題が違っていると、解答が得られないリスクがある。そのリスクを回避した聞き方が、「課題を推測しない」である。全体を聞いていれば、どのようなクエスチョンにも答えられると考え、【設問なし】「①全体理解が大切」だと、意図的に全体を聞くようにしている。①と同様の理由で【設問なし】「②内容把握に支障」をきたさないように、課題を推測しない。【設問なし】「③余裕がない」は、意図的に全体を聞くのではなく、結果的に全体を聞くしかなく、【設問なし】「④わからない語彙があり考えられない」は、課題を推測する情報が不十分であったからである。

設問がある場合、どのようにあるいは何を聞くべきかという本文を聞く準備を行うが、設問がない場合は、本文を聞きながら、どのようにあるいは何を聞くべきかという判断を、同時にしている。これは、堀井（2005, p.16）がアカデミック・ジャパニーズとして挙げている問題発見解決能力と考えられる。講義を聞きながら、重要となりうる部分を見つけ、何をノートに書き留めるべきか何を記憶すべきかを判断するのと同じである。しかし、設問がない聴解テストでは、受験者が重要だと考えた部分が必ず問わ

れるかはわからない。例えば、本研究で使用したアイテムに「ディベート」について教師の談話がある（p.153 5番を参照）。この談話の要点は、「1.ディベートの説明、2.学生の感想の1位と2位、3.それに対する教師の意見、4.次回への指示」である。このアイテムの設問は、「先生が授業で行ったディベートの感想文について話をしています。今回は、ディベートのやり方をどのように変えますか。」であった。設問がある場合は、談話を聞く前に「4.次回の指示」のみを聞けばよいとわかり、「1.ディベートの説明、2.学生の感想1位と2位、3.それに対する教師の意見」の部分が、たとえ理解できなくても解答が可能である。実生活の聴解は、自らの課題がある場合は（聞く目的と聞くポイントが明確）、その課題に沿った行動ができれば行動上は問題解決となりうる。上に挙げたテスト・アイテムの内容を実生活で聞くと置き換えた場合、テストで与えられた課題部分の理解だけでは、不十分である。談話全体の理解があつて、次回への指示が聞き取れることが必要であるため、このアイテムは設問がないほうが聴解能力測定としては有効であろう。

表 7.2.2 本文時の聴解行動のヴァリエーション②（留学生設問ありと設問なしの比較）
「情報に対する態度」

聴解行動	ヴァリエーション（留学生設問あり）	ヴァリエーション（留学生設問なし）
肯定的に 処理する	<p>言語形式・談話構造</p> <p>①音声の特徴がヒント ②スピードがよい ③男女の会話なのでどちらが言うかわかる ④語彙が簡単 ⑤統語的機能がわかる ⑥ディスコース・マーカ―に注意する ⑦構成が明確だ ⑧論理的な構造だ ⑨解答が明確に提示される</p> <p>テキスト外</p> <p>①常識から判断 ②常識と違う ③背景知識から判断 ④話すときの習慣から判断する ⑤話し手の性格を推測する ⑥惑わし選択肢、罠を推測する ⑦テスト・テクニックを使う</p> <p>キーワード</p> <p>①トピックから推測する ②プリタスクから推測する ③頻度が高い ④内容理解から推測する</p>	<p>言語形式・談話構造</p> <p>①統語的機能がわかる ②ディスコース・マーカ―をヒントにする ③構成が論理的 ④構成が対比</p> <p>テキスト外</p> <p>①背景知識・常識・経験を使う ②トピックが得意（既知語彙） ③トピックがわかると談話構成が予測できる ④テスト・テクニックを使う ⑤先読みする</p> <p>キーワード</p> <p>①頻度が高い ②全体を聞くとわかる</p>
否定的に 処理する	<p>言語形式・談話構造</p> <p>①スピードが速い ②発音が不明瞭だ ③母語と違うため（非漢字圏は統語的相違）・語彙がわからない・カタカナ語（漢字圏）が難しい ④統語機能がわからない ⑤談話構造に対するピリーフと違う ⑥トピック展開が速い ⑦解答が不明確だ ⑧量が多い ⑨情報量が少なすぎる</p> <p>テキスト外</p> <p>①背景知識がない ②常識と違う ③間違った常識を使う ④トピックに興味がない、苦手だ、専門外だ、遭遇頻度が少ない ⑤メモに集中しすぎる ⑥解答を確信した後聞かない ⑦集中力が低下する ⑧緊張 ⑨課題が予想外</p> <p>キーワード</p> <p>①キーワードの勘違い（音から判断するため）</p>	<p>言語形式・談話構造</p> <p>①スピードが速い ②発音が不明瞭だ（活用語の音を聞き間違える・特殊音） ③声・話し方に抑揚がない ④語彙が難しい（わからない・漢字熟語が多い） ⑤トピックが日常的ではない（専門用語） ⑥トピック展開が中途半端 ⑦談話構造が日常談話と違う ⑧情報量が多い</p> <p>テキスト外</p> <p>①背景知識・常識・経験がない ②トピックに興味がない・反発する ③設問がない形式に慣れていない</p> <p>キーワード</p> <p>①キーワードの語彙がわからない</p>

次に、「情報に対する態度」を分析・考察する（表 7.2.2 前頁）。聴解行動は「肯定的に処理する」と「否定的に処理する」の2つで設問がある場合と同様である。よって、聴解行動のヴァリエーションの違いを比較する（表 7.2.2）。設問があるアイテム時と比較すると、ヴァリエーションにばらつきはあるが、大きな違いはない。よって、設問がないための聴解行動と考えられるものだけに注目する。「肯定的に処理する」の、【設問なし】「⑤先読みする」は、「聞いてからメモするという後追い行動では間に合わないの、談話の先読みをしながら聞く」という例である。設問がある場合は、常にプリタスクを頭において聞けるため、解答と関係がある情報を聞いたときにメモをすればよい。しかし、設問がない場合は、何が解答となるのかわからないため、多くのことを記憶あるいは記録しなければならない。そのため、談話の先を読んで、重要となりうる情報を予測しておく、その情報が聞こえたときの処理スピードが早くできるからである。「否定的に処理する」に【設問なし】「③設問がない形式に慣れていない」という例があり、テストの聴解のみならず授業の聴解練習でも設問があるアイテムを大量に聞いていたことがうかがえる。インタビュー時、設問がないアイテムに対し、「普段の聞き方に近い」「普段も話し手のポイントは何かを考えながら聞くのでテストで設問があったらそれをポイントにし、設問がなかったら自分でポイントを考える」という例がある一方、「設問がないアイテムではタスクを考えることも同時にしたため本文理解がおろそかになった」「日常の聴解では設問など考えない」という例もあった。個々違う聞き方をしているという例が得られたことは、聴解研究には貴重である。「普段の聞き方」は、個人差もあるが、聞く目的によって課題の有無が生じる。聴解テストだから課題があるのではなく、日常聴解でも聞く目的があれば、それが課題になる。友人とおしゃべりの話題に興味がなければ、時々うなづく程度でも人間関係は保たれるが、試験の範囲を聞き逃してはいけない。日常聴解で、無意識に行っている課題設定を、意識的に行うことで、外国語聴解により変化をもたらすと考える。課題設定をする・しない聞き方、双方の効果を体験することで、自分の聴解のスタイルを認知できる。どちらがより効果的かは、個々で差があるかもしれないが、自分の聞き方を認知することは、聴解学習により効果をもたらすだろう。「⑥トピック展開が中途半端」に、「推測したタスクから考えて実験結果の話が続くと期待していたが、提案のみで終わってしまい簡単すぎる」という例があった。本文理解もしており、正答を選んでいるのだが、「テスト・アイテムに対して不満がある」と話した対象者もいた。日本語の大規模テストの形式は、3章の「表 3.2 聴解テスト・アイテムの設問、選択肢」（p.48 を参照）にみられるように新日能試の一部を除き、すべて設問があるタイプである。テストには実用性が求められており（Bachman & Palmer 1996 邦訳 p.41、静 2002, p.44、尾崎 2008, p.38）、日留試は日本語、理科、総合科目、数学を同日にテストするため、日本語のテスト時間は2時間程度である。2時間でより多くのテスト情報を得なければならないため、1～2分の短いテスト・アイテムになっていると予想される。本文談話が短ければ、一問一答になってしまうだろう。英語の TOEFL コンピュータ版（CBT、iBT）では、10分程度の講義や会話などが使用されており、特に iBT は CBT にあった短い会話の聴解がなくなった。聴解アイテムの時間は長く一問複答である。日留試も場面や状況、トピック、聞く目的や内容に適したアイテムの時間の長さ、複答形式を柔軟に取り入れる必要があるのではないかと思われる。

最後に、「不安材料」についてである（表 7.2.3 次頁）。設問がある場合は「わからない語彙や表現がある」と定義し「不安材料1」とした。設問がない場合は「わからない語彙や表現、複数の解答候補がある」と「複数の解答候補がある」が加わったため「不安材料2」とした。聴解行動は「不安を解決する」と「不安を解決できない」であり、設問がある場合と同様である。聴解行動のヴァリエーションの違いを分析・考察する（表 7.2.3）。設問がない場合はプリタスクがわからず、本文を聞きながら様々な情報をもとに課題を推測する（【設問なし】①～⑤）のだが、設問がある場合と推測元はほぼ同じである。設問がある場

合と大きく違ったのは、推測する際に複数の課題を推測する対象者がいたことである。語彙や表現などがわからない不安と、課題がなく解答が絞れない不安が重なり、緊張感は増したようである。この緊張を集中力に変えた対象者もいるため、緊張が必ずしも負の影響を与えないとは言えないが、長時間の緊張はストレスになるだろう。また、設問がある場合は、「解答が得られた時点で談話はすぐに終わるだろうと予測できる」のだが、設問がない場合は「いつ終わるかがわからない」という意見もあった。受験者あるいは聞き手にとって、話の終わりが予測できない談話は、構造的に不完全である。プリタスクがあるから答えられるという談話作成ではなく、談話として完結するようなテスト・アイテムにすべきである。

不安材料が解決できない場合は、パニックになり「その後聞けなくなってしまう」対象者もいた。また、本文がわからないため、消極的ではあるが「選択肢まで待つ」という聴解行動も見られた。

表 7.2.3 本文時の聴解行動のヴァリエーション②（留学生設問ありと設問なしの比較）

「不安材料」

聴解行動	ヴァリエーション（留学生設問あり）	ヴァリエーション（留学生設問なし）
不安を解決する	①文脈から推測する ②母語から推測する ③カタカナ語は英語から推測する ④語彙分析から推測する ⑤トピックの背景知識を使う ⑥回避する	①文脈から推測する ②母語から推測する ③英語から推測する ④発音から漢字を推測する ⑤語彙分析から推測する ⑥勘
不安を解決できない	①集中力が分散する ②その後の本文を聞けない ③思考が止まる	①集中力がきれる ②思考が止まる

7.1.3 クエスチョン

設問がある場合は、プリタスクでクエスチョンを一度聞いており、本文を聞いた後ポスト・クエスチョンとして聞く。聴解行動としては、「再確認する」、プリタスク時に聞き逃したような場合は「確認する」という行動が見られた（6章「6.4.6 ポストクエスチョン」表 6.6 p.136）。

表 6.6 ポスト・クエスチョン時の概念と聴解行動のヴァリエーション⑥（留学生設問あり）

7/14	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
ポストQ	確認1	もう1度聞く	再確認する	なし	
			確認する	なし	曖昧である

設問がない場合は、本文終了後、初めて聞くクエスチョンであるため、設問がある場合のポスト・クエスチョンとは機能が大きく違う。よって、クエスチョンに関しては、設問がある場合と設問がない場合の比較ではなく、「クエスチョン」として独立させ分析を行う（表 7.3 次頁）。

クエスチョンでは、「確認 2」と「照合」の2つの概念がある。クエスチョン時で、本文終了後に初めて課題を聞くため、本文中で推測した課題と同じなのか異なるのかを確認し、その直後に解答となりうる情報を記憶から取り出し、あるいは整理する。よって、確認後に照合となる。

「確認 2」の「肯定的に受理する」は、課題を推測しながら聞いていた対象者が、推測した課題に対して「①推測とあっている」「②聞き逃した部分ではない」と確認し、推測せず全体を聞いていた対象者は、

談話の要点や主張などを聞いており、それが課題であったことに「③納得する」。クエスチョン時に「否定的に受理する」場合は、本文中で推測した課題と違っていたり、推測できなかったため、「照合」はできず、その時点では「わからない」と感じ、「選択肢まで待つ」ことになる。「③語彙がわからない ④意味がわからない」は、クエスチョンがわからなかったために、「選択肢まで待ち」、選択肢時に解答しようとしている。この2つの「選択肢まで待つ」は行動の結果として起こった現象としての「選択肢まで待つ」である。ヴァリエーションの3つ目の「選択肢まで待つ」はクエスチョンを聞いても、そこで解答をしようせず、最初から「選択肢まで待つ」と決めていた対象者であり、行動の結果の現象ではなく、積極的に選択した行動である。この行動は設問が最初にある場合より、解答を決めるまでは慎重になっている結果と言えよう。

設問がある場合は、本文を聞きながら解答を絞り込むことができ、また選択肢の前にポスト・クエスチョンが入るため時間的に余裕があるが、設問がない場合はクエスチョンを聞くと同時に解答を絞り込まなければならない。よって「照合」時に「①解答の言語化」を行うが、それにかかる時間は設問があるアイテムより短く、選択肢まで解答を言語化しない対象者もいる。そして、「③選択肢で消去法が使える」のように、選択肢を聞く際の態度を決めることもある。「照合」時の「否定的に処理する」は、クエスチョン内容は理解したが、「①本文内容把握が完全ではない ②理解した部分ではない ③聞き逃した部分だ」のように、この時点では解答が得られず、選択肢まで待つか、「わからない」と感じてしまう。

表 7.3 クエスチョン時の概念と聴解行動のヴァリエーション③（留学生設問なし）

7/14	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
クエスチ ョン	確認 ²	推測したタスクかどうか確認する	肯定的に受理する	①推測とあっている ②聞き逃した部分ではない ③納得する	
			否定的に受理する	①推測できなかった ②推測と違う ③語彙がわからない ④意味がわからない	選択肢まで待つ
			選択肢まで待つ ¹	なし	
	照合	本文から得た情報を解答になりうるか照合する	肯定的に処理する	①解答を言語化する ②選択肢で消去法が使える ③推測と違っていたが、得た情報から探す	
			否定的に処理する	①本文内容把握が完全ではない ②理解した部分ではない ③聞き逃した部分だ	選択肢まで待つ

7.1.4 選択肢

選択肢時に見られた概念は「確信・確定」と「葛藤・吟味」であった（表 7.4 次頁）。選択肢では、本文から得られた情報と選択肢が一致、あるいは言語化した解答と選択肢が一致していた場合「確信・確定」となる。これは設問があるテスト・アイテムの場合と同様である。解答を「確信・確定」できなかった場合は、選択肢から解答を見つけなければならない。解答を「確信・確定」できなかった場合は、

¹ 聴解行動としての「行動の結果として起こる現象」の「選択肢まで待つ」との違いは、積極的に「選択肢まで待つ」という行動を行っている場合は「聴解行動」であり、わからないので消極的に「選択肢まで待つ」場合は「行動の結果として起こる現象」とした。

選択肢から解答を見つけなければならない。この概念を「葛藤・吟味」とした。設問がある場合は「葛藤」のみであったが、「吟味」を加えた理由は後述する。

まず、「確信・確定」の聴解行動のヴァリエーションの違いを比較する（表 7.4.1）。【設問なし】「②解答がある」は、本文を聞きながら推測した課題と解答が選択肢にあると確信できた。設問がない場合、解答を得るためのプリタスクもなく、また本文直後に聞くクエスチョンから選択肢まで時間がないため、選択肢に正答があったときの喜びが伝わってくる。このような感情は能力測定項目とは無関係であるが、テストの動機付けとしてはよい影響を与えるのではないだろうか。

表 7.4 選択肢時の概念と聴解行動のヴァリエーション④（留学生設問なし）

7行ム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
選択肢	確信・確定	得た情報との一致度が高い	得た情報と照合する	①自分の解答と一致 ②解答がある	
	葛藤・吟味	解答を見つけなければならない 選択肢の内容と本文内容を吟味する	ストラテジーを使う 選択肢に問題があると判断する	①背景知識・常識から推測する ②消去法 ③本文の語彙との一致・類似・相違から選ぶ ④メモから推測する ⑤メモにある語彙を選ぶ ⑥言い換えの意味と自分の答えを比較する ⑦言い換える語彙を漢字になおす ⑧可能性のある選択肢を比較する ⑨本文の最後の部分のものを選ぶ ⑩メイン・ポイントや主張を選ぶ ⑪テスト・テクニックを使う ⑫4つの選択肢からわからない語彙を推測する ①期待した選択肢がない ②言い換えがされている ③選択肢同士が似ている ④適当な答えがない ⑤聞き取れた情報だけでは解答できない	不満（→解答放棄になりうる）

表 7.4.1 選択肢時の聴解行動のヴァリエーション④（留学生設問ありと設問なしの比較）
「確信・確定」

聴解行動	ヴァリエーション（留学生設問あり）	ヴァリエーション（留学生設問なし）
解答を確信する 解答を確定する	①自分が決めた答えから推理する ②錯乱肢だとわかる	①自分の解答と一致 ②解答がある

解答を「確信・確定」できなかった場合は、選択肢から解答を見つけなければならない。表 7.4.2 で（次頁）、設問がある場合と設問がない場合の聴解行動の比較をする。設問がない場合の概念を「葛藤・吟味（定義：わからない語彙や表現、複数解答候補がある）」とした。設問があるアイテムでは、「葛藤（定義：わからない語彙や表現がある）」だけであったが、「吟味」を加えた理由は、「ストラテジーを使う」の【設問なし】「⑧可能性のある選択肢を比較する ⑫4つの選択肢からわからない語彙を推測する」のように、設問があるアイテム時より選択肢の情報に頼っているヴァリエーションが見られたからである。また、「解答を決める時間が設問がないアイテムではほとんどないので選択肢をよく聞かなければならな

い」という例もあった。「本文で情報がうまくとれなかった」「クエスチョン時まででは必要な情報が足りなかった」などの理由で、選択肢までに解答を得られなかった対象者にとっては、選択肢が唯一残された情報源である。選択肢から本文内容を推測し、選択肢の内容だけで【設問なし】「⑩メイン・ポイントや主張を選ぶ」というヴァリエーションも見られた。これは、クエスチョンがメイン・ポイントや主張を問うていない場合にも見られた行動である。プリタスクがなく本文を聞く場合は、その談話のメイン・ポイントや主張を考える対象者が多く、当然その部分が質問されるだろうと予想していたが、クエスチョンが予想と違っていった場合、他の情報は重要ではないと落としており、足りない情報から自分の解答を変えられず、間違いだとはわかっていても、【設問なし】「⑩メイン・ポイントや主張を選ぶ」ことしかできなかった。「クエスチョン」に対して、「予想と違い些細なことを質問された」という例があった。自分が推測した、あるいは大切だと思った部分とクエスチョンが違っていった場合、すぐに修正できた対象者もいればできなかった対象者もいる。対象者が感じた「些細なこと」を、テストの課題にすること自体は問題ではない。それは、プリタスクで指示されたことを本文から抽出できる能力を、測定しようとしているからである。そのような能力の測定をする場合は、設問が必要である。

表 7.4.2 選択肢時の聴解行動のヴァリエーション④（留学生設問ありと設問なしの比較）
「葛藤」と「葛藤・吟味」

聴解行動	ヴァリエーション（留学生設問あり）	ヴァリエーション（留学生設問なし）
ストラテジーを使う	①背景知識や常識から推測する ②常識だから誤答と考える ③消去法 ④唯一選択をする（唯一わからない選択肢、1つしかわからない選択肢、意味のない選択肢） ⑤本文の語彙と一致、類似、相違から判断する ⑥わからなかった語彙を音から漢字に置き換える（非漢字圏） ⑦統語機能から判断する ⑧本文内容から推理する ⑨選択肢から本文内容を推測する ⑩わからない語彙を回避する ⑪テスト・テクニックを使う ⑫勘、当て推量、運	①背景知識・常識から推測する ②消去法 ③本文の語彙との一致・類似・相違から選ぶ ④メモから推測する ⑤メモにある語彙を選ぶ ⑥言い換えの意味と自分の答えを比較する ⑦言い換えの語彙を漢字になおす ⑧可能性のある選択肢を比較する ⑨本文の最後の部分のものを選ぶ ⑩メイン・ポイントや主張を選ぶ ⑪テスト・テクニックを使う ⑫4つの選択肢からわからない語彙を推測する
選択肢に問題があると判断する	①選択肢同士が似ている ②曖昧だ ③言い換えがされている ④適当な答えがない ⑤正答が複数ある	①期待した選択肢がない ②言い換えがされている ③選択肢同士が似ている ④適当な答えがない ⑤聞き取れた情報だけでは解答できない

選択肢から解答が得られなかった場合、「選択肢に問題があると判断する」という行動は、設問がある場合と同様であった。設問がないアイテムの場合、解答放棄をした対象者はいなかったが、正答を聞いても納得ができず「本文や選択肢が間違っている、曖昧だ、誤答がより正解に近い」と不満を表した対象者がいた。この結果から、設問がある場合、本文を聞き終わった段階で解答がないと判断したり、選択肢に適当なものがない場合、課題に対する解答をかなり限定しているため、解答放棄をするのではないかと考える。設問がない場合は、自分の考えた正答に対して、設問がある場合より許容範囲が広く、解答放棄をせずに誤答を選んだと思われる。

ここから、留学生が設問の有無に関してどのように感じていたのかを分析する。「どちらも差がない、設問がないほうが煩わしくない、設問がないほうが普通の聞き方に近い」という意見もあったが、多くは「設問がないとメモがとりにくくすべてを記憶しなければならない、すべてを理解しなければ答えら

れない」と、記憶の負担や全文理解により困難度があがると感じる対象者がいた。留学生は設問がないアイテムの正答率が下がっている²事実からも、設問がないアイテムのほうが難しかったと言える。「設問があれば、そこからキーワードや本文内容の推測などを早めに準備して本文を聞くことができるため、言語知識を補うことができる」「心理的にも聞くべきポイントがわかっているので安心感がある」と述べている対象者もいた。このように、留学生は設問がないテスト・アイテムをより難しいと感じているのだが、聴解テスト・アイテムとしては以下のように高く評価をしている。設問がない場合は、「すべてを聞く必要がある、すべてを理解しなければ答えられない、何が重要かを選び出して答えを得ることが必要だ、設問があると聞き方のコツだけでわかってしまう」という意見があった。同様に、留学生は設問があるアイテムに対しても評価している。例えば、「普段の聴解では設問がある聞き方とない聞き方と両方あるため両方のタイプが必要だ、目的によって違うので両方あるべきだ、1つの形式だけだと違う形式の聴解になった場合できなくなるので両方必要だ」という意見があった。留学生は、日本語授業の聴解練習で日留試のような設問がある形式の練習と、設問がないタイプ、長い談話、あるいは生教材などでも練習をしているはずである。また、日本で生活し、日本語で授業や講義も受けている³。留学生は、多様な聴解経験から、それぞれの聴解方法の違いを自ずと獲得しているのではないかと考えられる。留学生の考える聴解能力は、①全文を理解すること ②何を聞かなければならないかを見つけること ③聞く目的によって聞き方を変えること、であることがわかった。日留試で、この3つの能力を測定することが求められるのではないだろうか。部分理解でよい場合は設問が必要であり、全文理解が必要かつ何が重要な情報かを判断する能力も含めて測定する場合は設問がないアイテムが適切だと思われる。

² 設問がないアイテムの正答率を記す。留学生は設問があるアイテムの正答率78%だったが、設問のないアイテムの正答率は72%と下がった。

³ 本研究の対象者の留学生は、外国語としての日本語授業以外に、自分の専門や母国の歴史などの日本語での授業を受けている。

7.2 日本人大学生の聴解行動の分析結果と考察（設問なし）

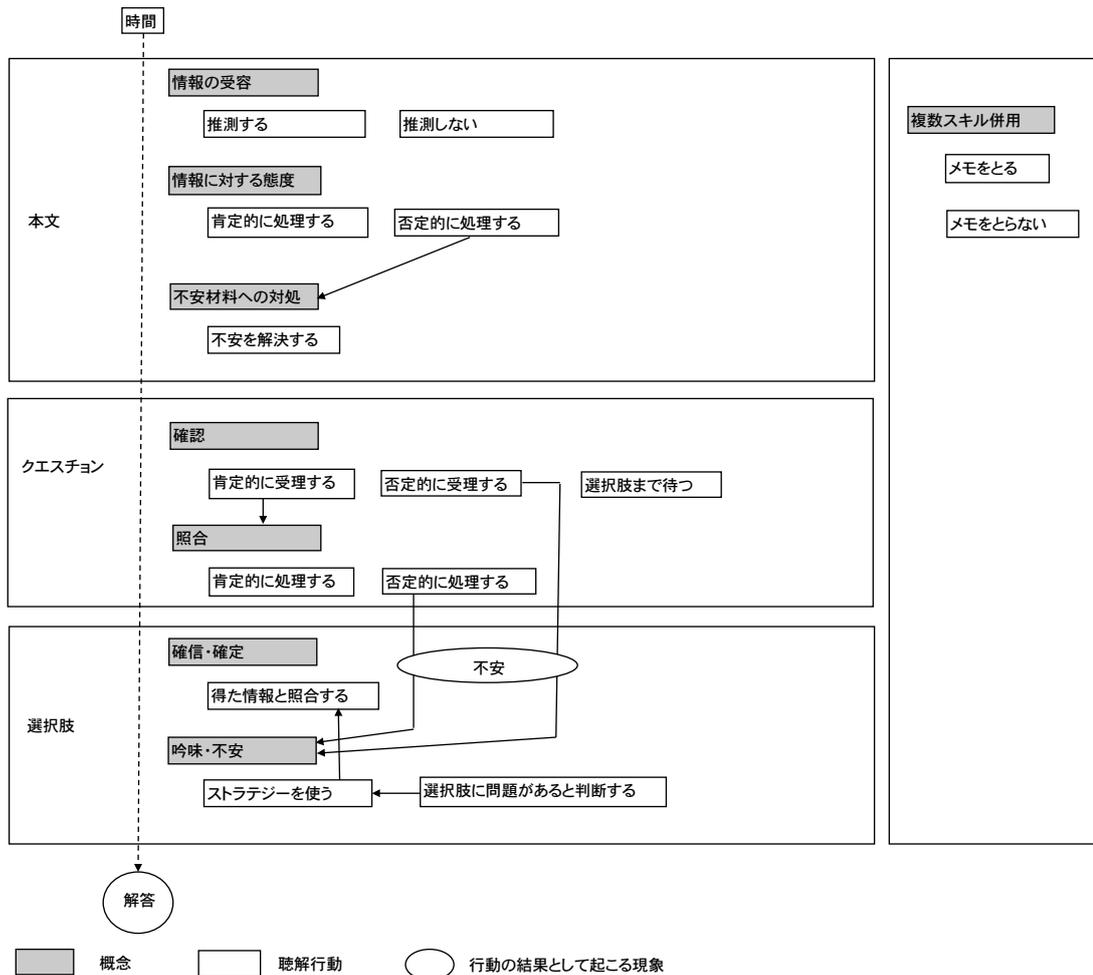


図 7.2 日本人大学生の聴解行動（設問なしのアイテムに対する行動）

図 7.2 は、日本人大学生の設問がないアイテム全体に対して見られた行動である。留学生の図 7.1 (p.155) と比較すると、各行動間の関係 (→) が少ない。設問があるアイテム時でも同様な関係が見られた。設問があるアイテムではなかった「不安」が、選択肢時に見られた。クエスチョン時に「否定的に処理」し、「不安」に繋がっていく。

7.2.1 全体

表 7.5 アイテム全体の概念と聴解行動のヴァリエーション①（日本人大学生設問なし）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
全体	複数スキル併用	聞きながらメモをとる	メモをとる	①解答になりそうな部分 ②記憶のためにする ③流れを思い出す ④全部・できるだけ書いておく ⑤整理する ⑥安心 ⑦習慣	
			メモをとらない	①聞きながらは難しい ②集中しにくい	

全体としてはメモをとるか、とらないかの行動である。設問のある場合と設問のない場合を比較したのが表 7.5.1 である。

表 7.5.1 アイテム全体の聴解行動のヴァリエーション①（日本人大学生設問なしと設問ありの比較）

聴解行動	ヴァリエーション（日本人大学生設問あり）	ヴァリエーション（日本人大学生設問なし）
メモをとる	①記憶のためにする ②課題やキーワードを書く ③整理する ④聴解テストの経験から	①解答になりそうな部分 ②記憶のためにする ③流れを思い出す ④全部・できるだけ書いておく ⑤整理する ⑥安心 ⑦習慣
メモをとらない	①記憶にある ②会話についていけなくなる ③集中しにくい ④解答しにくくなる	①聞きながらは難しい ②集中しにくい

設問があるアイテム時と違う部分は、【設問なし】「①解答になりそうな部分」を、メモにとることである。プリタスクがある場合は、「プリタスクに対するキーワード」をメモするのだが、課題がわからないので「解答になりそうな部分」という表現になった。【設問なし】「④全部・できるだけ書いておく」は重要だからメモをとるのではなく、テストだから出来るだけ多く書いておいたほうがよいと判断した。これは、【設問なし】「⑥安心」というヴァリエーションにも言える。【設問なし】「⑦習慣」は、話を聞く際に記憶すべきことをメモするのが習慣だからというものである。メモをとる際に、「設問がある場合は単語レベルでメモをとったが、設問がない場合は文レベルでとりたくなる」という例があった。焦点を絞ってメモがとりにくいこともあるが、文レベルのほうが本文全体の内容把握を記憶しておけるということだろう。メモをとる行動のヴァリエーションは、留学生とあまり違いは見られない(表 7.5.2)。

表 7.5.2 アイテム全体の聴解行動のヴァリエーション①（留学生と日本人大学生設問なしの比較）

聴解行動	ヴァリエーション（留学生設問なし）	ヴァリエーション（日本人大学生設問なし）
メモをとる	①推測の手がかりにする（語彙） ②記憶のためにする ③集中する ④全部・できるだけ書いておく	①解答になりそうな部分 ②記憶のためにする ③流れを思い出す ④全部・できるだけ書いておく ⑤整理する ⑥安心 ⑦習慣
メモをとらない	①本文が聞けなくなる ②集中できない	①聞きながらは難しい ②集中しにくい

7.2.2 本文

表 7.6 本文時の概念と聴解行動のヴァリエーション②（日本人大学生設問なし）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
本文	情報の受容	解答のための情報を得る	推測する (質問される可能性のある部分に注目する)	言語形式・談話構造 ①談話構造から ②談話のコミュニケーション機能から ③言い換え部分 ④ディスコース・マーカ一部分 テキスト外 ①メイン・トピック ②要点 ③細かい部分 ④言語テスト経験 ⑤質問を仮定 キーワード ①キーワード	
			推測しない (全体を聞く)	①全体理解が大切 ②ポイントを聞く ③余裕がない ④推測できない	
	情報に対する態度	難易に関わる要因から情報を処理する	肯定的に処理する	言語形式・談話構造 ①統語的機能がわかる ②ディスコース・マーカをヒントにする ③構成が明確 ④最後の部分 ⑤強調 テキスト外 ①背景知識・常識・経験は使わない ②既有知識部分は聞かない ③既有知識を使う ④テスト・テクニックを使う ⑤トピックに興味がある キーワード ①要点	
			否定的に処理する	言語形式・談話構造 ①声・話し方に抑揚がない ②語彙が難しい ③語彙使用が曖昧 ④トピックが日常的ではない ⑤くどい ⑥主題がない ⑦内容が具体的ではない ⑦構成が不明確 テキスト外 ①質問を推測できない キーワード ①キーワードになる語彙がわからない	
不安材料 ²	わからない語彙や表現がある	不安を解決する	①文脈から推測する ②回避		

本文の概念は「情報の受容」「情報に対する態度」「不安材料」の3つで、設問がある場合と同じであった(表 7.6)。

まず、「情報の受容」から分析・考察する。「表 7.2.1 本文時の聴解行動のヴァリエーション②(留学生設問ありと設問なしの比較)「情報の受容」(p.158 参照)で分析したように概念は設問がある場合と同様だが、聴解行動に違いが見られたため、聴解行動とヴァリエーションの比較をする(表 7.6.1 次頁)。設問がある場合は、選択的か全体的かを決めて聞くが、設問がない場合は、本文を聞きながら課題を推測し絞り込んで聞か、推測はせず全体把握をするかを定める。この行動は留学生と同様である。推測

する場合は本文内の「言語形式・談話構造」「テキスト外」「キーワード」から推測し、これも留学生と同様である。よって、表 7.6.2 で留学生と日本人大学生の設問がない場合の比較をしながら分析・考察を行う。

表 7.6.1 本文時の聴解行動のヴァリエーション②（日本人大学生設問ありと設問なしの比較）
「情報の受容」

聴解行動	ヴァリエーション（日本人大学生設問あり）	聴解行動	ヴァリエーション（日本人大学生設問なし）
選択的に受容する	①解答のための情報として認知する ②事前推測を訂正する ③不必要な情報は捨てる ④解答を得たあとはあまり集中しない	推測する（質問される可能性のある部分に注目する）	言語形式・談話構造 ①談話構造から ②談話のコミュニケーション機能から ③言い換え部分 ④ディスコース・マーカ一部分 テキスト外 ①メイン・トピック ②要点 ③細かい部分 ④言語テスト経験 ⑤質問を仮定 キーワード ①キーワード
全体的に受容する	①内容把握する ②解答があった後も正確な情報を得るために最後まで聞く ③内容を追う	推測しない（全体を聞く）	①全体理解が大切 ②ポイントを聞く ③余裕がない ④推測できない

表 7.6.2 本文時の聴解行動のヴァリエーション②（留学生と日本人大学生設問なしの比較）
「情報の受容」

聴解行動	ヴァリエーション（留学生設問なし）	ヴァリエーション（日本人大学生設問なし）
推測する（質問される可能性のある部分に注目する）	言語形式・談話構造 ①談話構造から ②談話のコミュニケーション機能から テキスト外 ①メイン・トピックを考える ②主張は何か考える キーワード ①トピック・キーワードを考える	言語形式・談話構造 ①談話構造から ②談話のコミュニケーション機能から ③言い換え部分 ④ディスコース・マーカ一部分 テキスト外 ①メイン・トピック ②要点 ③細かい部分 ④言語テスト経験 ⑤質問を仮定 キーワード ①キーワード
推測しない（全体を聞く）	①全体理解が大切 ②内容把握に支障が出る ③余裕がない ④わからない語彙があり考えられない	①全体理解が大切 ②ポイントを聞く ③余裕がない ④推測できない

「情報の受容」（表 7.6.2）において課題を「推測する」場合、日本人大学生も留学生と同様、複数の課題を推測している対象者が多く、推測した複数の課題の解答となり得る部分を記憶・記録していた。推測する場合、言語【日本人大学生】「③言い換え部分」は、「同じ内容を違う表現で2度言った場合は重要だ」と判断している。【日本人大学生】「④ディスコース・マーカ一部分」は、例えば、「そこで、次回は」という表現から、それまでに話したことに対して何か言うだろうと推測する。テキスト外では、全体の中の【日本人大学生】「②要点」、記憶しにくい数字や説明の【日本人大学生】「③細かい部分」、本文中の意見は課題になりやすいと知っている【日本人大学生】「④言語テスト経験」から解答を絞り込ん

でいる。そして、何を聞くべきかを絞るために【日本人大学生】「⑤質問を仮定」する。留学生より多くの情報元を活用していると言える。

留学生がテキスト外で【留学生】「①メイン・トピックを考える」「②主張は何かを考える」と、全体の中で重要だと思われる部分に焦点を絞っているのに対し、日本人大学生はより細かく本文の部分部分に焦点を当てている。これは、課題を1つに絞り込んでしまうと、推測が間違った場合、解答できなくなるリスクを避けようとした行動である。課題を「推測しない」場合でも、トピック、理由、数字、主張、説明など、問いになりそうな部分を押さえながら【日本人大学生】「②ポイントを聞いている」。しかし、日本人大学生にも、【日本人大学生】「③余裕がない ④推測できない」のように留学生と同様の行動があり、ネイティブにとっても設問がない場合は負荷が大きいことがわかる。

表 7.6.3 本文時の聴解行動のヴァリエーション②（日本人大学生設問ありと設問なしの比較）
「情報に対する態度」

聴解行動	ヴァリエーション（日本人大学生設問あり）	ヴァリエーション（日本人大学生設問なし）
肯定的に処理する	言語形式・談話構造 ①話し方が明確だ ②統語的機能がわかる ③ディスコース・マーカ―をヒントにする ④構成が明確だ ⑤論理的な構造だ ⑥解答が明確に提示される テキスト外 ①知識は使わない ②常識は使わない ③テストの談話構造と判断する ④複数解答候補から絞る ⑤解答は確定せず漠然としておく ⑥トピックから解答を推測する ⑦トピックが身近だ ⑧情報が多い キーワード ①プリタスクから推測する ②全体把握のために拾う	言語形式・談話構造 ①統語的機能がわかる ②ディスコース・マーカ―をヒントにする ③構成が明確 ④最後の部分 ⑤強調 テキスト外 ①背景知識・常識・経験は使わない ②既有知識部分は聞かない ③既有知識を使う ④テスト・テクニックを使う ⑤トピックに興味がある キーワード ①要点
否定的に処理する	言語形式・談話構造 ①発音が不明瞭 ②音声予測の間違い ③解答確信後に逆接で混乱する ④情報の時間差が大きい ⑤解答が不明確だ テキスト外 ①常識と違う	言語形式・談話構造 ①声・話し方に抑揚がない ②語彙が難しい ③語彙使用が曖昧 ④トピックが日常的ではない ⑤くどい ⑥主題がない ⑦内容が具体的ではない ⑦構成が不明確 テキスト外 ①質問を推測できない キーワード ①キーワードになる語彙がわからない

次に、「情報に対する態度」で日本人大学生の設問がある場合と設問がない場合を比較する（表 7.6.3）。「肯定的に処理する」では、設問があるアイテムには見られなかった【設問がない】「④最後の部分 ⑤強調」と、テストの談話であることにより注意を払っている。テキスト外でも、【設問がない】「②既有知識部分は聞かない」と聞く負担を軽減し、【設問がない】「③既有知識を使う ④テスト・テクニックを使う」のように言語知識を補助する行動をとっている。「否定的に処理する」では、設問があるアイテム時にはなかった【設問がない】「②語彙が難しい ③語彙使用が曖昧 ④トピックが日常的ではない

⑤くどい ⑥主題がない ⑦内容が具体的ではない」があり、本文を聞きながら解答が明確に得られないことへの苛立ちが感じられ、それが本文内容への不満となってヴァリエーションに表れている。これらは、プリタスクがないため、複数あるいは多様な推測をせまられた結果であり、「自分の推測が的中しないのではないか、推測を絞り込めない」という心理的な影響があったと考えられる。

表 7.6.4 本文時の聴解行動のヴァリエーション②（留学生と日本人大学生設問なしの比較）
「情報に対する態度」

聴解行動	ヴァリエーション（留学生設問なし）	ヴァリエーション（日本人大学生設問なし）
肯定的に処理する	言語形式・談話構造 ①統語的機能がわかる ②ディスコース・マーカ―をヒントにする ③構成が論理的 ④構成が対比 テキスト外 ①背景知識・常識・経験を使う ②トピックが得意（既知語彙） ③トピックがわかると談話構成が予測できる ④テスト・テクニックを使う ⑤先読みする キーワード ①頻度が高い ②全体を聞くとわかる	言語形式・談話構造 ①統語的機能がわかる ②ディスコース・マーカ―をヒントにする ③構成が明確 ④最後の部分 ⑤強調 テキスト外 ①背景知識・常識・経験は使わない ②既有知識部分は聞かない ③既有知識を使う ④テスト・テクニックを使う ⑤トピックに興味がある キーワード ①要点
否定的に処理する	言語形式・談話構造 ①スピードが速い ②発音が不明瞭だ（活用語の音を聞き間違える・特殊音） ③声・話し方に抑揚がない ④語彙が難しい（わからない・漢字熟語が多い） ⑤トピックが日常的ではない（専門用語） ⑥トピック展開が中途半端 ⑦談話構造が日常談話と違う ⑧情報量が多い テキスト外 ①背景知識・常識・経験がない ②トピックに興味がない・反発する ③設問がない形式に慣れていない キーワード ①キーワードの語彙がわからない	言語形式・談話構造 ①声・話し方に抑揚がない ②語彙が難しい ③語彙使用が曖昧 ④トピックが日常的ではない ⑤くどい ⑥主題がない ⑦内容が具体的ではない ⑧構成が不明確 テキスト外 ①質問を推測できない キーワード ①キーワードになる語彙がわからない

表 7.6.4 で、留学生と日本人大学生の設問がない場合の聴解行動を比較した。両者とも非常によく似た聴解行動を取っている。留学生が「肯定的に処理する」テキスト外の【留学生】「①背景知識・常識・経験を使う」を言語知識の補助に使用していることに対し、日本人大学生は【日本人大学生】「①背景知識・常識・経験は使わない」と、テストであるがゆえ本文内で表現されている内容だけで解答をすべきと考えていることがわかる。

表 7.6.5 本文時の聴解行動のヴァリエーション②（日本人大学生設問ありと設問なしの比較）
「不安材料」

聴解行動	ヴァリエーション（日本人大学生設問あり）	ヴァリエーション（日本人大学生設問なし）
不安を解決する	①メモする	①文脈から推測する ②回避

表 7.6.6 本文時の聴解行動のヴァリエーション②（留学生と日本人大学生設問なしの比較）
「不安材料」

聴解行動	ヴァリエーション（留学生設問なし）	ヴァリエーション（日本人大学生設問なし）
不安を解決する	①文脈から推測する ②母語から推測する ③英語から推測する ④発音から漢字を推測する ⑤語彙分析から推測する ⑥勘	①文脈から推測する ②回避
不安を解決できない	①集中力がきれる ②思考が止まる	

最後に、「不安材料」についてだが（表 7.6.5）、設問がある場合、日本人大学生はその定義が「記憶に残りにくい」であり、それは【設問あり】「①メモする」ことで解決していた。設問がない場合の定義は「わからない語彙や表現がある」であり、これは留学生の「わからない語彙や表現、複数解答候補がある」の前半部分と同様言語知識に関することであった。留学生に見られた「複数解答候補がある」であるが、日本人大学生の中で複数の課題を推測していた対象者は、それぞれの解答を本文中で得られたため、不安にならなかった。

不安を解決する場合、留学生は様々な情報を使用し推測するが、日本人大学生は【日本人大学生】「①文脈から推測する」のみであった（表 7.6.6）。留学生は不安が解決しないと「その後聞けなくなる」という現象が起こったが、日本人大学生はそのようなことはなく、【日本人大学生】「②回避」をして最後まで聞くという行動となった。日本人大学生に「わからない語彙・表現がある⁴」があったが、本文内容から推測し、不安を解決した。よって、日本人大学生は「不安を解決できない」という聴解行動はなかった。

⁴ わからない語彙として「ディベート」を数人が挙げた。

7.2.3 クエスチョン

表 6.13 ポスト・クエスチョン時の概念と聴解行動のヴァリエーション⑥（日本人大学生設問あり）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
ポストQ	確認1	もう1度聞く	再確認する	①プリタスクの情報の再確認 ②解答に自信がなかった	
			無視する	①覚えていない ②聞いたが要らない	

6章では、設問があったため、ポスト・クエスチョンは表 6.13 (p.149 と同じ表) のように確認するための機能だったが、7章では本文後に初めて聞くクエスチョンであり、機能が違う。そのため、ここでは、比較ではなく、設問がない場合の聴解行動として独立して分析をした(表 7.7)。その後、設問がない場合の留学生と比較し、分析・考察を行う(表 7.7.1 表 7.7.2)。

表 7.7 クエスチョン時の概念と聴解行動のヴァリエーション③（日本人大学生設問なし）

アイテム	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
クエスチョン	確認2	推測したタスクかどうか確認する	肯定的に受理する	①推測とあっている ②推測より簡単なQだ ③記憶にある	
			否定的に受理する	①推測できなかった ②推測と違う	不安
			選択肢まで待つ	なし	
	照合	本文から得た情報を解答になりうるか照合する	肯定的に処理する	①解答を言語化する ②推測と違っていたが、得た情報から探す	
			否定的に処理する	①本文内容把握が完全ではない ②時間的余裕がない ③選択肢に不安 ④曖昧な記憶	不安

留学生と同様、日本人大学生も、クエスチョン時に「確認2」と「照合」を行う。まず、「確認2」の聴解行動について、留学生と日本人大学生の比較をする(表 7.7.1 次頁)。「肯定的に受理する」では、【日本人大学生】「①推測とあっている ②推測より簡単なQだ ③記憶にある」と、自信を持っていることがうかがえる。「確認2」が済むと「照合」へすすむ(表 7.7.2 次頁)。「否定的に受理する」は、【日本人大学生】「①推測できなかった ②推測と違う」の2つであり、留学生に見られた【留学生】「③語彙がわからない ④意味がわからない」というヴァリエーションはなかった。日本人大学生は、ここで「不安」を抱える。日本人大学生も、留学生と同様計画的にクエスチョン時では解答を決めず、「選択肢まで待つ」対象者がいた。

「照合」時の「否定的に処理する」に【日本人大学生】「③選択肢に不安 ④曖昧な記憶」があり、本文情報の取得の不完全さが現れている。1～2分程度の長さの本文だが、設問がない場合は、何を記憶すべきか絞れず、多くの記憶容量が必要になるため、日本語ネイティブにとっても記憶に負荷がかかった。設問があるアイテムでは解答に自信があったが、設問がないアイテムではその自信が揺らいでいることがわかる。クエスチョンを「確認」し「照合」した時点で、「否定的に処理した」場合は「不安」を

抱えて選択肢を聞くことになる。本文理解の際の心理的な面では、日本人大学生のほうが多くの留学生より優位だったが、クエスチョンを聞き終わった段階では、留学生と日本人大学生に差はないと思われる。

表 7.7.1 クエスチョン時の聴解行動のヴァリエーション③（留学生と日本人大学生設問なしの比較）
「確認」

聴解行動	ヴァリエーション（留学生設問なし）	ヴァリエーション（日本人大学生設問なし）
肯定的に受理する	①推測とあっている ②聞き逃した部分ではない ③納得する	①推測とあっている ②推測より簡単な Q だ ③記憶にある
否定的に受理する	①推測できなかった ②推測と違う ③語彙がわからない ④意味がわからない	①推測できなかった ②推測と違う
選択肢まで待つ	なし	なし

表 7.7.2 クエスチョン時の聴解行動のヴァリエーション③（留学生と日本人大学生設問なしの比較）
「照合」

聴解行動	ヴァリエーション（留学生設問なし）	ヴァリエーション（日本人大学生設問なし）
肯定的に処理する	①解答を言語化する ②選択肢で消去法が使える ③推測と違っていたが、得た情報から探す	①解答を言語化する ②推測と違っていたが、得た情報から探す
否定的に処理する	①本文内容把握が完全ではない ②理解した部分ではない ③聞き逃した部分だ	①本文内容把握が完全ではない ②時間的余裕がない ③選択肢に不安 ④曖昧な記憶

7.2.4 選択肢

表 7.8 選択肢時の概念と聴解行動のヴァリエーション④（日本人大学生設問なし）

7/7M	概念	定義	聴解行動	ヴァリエーション	行動の結果として起こる現象の例
選択肢	確信・確定	得た情報との一致度が高い	得た情報と照合する	①自分の解答と一致 ②解答がある	
	吟味・不安	選択肢の内容と本文内容を吟味する 解答がないはずがない	ストラテジーを使う	①背景知識・常識（正誤は関係ない）から推測する ②消去法 ③本文の語彙との一致・類似から選ぶ（相違はない） ④メモから推測する ⑤言い換えの意味と自分の答えを比較する ⑥テスト・テクニックを使う ⑦語彙の解釈 ⑧選択肢の解釈を広げる ⑨正答誤答が明確	
			選択肢に問題があると判断する	①期待した選択肢がない ②言い換えがされている ③選択肢同士が似ている ④選択肢が曖昧	

クエスチョン時に得られた解答が選択肢にあれば、「解答を確信・確定」することになる。自分の解答が選択肢にない、あるいはクエスチョン時に解答が明確になっていない場合は「吟味・不安」に移る（表 7.8）。設問がある場合は、選択肢を聞いた時点で解答を「確定」させるのだが、設問がない場合は対象者

自身が推測した課題に対する解答を「確信・確定」する（表 7.8.1）。概念としては違うのだが、聴解行動としては「得た情報と照合する」である。

表 7.8.1 選択肢時の聴解行動のヴァリエーション④（日本人大学生設問ありと設問なしの比較）
「確定」と「確信・確定」

聴解行動	ヴァリエーション（日本人大学生設問あり）	ヴァリエーション（日本人大学生設問なし）
得た情報と照合する	①自分が決めた答えと一致している ②言い換えと照合する ③消去法	①自分の解答と一致 ②解答がある

表 7.8.2 選択肢時の聴解行動のヴァリエーション④（留学生と日本人大学生設問なしの比較）
「確信・確定」

聴解行動	ヴァリエーション（留学生設問なし）	ヴァリエーション（日本人大学生設問なし）
得た情報と照合する	①自分の解答と一致 ②解答がある	①自分の解答と一致 ②解答がある

設問がない場合の留学生と日本人大学生の選択肢時の聴解行動を比較する（表 7.8.2）と、両者とも同じ行動をしていることがわかる。

表 7.8.3 選択肢時の聴解行動のヴァリエーション④（日本人大学生設問ありと設問なしの比較）
「不信」と「吟味・不安」

聴解行動	ヴァリエーション（日本人大学生設問あり）	ヴァリエーション（日本人大学生設問なし）
ストラテジーを使う	①自分の解答の先に何があるかを考え選択する ②作成者の立場で考える	①背景知識・常識（正誤は関係ない）から推測する ②消去法 ③本文の語彙との一致・類似から選ぶ（相違はない） ④メモから推測する ⑤言い換えの意味と自分の答えを比較する ⑥テスト・テクニックを使う ⑦語彙の解釈 ⑧選択肢の解釈を広げる ⑨正答誤答が明確
選択肢に問題があると判断する	①選択肢が本文の要点ではない ②自分の解答と違う ③正確な答えがない ④正答が複数ある	①期待した選択肢がない ②言い換えがされている ③選択肢同士が似ている ④選択肢が曖昧

解答を確信・確定できなかった場合は、対象者自身が持っている情報を「吟味（定義：選択肢の内容と本文内容を吟味する）」したり、「不安（定義：解答がないはずかない）」を感じたりする。聴解行動としては「ストラテジーを使う」と「選択肢に問題があると判断する」の2つであり、これは設問がある場合と同様である（表 7.8.3）。設問がある場合は「自分の解答との一致度が低い」ため、アイテム自体に「不信」を抱きつつ、「ストラテジーを使って」解答を得ようとする、あるいは「選択肢に問題があると判断して解答放棄に至ることもあった。設問がない場合は「ストラテジーを使う」のヴァリエーションの多さからもわかるように、解答を得るために様々な行動をしていることが明らかである。

表 7.8.4 選択肢時の聴解行動のヴァリエーション④（留学生と日本人大学生設問なしの比較）
「葛藤・吟味」と「吟味・不安」

聴解行動	ヴァリエーション（留学生設問なし）	ヴァリエーション（日本人大学生設問なし）
ストラテジーを使う	①背景知識・常識から推測する ②消去法 ③本文の語彙との一致・類似・相違から選ぶ ④メモから推測する ⑤メモにある語彙を選ぶ ⑥言い換えの意味と自分の答えを比較する ⑦言い換えの語彙を漢字になおす ⑧可能性のある選択肢を比較する ⑨本文の最後の部分のものを選ぶ ⑩メイン・ポイントや主張を選ぶ ⑪テスト・テクニックを使う ⑫4つの選択肢からわからない語彙を推測する	①背景知識・常識（正誤は関係ない）から推測する ②消去法 ③本文の語彙との一致・類似から選ぶ（相違はない） ④メモから推測する ⑤言い換えの意味と自分の答えを比較する ⑥テスト・テクニックを使う ⑦語彙の解釈 ⑧選択肢の解釈を広げる ⑨正答誤答が明確
選択肢に問題があると判断する	①期待した選択肢がない ②言い換えがされている ③選択肢同士が似ている ④適当な答えがない ⑤聞き取れた情報だけでは解答できない	①期待した選択肢がない ②言い換えがされている ③選択肢同士が似ている ④選択肢が曖昧

表 7.8.4 は設問がない場合の留学生と日本人大学生の聴解行動を比較したものである。留学生には「自分が得た解答は本当に正しいのか」という「葛藤」があり、解答を得るために選択肢の内容と本文内容を「吟味」して解答を得ようとする。日本人大学生は「葛藤」ではなく、当然自分の得た情報の中に解答があると確信しており「吟味」するが、選択肢に適当な解答が見つからない場合は「不安」を抱える。日本人大学生は、聴解に自信があるため、「解答がないはずがない」と考えるため、再び自分が得た情報と選択肢の「吟味」を繰り返す。設問があるアイテムでは、日本人大学生が「ストラテジーを使う」のは、「選択肢に問題がある」場合のみで、解答を得るために初めから使用することはなかった。しかし、設問がない場合は、解答を得るために「ストラテジーを使う」。ストラテジーのヴァリエーションは、留学生と比較すると推測の情報元の数は少ないが、多様な推測をしていることがわかる。日本人大学生特有に見られたヴァリエーションとして、【日本人大学生】「⑦語彙の解釈 ⑧選択肢の解釈を広げる」があった。これが行なわれたのは、4番のアイテム（p.153）で、『循環』が物理的な意味か科学的な意味かで迷った」という例や、「蒸発は大気中で水が循環しているともとれる」という例があった。また、5番のアイテム（p.153）で、「賛成か反対かのどちらかの側にたって話す場合、相手を見て話していると考えられるのではないか」などの例が挙げられた。前述の語彙や選択肢の意味を広く解釈することが、解答を得るためには負の影響を及ぼした。プリタスクがなかったため、本文の内容も語彙の意味もあまり絞り込まなかったゆえ、様々な立場や視点を考えてしまった結果である。今回使用したアイテムは、設問があったアイテムから設問を外したため、設問があれば迷わなかったのかもしれない。しかし、談話としては、不完全だったということになる。選択肢の意味解釈も、設問があるアイテムに比べ広く捉えており、正答を選ぶ際に迷った対象者もいた。設問があるアイテムでは、テスト・アイテムに問題があるとして解答放棄をした対象者がいたが、設問がないアイテムでは全員解答を選んだ。これは留学生と同様の行動である。誤答を選んだ対象者もいたが、設問があるアイテム時に見られたテスト・アイテムに対しての「不信」はなく、また、留学生が設問がないアイテムに抱いた「不満」も見られなかった。これは、次に述べることに関連するが、設問がない聴解をより評価していることも影響していると思われる。

最後に、日本人大学生の設問の有無に対する意見を分析する。日本人大学生は、対象者全員が聴解能力を測定するのであれば設問が最初にあるアイテムはよくないと述べている。理由は、設問がないテスト・アイテムでは、①全体の流れを聞いて理解できる ②緊張感を持って聞かなければならないので最大限の力を使う ③全体を意識して聞く ④全体を聞いた中で判断する ⑤文章理解力が聴解能力には必要 ⑥クエスチョンをより正確に聞く（設問がある場合と比較して）からである。設問がある場合の欠点として、①単語のみの理解でも答えられる ②プリタスクの部分のみに集中し他はわからなくても答えられる ③選択肢の理解のみでも答えられる ④音の記憶のみでも答えられる を挙げた。これは、偶然性（まぐれ当たり）を極力避けるべきであるという意見である。留学生は、設問の有無に対し賛否両論であったが、日本人大学生は一律に設問がないテスト・アイテムの評価が高かった。日本人大学生は設問があるアイテムより、設問がないアイテムのほうが正答率が高かったことから⁵、ネイティブの直感として設問がないアイテムが自然だと感じ取っているのかもしれない。

7.3 設問がないアイテムにおける2者の総括と考察

設問がないアイテムの聴解行動においても、留学生の特徴は不安材料への対処である。本文とクエスチョンで不安材料があった場合、不安を解決できず、その負の影響が選択肢が終わるまで続いた。この行動は、設問があるアイテムでも同様であった。設問がないアイテムでは、解答となる情報を絞り込んで聞けなかったことが不安を助長し、記憶に負荷をかけた。一部の留学生にとって、設問がないアイテムのほうが困難度が高く感じられた原因でもある。日本人大学生の特徴は、設問があるアイテムでは見られなかった「不安」が生じたことである。設問があるアイテムでは、本文の不安材料は本文終了時までに解決したが、設問がないアイテムでは、本文終了後にクエスチョンを聞くため、クエスチョン時に本文で得た情報に対しての不安が発生した。本文を聞く際に、対象者はクエスチョン（タスク）を推測し、情報を絞り込んで聞いていた。そのため、クエスチョンと自分の予測との違いに不安を抱いたからである。日本人大学生は、設問があるアイテムでは自分の解答に絶対的な自信を持っており、適当な選択肢がない場合は、テスト選択肢に対して「不信」を抱いた。しかし、設問がない場合は、本文から得た情報とクエスチョンとの照合にあまり時間をかけられなかったため、自分の解答に対する自信は設問があるアイテムより低かった。留学生も日本人大学生も設問がある言語テストの経験が豊富であり、日常聴解の1つである大学の授業・講義で重要部分を絞り込む聞き方をしているため、課題を絞ったほうが正答を得やすいという聴解行動をとった対象者もいた。その場合、予測した課題とクエスチョンが違った場合の心理的ダメージは大きい。今回使用した設問がないアイテムは、元来設問があるアイテムであったため、一概にアイテムの問題とは決められないが、本文と選択肢に不完全さ、曖昧さがあったという指摘である。

設問がないアイテムの場合、留学生にも日本人大学生にも記憶の負荷が多くかかる。これは、テスト・アイテムがすべて短い談話だったため、できる限り多くのことを記憶しようとした、また記憶できた結果である。例えば、長い談話であつたら、クエスチョンの絞り込みをせざるをえなくなり、できる限り多くのことを記憶する行動は、大切なことを聞き逃すリスクが増加するためしないだろう。また、記憶の負担を軽減するために、メモをとるという行動が必然的になる。2010年に日留試のシラバスが改定された際、それまですべて一問一答だった読解のテスト・アイテムに、一問複答（二答と三答）のアイテ

⁵ 設問がないアイテムの正答率を記す。日本人大学生の設問があるアイテムの正答率は90%、設問がないアイテムの正答率は93%にあがった。

ムが加わった。一問二答の本文テキストは一問一答のものと比較して顕著に長くなったわけではないが、一問三答の本文テキストは一問一答の本文テキストの約3倍の長さになっている。読解の1アイテムの本文テキスト量が増加したことは、今後聴解テスト・アイテムでも、本文の長さが伸びる可能性がでてきたと考えられる。留学生が聴解能力を、①全文を理解すること ②何を聞かなければならないかを見つけること ③聞く目的によって聞き方を変えること（7.14「選択肢」p.164を参照）と考えていることは、長い談話を聞く必要性を示唆していると言えよう。

Bachman & Palmer (1996, 邦訳 pp.71-78)は、言語テストで言語能力を測定したい場合、言語知識と方略的能力以外にも ①個人的特性 ②話題の知識 ③情意スキーマ も考えなければならないとしている。なぜなら、これらはテスト得点に影響力を持っており、これらの特性が受験者の言語運用を促進するように言語テストを設計することが望ましいからである。「①個人的特性」は、年齢、性別、国籍、在住状況、母語、一般的教育の水準とタイプ、特定のテストに関する受験経験および準備と量と種類である。「②話題の知識」とは、長期記憶の知識構造であり、実生活の中で言語使用の情報の基礎となるものである。「③情意スキーマ」は、話題の知識と感情的に関係し合い、テスト内容と類似の文脈の過去の感情的経験から意識的、あるいは無意識的に評価することである。

表 7.9 留学生、日本人大学生の個人的特性と話題の知識の違い

		留学生	日本人大学生
個人的特性	年齢	20～34（平均 25.2 歳 中央値 23 歳）	18～19（平均 18.3 歳 中央値 18 歳）
	国籍	日本以外	日本
	在住状況	大学・大学院生	大学生
	母語	日本語以外	日本語
	一般的教育の水準とタイプ	高校卒業以上	高校卒業以上
	特定のテストに関する受験経験および準備と量と種類	言語テスト経験多	言語テスト経験多
話題の知識	話題の知識の経験年数（年齢）	20年以上	約20年

日留試の聴解テストに、「①個人的特性 ②話題の知識 ③情意スキーマ」が、どのような関わりを持っているのかを考えてみる。「①個人的特性」と「②話題の知識」の違いを見るために、今回の対象者である留学生、日本人大学生の特性を表 7.9 にまとめた。表 7.9 の灰色の部分には2者間の共通項目・近似項目である。なお、6章「6.3.1 対象者」(p.120)で対象者の条件としなかった性別は考慮しない。出身国の違いは、日本・日本語と日本・日本語以外とする。在住状況は、日本の教育機関で勉学をしている留学生と大学生であり、日本に住んでいることからこの時点での生活環境は近似であるとした。一般的教育の水準とタイプは、留学生が学部生から大学院生までいること、また各国により違いはあると考えられるが、高校卒業時までの教育を終えているという事実において、留学生と日本人大学生は類似であるとした。同様のことが話題の知識についても言え、実生活年数から、一般的知識は近似しているとした。

話題の知識は年齢に比例すると考え、留学生を20年以上、日本人大学生を約20年とした。

表7.9から、「①個人的特性」に関して、留学生と日本人大学生は国籍と母語以外はほぼ近似であると考えられる。「②話題の知識」に関してであるが、留学生と日本人大学生の経験年数に大差はない。よって、今回研究対象者とした留学生と日本人大学生、両者の日留試に対する意見や考え、感想などが聞けたことは、非常に有意義であった。彼らの発した言葉すべてに、日留試を改善するためのヒントがあると考えられる。「①個人的特性」の項目から考えると、日留試の聴解テストには大学場面が多いことから、妥当だと判断する。受験者が言語テスト経験を多く持っていることから、画一的なテスト形式ではなく、目的に応じたテスト形式が望まれる。「②話題の知識」に関しては、20年ほどの人生経験者にとって、興味があり、有益だと感じられるものがよい。西川・小牧(2002, p.26)は、聞き手は聴くための刺激を選ぶものであり、選ばれる刺激は適度な新奇性と興味、そして適度な容易さであると述べている(同掲 pp.26-27)。新奇性と興味に関しては、多様な受験者すべてに対応するのは難しいが、日本の大学生が日常的にどのような話題に興味を持っているのか、頻度的には何についての話が多いのかの調査が参考になる。テスト形式も、興味を得るために、出題する談話展開の複雑さと内容の密度をあげられる長い談話は有効であろう。

「③情意スキーマ」として、今回、留学生と日本人大学生に見られた「不安」「わからない」という感情と、日留試の聴解テスト・アイテムに対する評価を考察する。今回使用した聴解テスト・アイテムは、日本人大学生にとって簡単であったが、唯一わからない語彙があり、それを「回避」した対象者がいた。日常生活では、ネイティブであっても、わからない語彙や表現は実在する。その際、わからない語彙や表現に対して、回避行動をとることがあり、それは自然な行動である。よって、「回避」行動を行うことに負の感情は生まれにくい。留学生にも「回避」行動は見られたが、そこには、日本語能力の不足に対する負の感情が含まれており、また、「回避」行動の割合が多ければ、談話の意味理解ができなくなり、テストとしては良くない結果となる。本研究の留学生対象者は、中級レベルの修了者から超級レベルの対象者までを含んでいたため、日本語能力のレベルにより、わからない語彙や表現の量に違いがあった。さらに、日本語クラスの中級修了直後、上級前半のレベルのクラスの対象者は、大学では外国語としての日本語授業を主に受けており、わからない語彙や表現は回避せず、意味確認をすることが日常的だったと考えられる。また、初級から中級の日本語授業では個別的要素を測定するテストが頻繁に行なわれる傾向があり、テストの文脈から推測することより語彙や表現が既知であることに拘っているのではないか。このような学習経験から、言語知識が不十分な場合に、否定的に処理する行動が多く見られたと推測する。

日留試の測定対象能力は、「単に日本語に関する知識の有無や知識の量を測定するアチーブメント・テストではない」と公言されている。堀井(2003, p.115)は、日留試におけるアカデミック・ジャパニーズは、日本の大学で勉強するのに必要な日本語能力、すなわちコミュニケーション能力と思考(認知)能力であると説明している。場面・状況、タスクを理解する能力が思考(認知)能力であり、談話を聞きながら場面・状況を把握し、何が重要かを判断するのもまた思考(認知)能力である。テスト・アイテムの設問の有無は思考(認知)能力を測定する役割であり、6章と7章の留学生、日本人大学生の聴解行動の結果から、設問があるアイテム、設問がないアイテム、2つとも聴解テストに必要なであると結論づける。

7.4 提案と今後の課題

ここでは、6章と7章で得られた留学生・日本人大学生、2グループの分析結果をふまえ、日本語授業や中間テスト、期末テストなどの小規模な聴解テストと日本語授業で行う聴解練習について述べる。学習者は大規模テストの得点をあげるためだけに学習しているのではない。筆者は学習者の聴解能力をあげるために聴解行動の研究結果が聴解テスト作成、聴解教材開発、聴解練習に役立つことを願っている。

聴解テストの聴解行動の分析結果から、聴解能力には経験や環境が影響することがわかった。日本人とは違う環境で生活してきた留学生ではあるが、一般的話題に対する知識とストラテジー使用に関する経験値は日本人大学生とほぼ差はない。留学生に必要な要素は、日本の大学で必要とされている言語知識と日本の大学の授業形式、そして日本の大学のシステムに関する知識である。そして、これらすべての知識の根底には日本の社会文化的な要素がある。

山下(2000)は、留学生と日本人大学生の講義理解に関わる因子とストラテジーを比較した。その結果、留学生には教科書や板書などの「文字情報」、教師の話し方や講義の仕方などへの「慣れ」、教室内でのやりとりや人数などの「教師と学生の関係」、講義内容の「講義構成」が影響していることがわかった。日本人大学生は、教師と学生の関係と講義や授業などの「教え方」、専門用語や教師特有の「話し方」、「教師への慣れ」、「文字情報」が因子として挙げられた。日本人大学生は、教師の個々の話し方や進め方に対する「慣れ」だが、留学生は日本の大学の授業の独特な話し方やその背後の文化的側面に対するものも含めた「慣れ」であることに注目している。留学生にとってのストラテジーは、全体の意味やキーワードに注意するなどの「授業の聞き方を工夫する：聞き取りⅠ」、わかるか否か、大事か否かなど「聞き取りに関する第一段階：聞き取りⅡ」と「予習・復習」であった。日本人大学生のストラテジーは、留学生の聞き取りⅠとⅡの2つをあわせた「聞き取り」、「予習・復習」、「質問」であった。留学生は、意識的に理解を助けるためのストラテジーを使用している。留学生は、聞き方の工夫を言語知識に注目する「聞き取りⅠ」と、課題発見に注目した「聞き取りⅡ」にわけ、日本語知識不足をストラテジーで補っていることを明確にした。

山下(2000)の研究から留学生は講義理解のために、言語知識に加え日本の社会文化的知識を学習し、課題発見能力をのばし、ストラテジーの積極的な使用を勧めることがよいと言える。では、日本語授業での聴解練習に日留試のような聴解テストを利用して、前述のような知識の習得と聴解方法を学習できるだろうか。本研究で使用した講義場面のアイテムは、非常に短く、録音されたCDの音声の使用から、実際の講義に対する「話し方への慣れ」という点では難しいかもしれないが、トピックを予測する、キーワードを発見する、未知の語彙や表現の意味を推測する、経験や世界知識から推測する、などの練習は可能である。今後、多様な談話構成、多様な形式(設問がある・設問がないなどや談話時間が短い・長い、3人以上の会話など)、多様な話し方、日本語の講義場面での教師と学生とのやりとり、そして文字情報をともなった聴解テスト、聴解練習用のテキストなどの開発が望まれる。ストラテジー使用に関しては、本研究では聴解テストのアイテム構成要素ごとの聴解行動と、各行動の関連性を見た。留学生は、情報理解に問題が発生した場合、不安材料を抱えて次の構成要素の聴解を続けていく。聴解理解のための様々なストラテジー使用も見られたが、本研究で明らかになった「不安」への対処のストラテジーを、対象者1人1人のデータから丁寧に分析する必要がある。日留試は公開されており、授業での使用も可能である。多少画一的な形式ではあるが、そこを利用し、わからない語彙や表現があっても聞き飛ばし、負の心理的影響に対して慣れるような速聴を繰り返すことも「慣れ」となり、不安を和らげる効果が期待できるのではないだろうか。

福田 (2005, p.77) は、日本語母語話者と日本語学習者 (1・2級) の文聴解 (1文の聴解) の即時処理を比較した結果、作動記憶容量の個人差による影響は見られなかったが、語の曖昧さの影響が見られ、多義性のある語 (同音異義語など) の処理時間は、日本語母語話者に比べ日本語学習者のほうが長くかかる」と述べている。本研究で使用した日留試の聴解テスト・アイテムの談話内の語の曖昧さは、留学生、日本人大学生双方から指摘された⁶。これは、福田が挙げた多義語の知識の問題のほかに、課題の曖昧さ、選択肢の言い換え、短い談話が原因と考えられる。日本語教師がオリジナル聴解教材や聴解テストを作成するにはこのようなことに注意する必要がある。

富田・小栗・河内 (2011, pp.77-79) は、聴解は音声テキストとして受け取る「テキスト化」の段階と、次に語の連続として内容理解をする「意味化」の2つであり、第二言語学習者には第一段階の「テキスト化」が重要であるとした。その言語の音として捉えられるのは、その言語の音声体系の枠組みを持っているからである。もし、その枠組みが有効でない場合、音は捉えられず「分類できない音 (雑音)」となりテキスト化ができなくなる。このために処理が遅れ、テキスト化の乱れを生じ、次の段階の意味化にも支障をきたすようになる。同時に、分類不可能な音があるために、その前後の音も、余分な作業に注意力をとられたために聞き逃す原因となることもあるとしている。Buck (2001, pp.4-8) も、聴解は自動的処理とコントロール処理が必要で、第二言語学習者は、スピードが速い発話などで自動的処理ができなくなると、語彙や文法などに注意を向けてしまい、広く意味を捉えようとしなくなると述べている。本研究で、この特徴は留学生の日本語レベルが低い対象者に多く見られた。わからない言葉 (音) があると、そこで理解のための思考が止まり、行動も止めてしまうという例である。富田・小栗・河内 (2011, p.84) は、母語話者は話し手の言外の意図などを心理的な表れと捉えたり、話し手の出身地や教育レベルを知る鍵として、次に何を言うのかの予測などに利用していると説明する。本研究では、留学生にも日本人大学生にも、自らの経験から推測や予測をたてる行動が見られた。しかし、例えば、話し手の地位や教育レベル、話し手と聞き手の関係、習慣や風習などの社会文化的な経験や知識から予測するような内容の場合、留学生は日本人大学生より不利となるだろう。日本語の特徴として、間接的表現、「以心伝心」などは話者同士の距離や立場、状況などによって多様であり、成人の日本人でもすべて理解できるわけではないが、これが日本語コミュニケーションを円滑にする要素の1つでもある。テストでは、意味解釈が複数のものは正答が1つに定まらず出題を避ける傾向にある。しかし、これを避けている限り、文の表面的な意味解釈しかできず、言語知識が高く、流暢な日本語を使用してもコミュニケーションにおいて円滑さに欠ける日本語学習者が、テストではよい得点をとる危険性がある。大規模テストでは、結果の信頼性が重要であること、テスト場面という日常とは異なる聴解を強いることから、正答は1つでなければならない。しかし、日本語授業内では、教師が対応でき、解答としてどのような可能性があるのかなども学習者と話し合うことができる。市販されている聴解テキスト、公開されている聴解テストを利用し、解答形式を自由解答にする、設問を外すなどの工夫が有効であると考えられる。

日留試のような大規模テストは、その波及効果⁷を考え、語用論的な内容をどのように出題するのかを

⁶ 6章と7章で日本人大学生は本文を聞いた際「否定的に処理する」で「発音が不明瞭」「解答が不明確」「構成が不明確」などを挙げている。また、選択肢に対しては「正確な答えがない」「本文の要点ではない」「選択肢が曖昧」などが挙げられている。

⁷ Heaton (1988, p.252) はテストの波及効果について以下のように述べている。言語テストが言語指導に対する影響は思った以上に高いものであると強調したい。なぜなら、テスト作成者は、テストの統計や他のテスト基準に容易にのめり込んでしまうものだからである。客観テストは、流暢さより正確さの強調になってはいないか、教師の努力にも関わらず言語学習のコミュニケーション側面より正確さを強調し、テスト・テクニックなどで言語学習の否定的側面を出してはいないか、など人工的に作られた状況での言語教育を奨励してしまい、それゆえ言語学習の動機付けをそいでしまっている。ある技能を習

示していく責任がある。日留試は、その得点で留学が決まるため、留学生の人生を左右する重要な役割を担っている。日留試の得点に、言語知識以外に社会文化的知識がどの程度取り入れられるのか、得点の精度は高いのか、そもそも点数化が可能であるのかと、課題は多い。テストの限界として言語知識以外、不確定な要素は測定しないという選択肢も含め、今後の大規模テストとテスト研究の動向に注目したい。

富田他（2011, pp.83-87）は、学習者が母語話者の聴解に近づくために以下のような練習を提示している。

1. 聴解練習のみでなく音声言語の習得の土台となる文字レベルの処理のスピードをあげるために、音を聞きながら読む。その際、文字列を音声と同じように捉え、音より先に見たり後に戻ったりをしてはいけない。
2. テキスト化では語彙に注目するだけではなく、語彙レベルを超えて感情や意図などに関わる面、ストレス、イントネーション、ポーズなどを考慮する。
3. 句読点の意味に注目させる。発話の完結を表す下降調と未完の低いままの上昇調、質問の上昇調などは、句読点、クエスチョン・マーク、立ち消えの「・・・」などの記号と声の表情をあわせることで感覚が育てられる。

富田らは、母語話者は聴解を自動的に処理しているために学習者の難しい点に気づけない。だからこそ、努力して処理するという経験を持つ、目標言語の学習経験者の聴解指導がもっと認知されるべきであると述べている。

留学生にとって、上記の3つの練習方法は〈日本人のような日本語聴解〉を目指さなくとも、有効だと考える。本研究では、**think aloud**で対象者自身が聞き取った内容を話しながら聴解行動を振り返るという方法を使用したが、その際、日本人大学生のテキスト再生率が1番高かった。テストとして考えた場合、正答を得るための情報さえ聞き取れていれば問題はないのだが、聴解の練習では、「テキスト化できる」能力を伸ばすために、ディクテーションや口頭要約などが有効だと考えられる。ディクテーションは、捉えた音のテキスト化の正確さを見ることができ、口頭要約は意味化の正確さを見ることができるからである。日留試のようなテストを利用する場合は、選択肢を聞かず正答を口頭あるいは文字で書かせる方法が、同様の効果をあげるのではないかと思われる。

竹蓋（1984, p.88）は、音の情報以外、言語活動の行われている環境、話題の内容、話者の態度や非言語、発想パターンなどの情報を上手く使えば、文字や音声がなくてもその不確定程度の約75%が自動的に埋められると述べている。1963年に、アメリカの大学で普通の母語話者と、聴覚機能に若干の問題のある（**Mild hearing loss**）母語話者、それに同大学大学院に留学中の外国人学習者、各32名の聴解能力を比較する研究を行った。その結果、聴覚に問題のある母語話者は普通の母語話者より5%低く、外国人学習者はさらに約15%劣ることが判明した（同掲 pp.73-74）。例えば、状況説明のみを聞かせたあと日本語で話している話者を音声無しで見せ、感情の動きがあるか、どのような内容を話しているのかなどの推測をするのも練習になるだろう。また、学習者同士どのように予測したか、なぜそのように予測したかなどについて話し合うことで、話し手の意図の理解を様々な面から考えることができることに気づく。

Buck（2001）は、背景ノイズ、注意力散漫なども聞き間違いのもととなる、聞き手の背景知識によって、同じセンテンスでも違う意味に取る場合もある、聞き手が何に興味を持ち、何が必要かによっても、

得する手段として、テスト・テクニックを用いることはできるだけ避けるべきである。このようにテストの波及効果に憂う反面、言語教育の方向を変え、教師に自分の教授法だけでなく教えている言語について調べてみようという気を起こさせる1つの有益な力となっているとも述べ、波及効果の悪い面と良い面を挙げている。

違う解釈をすると述べている。同じ談話でも個々によって違うことを知り、その違いはどこに由来しているのかを知ることによって聞く際の態度、注意の向ける箇所、既存知識の使い方、談話内のヒントなど、聴解時の行動選択肢が豊富になる (Buck, 2001, pp.6-8)。Buck (2001, pp.54-55) は、コミュニケーション・アプローチの聴解スキルとして以下の7つを挙げている。

- ①語彙理解、わからない語彙は推測する
- ②統語パターン、形態の型、談話パターンを理解する
- ③ストレス・非ストレス、イントネーション、句読点を理解する
- ④話者の目的を知る
- ⑤社会的状況、話者の意図、文脈から正しいまとめや推測をする
- ⑥話者の聴者に対する態度と議論の話題を認識する
- ⑦話者が使用しているテクニックとレトリックを知る

Buck (2001) が挙げた、聞き手による聞き方の違いを認知すること、何をすることが聴解理解を助けるのかを意識することを、聴解練習に組み込めば、能力は向上する。聴解テストや聴解練習の談話のスク립トを利用し、統語パターン、談話パターン、句読点の理解などを聴解練習の前に行うなど、今後聴解スク립トと聴解の関係の研究を進めていくことも聴解能力向上に成果をあげるのではないだろうか。

坂本 (1994) は、大学生活の聴解力として講義形式に慣れること、ノートの取り方、聞いた内容をまとめることが必要だと考え、聴解練習として大意聞き取りと詳細聞き取りの2種類を挙げている。大意聞き取りでは、上級の留学生は既にスキミングはできるため、スキミングの内容確認の必要性を強調している。詳細聞き取りでは、口頭で繰り返させ、正確に文を組み立てられるかどうかの練習もする。聴解練習は、聴解力だけではなく表現力をのばす練習としても適当であると述べている (坂本 1994, pp.472-477)。「慣れ」に関しては前述しているので省くが、聴解練習が表現力を伸ばす、つまり発話や作文などの記述能力に影響を及ぼすことについて考えてみる。日留試では聴読解という文字情報を見ながらの聴解分野がある。新日能試の聴解テストでは、「即時応答」や「発話表現」で発話に繋がる能力について問うことを始めた。「即時応答」と「発話表現」で会話能力を測定できるとはしていないが、可能性の方向を示していると考えられる。聴解は聞いて理解した事柄を土台に、発話をしたり、記述をしたりする情報源である。聴解スキルのみでの測定にとどまらず、読解と組み合わせた複合スキル、聴解後に発話する・記述するなどの発信力も含めた統合的なテストも将来に期待したい。会話テストの実施は人員、評価など難しい点が多いが、パソコンやインターネットなどの発達と普及に伴い、直接対面ではない会話テストの試みがヨーロッパでは始まっており、会話テストが簡単に実施される日も近い。

最後に、「わからない言葉」の処理について述べる。水田 (1996) は、日本語話者と中国人日本語学習者の独和聞き取りに関する研究で、両者に見られる共通問題点として ①文節化 ②辞書の意味 ③テキスト上の意味 ④テキストからの派生問題 ⑤話し手の意図 ⑥その他の場面要素 の6つを報告している。これらは日本語話者と中国人日本語学習者ともに共通してみられた問題点ではあるが、日本語話者は文脈理解に問題が生じることが多く、学習者は日本語の知識不足から生じることがわかった。問題処理戦略として両者に共通に見られたのは ①問題特定だけで終える処理法 ②推測する処理法 ③問題を解決し確認、精緻化へと連鎖させること であった。問題処理戦略では、日本語話者は「自己モニター」や「聞き流し」への連鎖が見られたが学習者には見られなかったため、これらを使いにくい戦略であるとしている。本研究では、留学生にも「回避」行動 (聞き流し) が見られたため、水田の研究結果とは異なっているが、「回避」行動に心理的な負の感情があることは、既に記述した。富田らや Buck が指摘しているように、テキストとして認知できなかった場合、留学生は処理が遅れる、その遅れが心理的に負の方向に働くと次を聞く準備ができなくなる、あるいは安易に聞き取れた語彙のみから推測しようとしてしまった。聴解で1番大切なことは、テキスト化できなかった語彙などは空白のままにし、次に聞こえてくる情報からテキスト化できなかった部分が推測できる可能性が

あること、空白の部分があっても意味化に支障が出ないことも多々あることを、聴解練習や日常生活での聴解経験を重ねることである。日本語学習者は、初級の段階から既習事項のみで作成された談話を聞くのではなく「わからない言葉があるのは当然」という談話を繰り返し聞くことで、その量に違いがあれどネイティブが日常的に経験している環境に近づけられる。常に正確に聞き取ることを教師も学生も期待しがちだが、聞く目的によって不備がある聞き取りでも最終的に意味がとれればよい、聞き間違えてもよいから聞き方の様々な方法を探るなどの目的にそった聴解練習をすることも必要であろう。

8章 結論

本研究の目的は、日留試の聴解テストに注目し、以下のことを明らかにすることであった。

- 1 日本留学試験聴解テスト・アイテムの特徴を明らかにする。
- 2 日本留学試験聴解テストが留学生に期待している聴解能力を、アカデミック・リスニングの視点から明らかにする。
- 3 聴解テスト時、受験者はどのような聴解行動をとっているのかを明らかにする。

3つの目的について、結論をまとめる。

8.1 日留試聴解テスト・アイテムの特徴

日留試は、2002年開始から2009年までと、2010年の改定後で聴解テスト・アイテム数が変わったこともあり、2002年から2009年を初期、2010年以降を改定後として、両者に大きな違いが見られる場合は、分けてまとめる。テスト・アイテムの特徴を ①理解の種類 ②場面 ③形態 ④トピック ⑤人数 ⑥コミュニケーションの目的 ⑦機能 ⑧情報を得る理由 の8つに分類して分析した結果、以下のことがわかった。(詳細は4章を参照)

「①理解の種類」では、本文の内容の部分理解が7割以上であった。全体理解は8%と低く、特にアカデミック場面の講義内容のアイテムに集中していた。アカデミック場面は、2002年1回目は出題アイテム全体の10%であったが、毎回増加傾向にあり2010年改定後は常に60~75%の出題に落ち着いている。アカデミック場面の増加は、講義のアイテムが増加したことであり、留学生にとって講義理解能力の必要性が高いことがわかった。

「②場面」は、2002年から2013年までの24回の合計では、アカデミック場面とキャンパス場面とライフ場面でほぼ均等に出题されている。日本の大学内・大学外の生活のあらゆる場面で、日本語能力を測定しようとしているからだと考えられる。しかし、毎回アカデミック場面が増加していることから、日常生活より大学の授業に重点がおかれつつあることがうかがえる。一方、ライフ場面は2002年1回目が全体の55%だったが、2010年改定後は全体の20%前後と減少している。これはアカデミック場面の重視傾向と対照的である。キャンパス場面も、アカデミック場面の増加に反し、2002年1回目の35%から2010年改定後は10%前後に減少してきている。しかし、キャンパス場面は日本の大学システムについての情報提示という機能があり、今後もある一定数は出题されると予想する。

「③形態」は「授業」40%、「私的会話」26%で全体の66%を占め、上記の場面と連動している。アカデミック場面が増加傾向であるため、今後は「授業」の比率が増加すると考えられる。

「④トピック」は、本研究ではテスト・アイテムのタイトルと位置づけたため分類はしなかった。今後講義内容や講演などを文系・理系あるいは専門分野別にわけること、専門用語の分類や必要語彙の整理ができるだろう。

「⑤人数」は、モノログ51%、ダイアログ48%で、ほぼ同じ割合であった。アカデミック場面での「講義」の増加に伴い、今後独話が増えていくと考えられる。3人以上の談話は1%で、これは音声の違いのみで話者を特定することの難しさからであると思うが、留学生にとって複数参加の談話理解、複数参加の談話に参入する必要性は高いと思われ、今後何らかの工夫を講じて出题されることを望む。

「⑥コミュニケーションの目的」については、情報提供に関わる「知識、問題解決、理解」で81%を占めていた。コミュニケーションによる影響は、「行動変容」が77%であった。本文の談話に参加する2人は、一方が話し手、他方が聞き手の役割を持っていることが多かった。このような談話で、聞き手の

行動を変えさせるという具体的な目的をタスクにすることが容易だったのではないかと考えられる。今後、アカデミック場面のゼミ、共同研究、共同作業などの場面において、お互いに意見交換をし、議論し、相手の意見を変容させる目的の談話の増加を期待したい。

「⑦機能」は、情報の授受に関して「講義、教授、説明」など一方向から情報を与えるものが多かった。2人の会話でも、「質疑応答」は情報の多い話し手が情報の少ない聞き手に情報を与えるという点で、「講義、教授、説明」と機能的には同等である。アカデミック場面においては、情報が一方向から流れるだけではなく、対等に「議論する」ような情報のやりとりが必要だと考えられる。2人の会話では、話し手が「提案」し、聞き手が「承諾」するというパターンが50%程度を占めた。日留試は四肢選択であるため、談話内容が提案に対して肯定1回、否定3回が含まれることが容易に予測できる。肯定と否定は瞬時に判別できるため、「提案と承諾」の内容では困難度を高くすることは難しい。「承諾するか否か」ではなく、「承諾の条件」「承諾の理由」「承諾のタイミング」などの、部分理解を問うことも必要だと考えられる。

「⑧情報を得る理由」は、「学術・研究」に関する情報が41%、「学習」に関する情報が14%で、この2つで全体の50%以上を占めた。情報を得る理由もテスト・アイテムの場面と連動している。

以上の結果から、テスト・アイテムの「場面」が、他の7つすべてに影響を与えていることがわかった。テスト・アイテムの「場面」設定のための現状調査と分析が、テスト・アイテム作成に重要であると言える。日留試は、大規模日本語テストで初めてアイテム・バンクを導入し、アイテム・ライターによってアイテムを作成している。多様な視点で作成したほうが、多様なアイテムができるという利点から、多くのアイテム・ライターが関わっていると思われる。しかし、今回分析をしたアイテムはテキストの展開、タスクの内容など類似のものが多かった(4章参照)。多くのアイテム・ライターが、同じような視点で、同じようなアイテムを作成した結果であろう。そして、一律なテスト・アイテム形式の弊害もある。これを避けるためには、アイテム・ライターに対し、テキスト展開やタスクの種類などの多様かつ多量の提示(サンプル)が必要である。日本語の自然談話を切り取って分析すること(job analysis)のほかに、言語テストのテキストやタスク研究が有効である。本研究では集められたテキストやタスクを、テキスト・バンクとして整理・管理されていれば、それを利用して将来アイテム作成ができることを提案した。

8.2 日留試聴解テストのアカデミック・リスニング能力

聴解テスト・アイテムのスクリプト分析より、以下のことが明らかになった。(詳細は5章を参照)

文法的知識では「講義に特徴的なディスコース・マーカー」の認知のみであった。「ディスコース・マーカー」は、談話構成の理解に必要であり、語彙や表現レベルの日本語知識を測定しているのではないことがうかがえる。テキスト知識として、「トピック展開」「結束性のマーカー」「テキスト内の関係」「構成要素の関係」の理解を求めるものが多く、テキスト全体の流れと各部分をどう関係づければ意味解釈ができるかという能力を測定しようとしている。機能的知識では、「コミュニケーション機能」の理解が必要であることから、タスク理解と談話の目的、場面による参加者の役割の把握を問うていると考えられる。社会言語学的知識は、「慣用的な言い回し」の理解のみが問われており、「言語使用の違い、関連事項、特定の場の約束事」は問われていなかった。「慣用的な言い回し」を表現の一種として捉え、文法的知識に含める考え方もあるが、社会文化的背景知識がなければ適切な使用はできない。言語使用が適切か否かをテストで問う場合、正答誤答が明確でなくなることもあり、テストでは出題しにくい可能性はある。聴解テストで問うことが難しいのであれば、適切な使用をしている談話を聞かせることが学習効

果をもたらすので有益である。日留試で、日本の大学の授業展開や学生たちの役割・行動などを扱ったテスト・アイテムを出題することで、留学希望者に日本の大学の情報を提示できる。日本の大学の授業の特徴には、大学のシステム、大学教育の歴史、習慣など社会文化的な要素が入っている。留学生にとっては大切な情報であり、日留試のテスト・アイテムに取り入れられることを期待したい。方略的能力については、スクリプト分析という方法からは捉えることが難しかった。

アカデミック・リスニングには不可欠な「講義の目的」「キーワード」がわかる能力は、テスト・アイテムの長さが短く、ノートとりのタスクもないことから、日留試のアイテムには見られなかった。日留試の聴解テスト・アイテムの形式の制約が、測定能力を限定していると言える。

以上の分析結果と、「8.1日留試聴解テスト・アイテムの特徴」(p.185)から、「日留試のアカデミック・リスニング」とは、留学生が日本で生活する際、特に大学内で遭遇する場で行なわれる談話の理解能力であると結論づける。聴解テスト・アイテムでは、プリタスク(課題)の理解と、本文の内容理解から課題達成の能力を測定していると考えられ、個別的に言語知識を測定しているのではない。しかし、選択肢では語彙や統語的な言い換えが5割近くされており、文法的知識が要求されている。文法的知識を問うことに終始してはならないが、文法的知識の必要性を排除することもない。ただ、本文内容理解で問うている知識が、テキスト知識であっても、選択肢で文法的知識が問われるのであれば、最終的に文法的知識を測定していることになる。選択肢では、本文理解能力を阻害するような言い換えを避けるように注意していく必要がある。

8.3 留学生¹の聴解テスト時における聴解行動

受験者は聴解テストが始まった時点、設問を聞いたときに、その後の聴解行動を決定する。1つはテストであることを重視し、タスクの解答に集中する選択的な聞き方で、もう1つは本文理解を重視して全体的に聞くという聞き方である。設問も含め、本文中にわからない部分があった場合は、その部分を記憶に保持しながらテキスト処理を行わなければならないが、心理的には不安を抱えるため、テキスト処理に負の影響が生じる。不安が大きすぎると、聴解行動ができなくなる場合もある。選択肢を聞く前に正答を言語化する場合と、選択肢を聞くまでは明確に言語化しない場合があるが、聴解テストであるため選択肢から正答を選ぶことを優先する傾向が見られ、正答を言語化している場合でも選択肢との照合時、言い換えなどに注意を払っている。選択肢に納得できるものがない場合は、解答放棄をする行動が見られた。聴解テストは一般の聴解とは違い、プリタスクが与えられ、選択肢から選ばなければならないという非現実的な点が浮き彫りとなり、それがテスト聴解であると留学生は認識している。

設問がない聴解テスト・アイテムの場合、本文を聞きながらタスクを推測しつつ緩やかに情報を絞る場合と、推測せず全体的に聞く場合の2つの聞き方であった。本文中にわからない部分があった場合は、設問があるアイテムと同様の行動が見られた。本文終了後にタスクがわかるため、タスクに対する情報を的確に得ていたかどうかで、選択肢時の処理が変わる。正答となる情報を得ていた場合は、設問があるアイテムと同様の行動になるが、情報量に過不足があったり曖昧な場合は、個々の選択肢と持っている情報を吟味し解答を決めた。設問がないアイテムの場合、課題推測と本文の情報の必要度から情報収集の量を決める、課題を推測せず満遍なく情報を保持するという行動になり、記憶への負担は増えた。その記憶負担を軽減するために、メモとりのスキルが設問のあるアイテムより重要性が高くなる。クエ

¹ 本研究では留学生と日本人大学生を対象として分析・考察を行ったが、日本人大学生は留学生に対する比較対象であり、日留試の受験者は留学生であるのでこの節では、留学生の聴解行動の結論を記述する。詳細は、6章と7章を参照。

スチョンで初めてタスクを知り、すぐに選択肢に移るため、情報処理スピードが、設問がある場合に比べ求められていた。設問がないアイテムの聴解のほうが、スキルや能力がより多く求められている。

設問の有無により聴解行動にかかる負荷の違いがあることがわかったが、実際の聴解の場合、談話を聞く前に聞く目的が明確な場合と、あまり明確ではない場合がある。例えば、ある人がテストの説明をする場面の聴解では、日時、場所、テスト範囲、持ちもの、禁止事項、注意事項などは事前に予測でき、明確な目的を持って聴解ができる。一方、講義を聞く場合、講義内容を理解するための知識情報、テキスト（教科書など）に記述がないためノートに書く必要がある情報、テストで出題される可能性がある情報、興味を持てる情報などが、どのようなタイミングで、どのような提示の仕方ですれるか予測しにくい。前者のような場合は、プリタスクがあり、目的の1つを解答させるという形式が妥当である。後者の場合は、プリタスクは必要なく、状況判断と課題設定を受験者自らが発見できる形式が妥当と考えられる。

さらに、聞く目的の重要度や、興味の強さも聴解行動に影響を及ぼす。聴解後のインタビューで、「興味を持てた内容のテスト・アイテムは集中して聞いている」という意見があった。一方、「テストだから、内容に関係なく聞いている（重要か興味があるかは無関係）」という意見もあった。両者のテスト聴解に対する態度に注目したい。例えば、テストの説明は情報を正確に聞き取ることが、その後の自分の学業や行動に影響を与えるが、友人がどの選択科目を選ぶかは人間関係維持には必要だが、その後の自分の生活や行動に影響をほとんど及ぼさない。友人の選択科目については、談話を聞く前のプリタスクも必要なく、タスクを絞り込んでいく必要もない。前者のような受験者には、友人の選択科目を聞き取ることに関心を持つことが求められ、後者のような受験者には聴解を強制していることになる。テストでは、既に聞く目的をプリタスクにより強制的に与えているが、実際の聴解ではタスクを自ら決めることもでき、また聞く行動自体を回避することもできるのである。テストの持つ強制的性に頼るテスト・アイテム出題は避けるべきである。すべての受験者が、聞く必要性を認め、興味を持つテスト・アイテムを作成することは難しいだろうが、多くの受験者が納得できる内容のテスト・アイテムを作成する義務が、テスト作題者に求められる。

日本語教育における聴解練習は、大規模聴解テスト形式の影響を強く受けているといつてよい。テスト形式と同様の練習問題は、問題を何度も解くことで「慣れ」につながり、テスト時に正の効果をあげることが期待できる。「慣れ」以上に、効果が期待できる練習は、タスクの予測、正解するまでの過程、不安材料の原因と処理、誤答を選んだ理由、選択肢に対する意見などについて、学習者同士が聴解過程を語り合うことである。本研究の対象者となった留学生の言語知識、経験、世界知識には差があり、聞き方も多様であった。唯一正しい聴解方法は存在しない。教師がよいと考える聴解方法は、ある学習者にとって必ずしも最適な聴解方法とは言えない。学習者同士が話し合うことで、学習者自身は聴解モニタリングを体験でき、他者の様々な聴解方法を学ぶことができ、その中から取捨選択をすることもできる。自律学習という学習者個人が独立して学習することも大切だが、学習者同士の協同学習（ピア・ラーニング）²も大切である。

8.4 今後の課題

本研究を始めるにあたり、

- (1) 日本留学試験テスト・アイテムの基礎的研究が少ない

² 詳しくは池田・館岡（2007）、館岡（2007）を参照。

- (2) アイテム・ライター制度の認知度が低く、アイテム・バンクの研究が少ない
- (3) 日本留学試験はアカデミック・ジャパニーズに特化したテストであるが、アカデミック・ジャパニーズの定義が曖昧である
- (4) 日本留学試験における聴解領域の研究が少ない

の4点の問題点をかかげた。

「(1) 日本留学試験テスト・アイテムの基礎的研究が少ない」に関してであるが、日留試は毎回実施された中の1フォームが公開され、テスト・アイテムの研究をすることができる。本研究では、そのテスト・アイテムを使用し基礎的データの分析ができた。この分析は今後も続けていきたい。しかし、テスト・アイテムだけの分析では推測の域を出ないこともあり、本研究での限界を感じた。テストの機密性は理解できるが、日本語テストの将来のためにもテスト・アイテムの結果のデータなど、最大限の情報開示を行ない、多くの研究者が研究できる環境が作られることを望む。研究されることで、テストにフィードバックがされ、テスト改善への方向性が示される。多くのテスト利用者に、日留試の理解を促し、広く知ってもらうことにも役立つ。更に、日本語教育にもフィードバックがされる。日留試の得点だけを利用するのではなく、日留試のテスト・アイテムは、日本の大学について知識を深める、日本語の社会文化的な側面を学ぶ、留学生の日本語、アカデミック・ジャパニーズは何かを考える、などにも利用できる。アカデミック・ジャパニーズは、日本語教育だけの問題でなく、日本人学生も含め、大学に関わる人たち全てが、検討していくべき問題である。日本語教育関係者以外が関わることで、「(3) 日本留学試験はアカデミック・ジャパニーズに特化したテストであるが、アカデミック・ジャパニーズの定義が曖昧である」ことが解決されていく。日留試の実施団体、日留試利用者（学習者、教師、大学）、研究者の協力があって日留試は育てられる。多くの人たちを巻き込んで、発展していくことに意義がある。

「(2) アイテム・ライター制度の認知度が低く、アイテム・バンクの研究が少ない」ことについて、考える。日留試が始めたアイテム・ライター制度は、日本語教師にとって大規模テスト作成に関わることができるよい機会を与えてくれた。日本語教師をめざしている人々にとっても、教壇に立って教える以外の仕事の可能性を広げている。大規模テスト・アイテム作成に関わった経験が、自らのクラスのテスト作成や日本語教育現場で使用するプレースメント・テストやアチーブメント・テストなどの作成に役立つ。テスト評価の重要性と難しさ、そしてその必要性を日本語教師が認知するためにも、アイテム・ライターの経験は有効である。アイテム・バンクの研究が少ないことは、すなわちアイテム・バンクがほとんど使用されていないことを示す。アイテム・バンクは、クラスで行なう小クイズからも作ることができる。テスト・アイテムをラベル化し、結果を付加しておくだけで、再度アイテムを利用することができる。同じテスト・アイテムを複数回使用することで、結果の数値が蓄積され、それが精度の高いテストを作る材料となる。本研究では、聴解テスト・アイテムのスキ립トのテキスト・バンクの提案をしたが、結果の数値を参考に少しずつテキストに変更を加え、より精度の高いテストを作成することも可能である。小規模テストであっても精度の高いテストが望ましく、アイテム・バンクは精度を高めるために役立つ。アイテム・バンクの実用性と利便性、そして簡便であることを広めるためにも、日留試のアイテム・ライターに日本語教育関係者だけでなく、多くの教育関係者が携わり、アイテム・バンクについての研究がされることを望む。

「1.1 本研究の背景」(p.3) で、聴解テスト・アイテム作成は、「作り手（アイテム・ライター、テスト実施機関）の視点と受け手（受験者）の視点が異なっていることである」とし、視点の違いに注目した。今回、作り手は、聴解スキ립トに文法的知識、テキスト知識、機能的知識を多く配していた。機能的知識では、日本語の機能的表現、機能語が重視されていた。聞き手（6章、7章の対象者）は、言

語形式や談話構造に注目して聞いており、アイテム・ライターが測定対象として意図した能力を使用していたと言える。しかし、談話構造の不自然さ、通常の会話音より明確な音声音などの指摘があり、テスト形式の議論が望まれる。選択肢で語彙や統語の言い換えが約半数行われていることは、「音」の聞き取りだけで正答を選べないよさという考えがあると思われ、設問から選択肢まで音を重視していることがわかる。本研究では選択肢に対しては対象者へのインタビューを行っておらず、今後本文との関係も鑑み、選択肢に関してさらに研究を進めなければならない。書き手と聞き手の視点の違いについては、聴解時の過程を詳細にすることで明確にできるのではないかと考えており、今後研究を深めていきたい。

「(3) 日本留学試験はアカデミック・ジャパニーズに特化したテストであるが、アカデミック・ジャパニーズの定義が曖昧である」についてだが、日留試で使用し始めた「アカデミック・ジャパニーズ」は、残念ながら2010年の改定でその表現はシラバスから削除されてしまった。だが、日本語教育では「アカデミック・ジャパニーズ」という用語は定着した。本研究では、「アカデミック・リスニング」を大学の講義・授業に関わる聴解と定義したが、それは限られた解釈でもある。今後、多くの研究者がそれぞれの定義付けを行い、「アカデミック・ジャパニーズ」の研究がされていくだろう。日本語教育における「アカデミック」の解釈は、始まってからまだ十数年である。今後、日本の高等教育に関わる講師と日本語教師の協力により、留学生のためだけの「アカデミック・ジャパニーズ」ではなく、日本人大学生にとっての「アカデミック・ジャパニーズ」として徐々に共通解釈が形成されると思われる。

最後に、「(4) 日本留学試験における聴解領域の研究が少ない」ことについて述べる。聴解に関する研究の難しさ、特に質的研究において本研究で感じられたのは、聴解をしながら対象者はその瞬間を口頭で即座に表現できないことである。研究者は聴解行動が終了した後、対象者の記憶を刺激することによって語ってもらうしかない。対象者は、聴解行動時には談話を聞きながら、その意味理解を瞬時に行なう。意味理解ができなかった場合は、不安と戦いながら聴解行動を続けなければならない。そのような聴解終了後、対象者が自らの聴解行動を語ることは、記憶の底から表出した内容に限られる。対象者は記憶の深層を語ることは難しく、研究者は対象者の記憶の深層を覗くこともできず、ブラックボックスの状態である。聴解研究は、対象者が語る情報を集め、多くのケースを累積していくことで徐々に解明していくほかないのである。今回、質的研究に拘り、一人の発言、一回の発言も丁寧に拾った。言語行動は、個々人のものであり、個々の場面、個々の状況、個々の判断、個々の決定から生じる個々の行動である。聴解のブラックボックスを開くヒントが、どこに隠されているか、まだ混沌とした状況である今日、質的研究のデータの重要性は高い。

日本語教育において、旧日能試が与えた影響は大きい。旧日能試の結果が留学基準として使用され、また職場へも影響を与えた³。旧日能試のテスト形式は日本語教育の場に持ち込まれ、テスト対策という名目で授業が行なわれている。テスト結果がその後の人生に多大な影響を与える場合、学習者はいかにテスト得点をあげるかを目標にするだろう。日本語学校では、日能試の合格率が学校の評判に繋がることもある。しかし、テストで測定する能力には偏りがあり、限界がある。例えば、大学生にとって課題発見能力が必要であると多くの研究者が論じているにも関わらず、現行の日留試の聴解テストは、課題発見能力を測定するという形式にはなっていない。講義理解には文字情報があり、ノート・テイキングのスキルも必要だが、現行の日留試の聴解テストでは測定していない。大規模テストは、実施上の時間の妥当性、採点の信頼性などを確保しなければならず、多肢選択、ブリタスクの必要性など、実際の生活で行なわれる聴解活動とは違う状況にある。我々言語教育に関わる者は、テストの「得点」をあげる

³ 新日能試に改定されてからも、その得点結果は、入社時の条件や職場での地位に影響を与えている。

ことを目標に教育をしてはならず、同時に学習者がテスト結果だけに拘るような環境を作ってはならない。テストは、評価の1つの道具である。それを、どのように学習に利用し、学習効果をあげるのに有益であるかを念頭に指導していく必要がある。

第1章で述べたが (p.8 を参照)、日留試は日本の高等教育への入り口である。留学生の将来の日本語能力を測定することで、日留試の結果を利用するすべての人々に、大学生活を送るための十分な日本語能力が受験者に備わっていることを証明する。そのためには、聴解・聴読解、読解、記述の3分野で、どのような日本語能力を測定しているかを明示することが求められる。点数だけではなく、その点数が具体的にどのような能力を示しているのかを、万人にわかってもらう表現も有益である。Can-do statements である CEFR (Common European Framework of Reference for Languages) や JF 日本語教育スタンダードの利用、独自の点数に対する説明なども視野に入れ、テスト・ユーザーに理解してもらえよう実施側の努力を期待したい。同時に、我々研究者もテスト結果の点数と実際の留学生の行動や理解力の関係を研究していかなければならない。

参考資料 * () は論文本文の関連ページ

1 日留試日本語シラバス 各領域の概要 2009年12月25日変更 (本文1章 p.6)

(日本学生支援機構 HP より転載 http://www.jasso.go.jp/eju/syllabus_01.html)

1. 読解、聴解、聴読解領域

読解は、主として文章によって出題されるが、文章以外の視覚情報(図表や箇条書きなど)が提示されることもある。聴解は、すべて音声によって出題され、聴読解は、音声と視覚情報(図表や文字情報)によって出題される。

(1) 問われる能力

読解、聴解、聴読解領域では、文章や談話音声などによる情報を理解し、それらの情報の関係を把握し、また理解した情報を活用して論理的に妥当な解釈を導く能力が問われる。具体的には以下のような能力が問われる。

ア) 直接的理解能力:

言語として明確に表現されていることを、そのまま理解することができるかを問う。たとえば、次のようなことが問われる。

- ・個々の文・発話内で表現されている内容を、正確に理解することができるか
- ・文章・談話全体の主題・主旨を、的確にとらえることができるか

イ) 関係理解能力:

文章や談話で表現されている情報の関係を理解することができるかを問う。たとえば、次のようなことが問われる。

- ・文章・談話に含まれる情報のなかで、重要な部分、そうでない部分を見分けることができるか
- ・文章・談話に含まれる情報がどういう関係にあるかを理解することができるか
- ・異なる形式・媒体(音声、文字、図表など)で表現されている情報を比較・対照することができるか

ウ) 情報活用能力:

理解した情報を活用して論理的に妥当な解釈が導けるかを問う。たとえば、次のようなことが問われる。

- ・文章・談話の内容を踏まえ、その結果や帰結などを導き出すことができるか
- ・文章・談話で提示された具体的事例を一般化することができるか
- ・文章・談話で提示された一般論を具体的事例に当てはめることができるか
- ・異なる形式・媒体(音声、文字、図表など)で表現された情報同士を相補的に組み合わせて妥当な解釈が導けるか

(2) 出題される文章や談話の種類

(1)で挙げられた能力は、大学等での勉学・生活の場において理解が必要となる文章や談話を題材として問われる。具体的には以下のような文章・談話である。

読解

- ・説明文
- ・論説文

- ・(大学等での勉学・生活にかかわる) 実務的・実用的な文書／文章 など

聴解、聴読解

- ・講義、講演
- ・演習や調査活動に関わる発表、質疑応答および意見交換
- ・学習上または生活上の相談ならびに指導、助言
- ・実務的・実用的な談話 など

2. 記述領域

(1) 問われる能力

記述領域では、「与えられた課題の指示に従い、自分自身の考えを、根拠を挙げて筋道立てて書く」ための能力が問われる。具体的には以下のようなことが問われる。

- ・与えられた課題の内容を正確に理解し、その内容にのっとった主張・結論を提示することができるか
 - ・主張・結論を支えるための、適切かつ効果的な根拠や実例等を提示することができるか
 - ・主張・結論を導き出すに当たって、一つの視点からだけでなく、多角的な視点から考察をおこなうことができるか
 - ・主張・結論とそれを支える根拠や実例等を、適切かつ効果的に、また全体としてバランスのとれた構成をなすように配列することができるか
 - ・高等教育の場において、文章として論述をおこなう際にふさわしい構文・語彙・表現等を、適切かつ効果的に使用できるか

(2) 出題される課題

- ・提示された一つまたは複数の考え方について、自分の意見を論じる
- ・ある問題について現状を説明し、将来の予想や解決方法について論じる 等

2 個別的要素テストの評価項目 (本文2章p.11)

Carroll (1961)は ①聴解力 ②発話力 ③読解 ④作文 の4つのスキル (skill) と、①音韻論 ②形態論 ③統語論 ④語彙論 の4つの相 (aspect) の組み合わせで 16 項目がテストで評価することが理想だが、事実上不可能であるため以下の 10 項目をあげている。

1. 構造知識 (形態・統語)
2. 語彙の一般的使用 (単語・慣用句)
3. 聞き分け (音素、異音、超文節音素：いくつかの文節素にまたがって表れる音声的特徴で音の高低、強勢、接続)
4. 発話産出 (音素、異音、超文節音素：いくつかの文節素にまたがって表れる音声的特徴で音の高低、強勢、接続)
5. 読解 (文中の文字を音に変換する。つまり、語彙、発音、ストレスパターンの習得)
6. 作文 (文中の音を文字に変換する、つまり、綴り)
7. 聴解率と正確さ
8. インタビューでの発話の量と質
9. 読解率と正確さ
10. 作文の量と正確さ

Carroll は7.-10.が重要であり、個々の語彙や構造などより発話全体としての力を測定することが大切だと述べている。(Harris, 1969 邦訳 pp.157-158)

3 z 得点、T 得点、CEEB 得点の算出方法 (本文2章p.16)

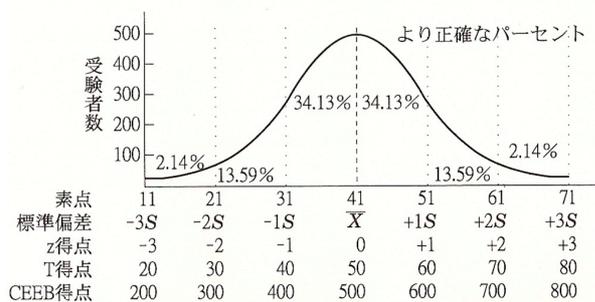


図1 標準得点の正規分布

図1は2章の「2.3.1 集団基準準拠テスト」に記載した図2.2 (p.14) と同様のものである。ここに記されている z 得点、T 得点、CEEB 得点の算出方法について説明する。

z 得点は素点が平均値からどのぐらい離れているかを標準偏差を利用して換算したものである。z 得点は以下の式によって導き出される。図1のテストの場合、素点 51 点の受験者は z 得点は +1 となる。

$$z = \frac{X - \bar{X}}{S} \qquad z = \frac{51 - 41}{10} = 1$$

z 得点は平均値から標準偏差だけ上回るのか下回るのかを表しており、-3 から +3 の間の得点となる¹。z 得点を使用して報告する場合は少数第2位までの得点が出されるため、受験者には -0.84 点、+2.00 点などが報告される。しかし z 得点は受験者やテスト結果使用者などにはわかりづらいため、わかりやすくしたのが T 得点である。

¹ テスト結果が正規分布を示してもその範囲を超えた得点をとった場合、例えば、図1のテストで受験者1名だけが81点をとった受験者のz得点は+4となる。しかし、この1名は他の受験者集団とあまりにかけ離れているため、この集団に属さないと考え、はぐれ値として計算から外す。(Brown 1996 邦訳 pp.123-124)

T 得点は以下のように z 得点を 10 倍して 50 を加える。よって、z 得点が -0.84 の場合 T 得点は 58.4、z 得点が +2.00 の場合 T 得点は 70 となる。

$$T=10(-0.84)+50=58.4$$

$$T=10z+50$$

$$T=10(2.00)+50=70$$

CEEB 得点は z 得点を 100 倍し、500 を加えるので T 得点の 10 倍の結果となり小数点はなくなる。T 得点は z 得点から換算しているため 20 から 80 の結果となる。受験者が素点最高点の 71 点の場合、T 得点は 80 となるが、80 点という得点を見たときに受験者は自分が受験者群のなかで最高得点をとったということが実感しにくい。100 点満点中何点だったかというテスト結果に慣れている受験者にとってわかりやすいのがパーセンタイル得点である (表 1) (Brown, 1996 邦訳 p.156)。この換算表はテスト結果が正確な正規分布となっていることが前提である。

表 1 標準化された得点のパーセンタイル得点への換算

z	T	CEEB	パーセンタイル	z	T	CEEB	パーセンタイル
3.0	80	800	99.9	-0.1	49	490	46.0
2.9	79	790	99.8	-0.2	48	480	42.1
2.8	78	780	99.7	-0.3	47	470	38.2
2.7	77	770	99.6	-0.4	46	460	34.5
2.6	76	760	99.5	-0.5	45	450	30.9
2.5	75	750	99.4	-0.6	44	440	27.4
2.4	74	740	99.2	-0.7	43	430	24.2
2.3	73	730	98.9	-0.8	42	420	21.2
2.2	72	720	98.6	-0.9	41	410	18.4
2.1	71	710	98.2	-1.0	40	400	15.9
2.0	70	700	97.7	-1.1	39	390	13.6
1.9	69	690	97.1	-1.2	38	380	11.5
1.8	68	680	96.4	-1.3	37	370	9.7
1.7	67	670	95.5	-1.4	36	360	8.2
1.6	66	660	94.5	-1.5	35	350	6.7
1.5	65	650	93.3	-1.6	34	340	5.5
1.4	64	640	91.9	-1.7	33	330	4.5
1.3	63	630	90.3	-1.8	32	320	3.6
1.2	62	620	88.5	-1.9	31	310	2.9
1.1	61	610	86.4	-2.0	30	300	2.3
1.0	60	600	84.1	-2.1	29	290	1.8
0.9	59	590	81.6	-2.2	28	280	1.4
0.8	58	580	78.8	-2.3	27	270	1.1
0.7	57	570	75.8	-2.4	26	260	0.8
0.6	56	560	72.6	-2.5	25	250	0.6
0.5	55	550	69.1	-2.6	24	240	0.5
0.4	54	540	65.5	-2.7	23	230	0.4
0.3	53	530	61.8	-2.8	22	220	0.3
0.2	52	520	57.9	-2.9	21	210	0.2
0.1	51	510	54.0	-3.0	20	200	0.1
0.0	50	500	50.0				

4 旧日能試と新日能試のレベル (本文2章 p.20, p.24)

(2013年7月 <http://www.jlpt.jp/about/pdf/comparison01.pdf#search=旧日本語能力試験>)

旧試験

級	認定基準
1	高度の文法・漢字(2,000字程度)・語彙(10,000語程度)を習得し、日常生活をする上で必要な、総合的な日本語能力(日本語を900時間程度学習したレベル)
2	やや高度の文法・漢字(1,000字程度)・語彙(6,000語程度)を習得し、一般的なことからいって、会話ができ、読み書きできる能力(日本語を600時間程度学習し、中級日本語コースを修了したレベル)

新試験

レベル	認定の目安 各レベルの認定の目安を【読む】【聞く】という言語行動で表します。それぞれのレベルには、これらの言語行動を実効するための言語知識が必要です。
N1	幅広い場面で使われる日本語を理解することができる 読む ・幅広い話題について書かれた新聞の論説、評論など、論理的にやや複雑な文章や抽象度の高い文章などを読んで、文章の構成や内容を理解することができる。 聞く ・さまざまな話題の内容に深みのある読み物を読んで、話の流れや詳細な表現意図を理解することができる。 ・幅広い場面で自然なスピーチの、まとまりのある会話やニュース、講義を聞いて、話の流れや内容、登場人物の関係や内容の論理構成などを詳細に理解したり、要旨を把握したりすることができる。
N2	日常的な場面で使われる日本語の理解に加え、より幅広い場面で使われる日本語を理解する程度 読む ・幅広い話題について書かれた新聞や雑誌の記事・解説、平易な評論など、論旨が明快な文章を読んで文章の内容を理解することができる。 聞く ・一般的な話題に関する読み物を読んで、話の流れや表現意図を理解することができる。 ・日常的な場面に加えて幅広い場面で、自然に近いスピーチの、まとまりのある会話やニュースを聞いて、話の流れや内容、登場人物の関係を理解したり、要旨を把握したりすることができる。

N3【新設】 日常的な場面で使われる日本語をある程度理解することができる

- 読む**
- ・ 日常的な話題について書かれた具体的な内容を表す文章を、読んで理解することができる。
 - ・ 新聞の見出しなどから情報の概要をつかむことができる。
 - ・ 日常的な場面で目にする難易度がやや高い文章は、言い換え表現が与えられれば、要旨を理解することができる。
- 聞く**
- ・ 日常的な場面で、やや自然に近いスピーチのまとまりのある会話を聞いて、話の具体的な内容を登場人物の関係などとあわせてほぼ理解できる。

3	基本的な文法・漢字(300字程度)・語彙(1500語程度)を習得し、日常生活に役立つ会話ができ、簡単な文章が読み書きできる能力(日本語を300時間程度学習し、初級日本語コースを修了したレベル)
4	初歩的な文法・漢字(100字程度)・語彙(800語程度)を習得し、簡単な会話ができ、平易な文、又は短い文章が読み書きできる能力(日本語を150時間程度学習し、初級日本語コース前半を修了したレベル)

N4	基本的な日本語を理解することができる 読む ・ 基本的な語彙や漢字を使って書かれた日常生活の中でも身近な話題の文章を、読んで理解することができる。 聞く ・ 日常的な場面で、ややゆっくりと語られる会話であれば、内容がほぼ理解できる。
N5	基本的な日本語をある程度理解することができる 読む ・ ひらがなやカタカナ、日常生活で用いられる基本的な漢字で書かれた定型的な語句や文、文章を読んで理解することができる。 聞く ・ 教室や、身の回りなど、日常生活の中でもよく出会う場面で、ゆっくり話される短い会話であれば、必要な情報を聞き取ることがができる。



5 テキスト・バンクの説明時に使用した聴解テスト・アイテム①と② (本文4章p.73)

①2006年第2回3番「サクラの交配」

アイテム情報「場面：アカデミック、形態：授業、人数：パブリック、機能：情報を授受する」

先生がサクラについて話しています。この先生は、ソメイヨシノというサクラは、どんな方法で増やされたと言っていますか。

一口にサクラと言っても様々な種類がありますが、一番よく見かけるサクラはソメイヨシノという種類です。日本にあるサクラのうち、7、8割がソメイヨシノだと言われています。今から150年ほど前の江戸末期に、二種類のサクラを交配させて、一本のソメイヨシノができました。しかし、サクラの仲間と同じ木に咲いている花同士では受精ができず、種ができません。そこで、最初のソメイヨシノから、枝の一部などを切り取って、それを増やしていきました。このようにして一本の木から作られたので、日本中のソメイヨシノはすべて、同じ遺伝子を持っているのです。

②2005年第2回7番「コミュニケーショントレーニング」

アイテム情報「場面：アカデミック、形態：説明、人数：パブリック、機能：情報を授受する・聞き手に働きかける」

授業で、先生が「コミュニケーショントレーニング」の方法を説明しています。このトレーニングとする質問の種類として、最も適切なものはどれですか。

今日は「コミュニケーショントレーニング」をしましょう。コミュニケーションと言ってもこれは、聞くほう、つまり、良い聞き手になるためのトレーニングです。まず、グループに分かれて、ひとり3分ずつ、あるテーマについて話してもらいます。その間、聞いている人は「あなたの話をきちんと聞いていますよ」ということを示すために、相手の目を見てうなずいたり、あいづちを打ったりしてください。話が終わったら、聞いていた人は質問をします。この時、話した人が伝えようとしたポイントを捉えて、それについて質問するようにしてください。後で、どんな質問をしたのか発表してもらいます。

6 テキスト・バンク分類箱の説明時に使用した聴解テスト・アイテム③と④ (本文4章p.74)

③2006年第2回2番「高校生の大学生活体験」

アイテム情報「場面：キャンパス、形態：私的会話、人数：対人間、機能：情報を授受する・聞き手に働きかける・相手の働きかけに応える」

女子学生が男子学生に大学の行事の手伝いを頼んでいます。この男子学生はどんなことをしますか。

女：今度、高校生が大学生活を体験しに大学に来るんだけど。

男：うん、掲示板で見たよ。講演をきいたり、実験をやったりするんでしょう？

女：そう。私、それを手伝うんだけど、一緒にやらない？

男：えっ？ 何をすればいいの？

女：会場のセッティングだけとか、実験の手伝いとか、ディスカッションに参加するとか。

男：ディスカッションは、ちょっと苦手だなあ。

女：実験はどう？ たぶん、高校生がうまくできなさそうだったら手伝うぐらいだと思うけど。

男：それならいいかな。じゃあ、やってみるよ。

④2006年第2回17番「人形を使った育児体験」

アイテム情報「場面：キャンパス、形態：私的会話、人数：対人間、機能：情報を授受する・自己の感情を表す・個人社会化に関する」

女子学生が男子学生に、人形を使った育児体験について話しています。女子学生は、この育児体験で何が一番大変だったと言っていますか。

女：私、前にね、赤ちゃんの人形を使って、育児体験をしたことがあるんだ。

男：へえ、人形。どんなこと、やったの？

女：抱いたり、おんぶしたりして、一週間、家でも学校でもずっと一緒にいたの。

男：うーん、人形だったら、食べさせなくていいから、あんまり大変じゃなさそうだけど。

女：それが、そうでもないの。ずっと抱いてるのって疲れるんだから。

男：へえ、そうなんだ。

女：でも、それよりね、どこかにぶつけるんじゃないかって、それが心配だったな。授業中も膝にのせてたんだけど、机にぶつからないようにとか、気になって。

男：ふーん。始終そばにいて注意しなくちゃいけないってことか。

女：うん、そうなの。その上、本物の赤ちゃんだったら、泣くわけでしょう。育児って大変ね。

7 解答用紙 (本文6章p.122)

名前： _____

例 ○答え ×答えじゃない △答えかもしれない ?わからない

1	2	3	4
△	?	○	○

1番 (16番)

1	2	3	4

2番 (15番)

1	2	3	4

(後略)

8 インタビュー項目 (本文6章p.123)

以下が半構造化インタビューの項目である。対象者によって、文言および内容の変更は随時行った。以下の項目は対象者が思い出すきっかけとなり、また概念図を構築する際に必要な項目と考えた。

【設問部分】

1. 設問では状況説明と質問があるが、これについて何か自分の心の変化とか、態度の変化など覚えていますか。(それぞれどう聞いているか。意識して分けて聞いているのか)
2. 質問を聞いた時に何か構えるというか、準備して本文の会話を聞くのか。(設問の答えを探すように聞くのか。) =設問なしとの比較のため

【本文】

3. 不安とか、びっくりした、やった!というような感情的なことはありましたか。(情報処理結果からの心的印象)
4. 聞きながら、これは答えだとわかる瞬間ってあるんですか。それは質問の答えを探しているから?

解答だと思うのはどんな状況なのか。何かヒントになる言葉とか、声とか、ポーズとかあったのか。(要因)

5. 聞きながら必要な情報、必要じゃない情報というように分けるのか。
6. 解答だと思った情報があったら、そのあとは聞かないのか。
7. 解答ではない情報(状況説明など)は何のために聞くのか。

【ポスト・クエスチョン】

8. 会話のあともう一度、質問があるがそれはどうか。
9. このときに頭の中に答えの文はあるのか。(明確に、何となく、選択肢を聞いたときに明確になるので今の段階ではないなど)

【選択肢】

10. 選択肢からはどうやって選んだか。
11. 選択肢について何か意見はあるか。どう感じたか。

【全体】

12. メモの存在はどうか。
13. 何が難しいか。簡単か。
14. 何か問題について意見はあるか。

9 留学生のデータ対象者 No. 9 (聴解行動設問あり 1 番) (本文 6 章 p.123)

Q : インタビュアー

A : 対象者

■ : 具体例

[] : テスト・アイテムの内容

インタビュー後、正答を確認し、わからない部分については対象者が納得するまで、説明を加えた。これは、次のテスト・アイテムを聴く前に、心理的な不安を1度解除してもらうためである。

1 番

A : まずその話は雨水は何のために使えるか、ためたら何のために使えるかという質問でした。例として、トイレの水とかしか使えませんね、という話が答えとしてもあるんだけど、**下水の問題になりそう**と思いました。とりあえず、1と2番の答えは×です。やっぱり3番と4番の間に、4番の答えは100%理解できなかったもので、ちょっと曖昧になりましたけど、おそらく3番です。

Q : 話の内容を聞いていて下水に関係あることというふうに思いましたか。

A : はい。

Q : こういう問題は質問があってから聞くんですけども質問の答えを聞こうと思って聞くんですか。

A : **まず話全部を聞いてそこから要点をちょっと考えて答えが来るときにこれがあうかどうか。**この質問を例としたら、雨水がトイレの水としてしか使えないねと言って、それははっきり×。**わざと**そうやって答えが、×の例として入ったと思うので、**作戦として使えます。**(テスト作成者の作戦)話を聞いたら雨水がたまったら、トイレの水にしか使えないんだよね。はっきり言われたのにあと

で話の説明にくると、ちょっと1番ははっきりした答えなのでおそらく×です。

Q：こういうテストの練習をいっぱいしたんですか。

A：うん、そうです。

Q：そういうときにこういう答えってあとで×になるなって。

A：はい、この話の中で話している女の子がすぐに水はトイレにしか使えないねという答えを拒否したので×です。

Q：どの言葉か、あるいはヒントでこれが答えだと思いましたか。

A：別に言葉をヒントとして使いません。もちろん聞きながら、あ、これが答えですねとはっきりわかりますけど、何の言葉だったかは覚えていない。

Q：聞きながらこれは答えだとわかるんですね。

A：聞きながら大事なものは、正しい答えは何ですかなので、言葉自体は覚えておきません。内容だけ。

Q：メモはしない？

A：しない。メモしたら、二つの考えになりますね。一つは聞こえる答え、もう一つは紙の答えになると邪魔になるので、個人のことだと思いますけど。

Q：四つの選択肢を聞く前にだいたい頭の中では答えだというのがあるのですね。

A：はい。この質問の問題は答えの問題でした。話しながら内容がはっきりわかりましたけど、3番と4番の答えは一つの言葉がちょっとわからないと言葉が曖昧になるのでそれが問題でした。

[3番 雨水で植木に水をやること 4番が水道代を節約すること]

A：おそらく3番だと思います。

Q：3番だと思ったのはどうして3番だと思ったんですか。

A：話の中にまずは雨水をたまっています。なぜかというと水の使う、節水のことではありませんでした。問題は雨水が降ったら下水と問題になりそうだったので、そういう話だったので、そういうふうになりました。

Q：下水と植木とちょっとつながらないような気もするんだけど。

A：かもしれない。

Q：簡単でした、この問題。

A：いえ。話は簡単でしたけど、答えは難しかったです。3番と4番の植木と節約はわかるんですが、そこがちょっと難しかったです。

Q：もうちょっとはっきりした原因はわかりますか。

A：やっぱり聞くときに言葉が、文の意味がよくわからなかったのがちょっと困りました。答えの文がわからなかった。

Q：植木？

A：うん。

Q：植木はポットに木が植えてあると植木というんですけど。

A：ああ、そうですか。ああ、やっぱり×です。

[1番 道路に水があふれないようにすること 2番 トイレの水に使うこと]

A：そしたら1番ですね。

Q：下水というのは、下水道、わかりますか。

A：うん。

Q：答えは1番なんですが。

A：植木の問題は、水が下水道から上にくると思ったんです。

Q：会話の内容はわかっていたんだけど、ということですね。

A：はい。

Q：こういうことって時々ありますか。

A：はい、そうですね。こういう質問は四つの答えから二つが見えていますので、いつもその二つの答えから探します。そこから正しい・・・ほとんど作戦がよく使うと思いますけど、たまには失敗しますけど。

Q：その二つの答えにどうやって決めるんですか。

A：答えを聞きながらこれかもしれません、これは絶対に違うという区別をまずして、で残っている答えから、ほとんど二つが残っていますから、選ぶことができます。たいてい頭の中にはパーセンテージでやります。あ、これは絶対に×は0パーセントで、これは正しいかもしれませんが、言葉の一つの意味がわからなかったので60%かなと思って、次はこれは90%かなと思って、あとで、どっちにしようかな、やっぱりこれですねとなります。

Q：四つの答えのうち二つが0%なんですね。

A：うん。

Q：この問題について何か。

A：内容はまあよかったけど、答えについてはちょっと気になりました。

参考文献

- 青山眞子・廣利正代・野口裕之 (2003) 「日本語能力試験の因子分析的検討」『日本語国際センター紀要 第13号』国際交流基金 pp.19-28
- 赤木浩文 (2003) 「言語テスト作成時の諸問題日本語の聴解試験のタイプと形式に関して」菅井英明 (代) 『日本語教育における評価法に関する基礎的資料整備とその分析』文部科学省科学研究費補助金報告書 国立国語研究所 pp.48-62
- 穂田照子 (2009) 「聞く」「聴く」「訊く」: 3つの「きく力」を育む取り組み」『桜美林 Today 第9号』桜美林大学教育センター群 pp.97-112
- 足立章子 (2008a) 「日本留学試験聴解アイテムの下位スキル分析と考察」『2008年日本語教育学会春期大会予稿集』: 首都東京大学 日本語教育学会 pp.151-156
- 足立章子 (2008b) 「日本留学試験聴解スクリプト分析からみる言語能力の考察」『立教大学ランゲージセンター紀要第19号』立教大学ランゲージセンター pp.103-113
- 足立章子 (2009a) 「聴解試験の入力時におけるファセット (facet) —設問について—」『立教大学ランゲージセンター紀要 No.21』立教大学ランゲージセンター pp.59-70
- 足立章子 (2009b) 「日本留学試験が測定している聴解能力—下位スキル分析より—」『専門日本語教育研究』第11号 pp.55-62
- 足立章子 (2010) 「留学生の聴解テスト時における聴解行動の概念モデル」『立教大学異文化コミュニケーション学部紀要 ことば・文化・コミュニケーション第2号』立教大学異文化コミュニケーション学部 pp.122-132
- 足立章子 (2014) 「アカデミック・リスニングにおける留学生の聴解行動—外国人留学生と日本人中学生の比較より—」7月『アカデミック・ジャパニーズ・ジャーナル』第6号 pp.84-93
- 阿部純一・桃内佳雄・金子康朗・李光五 (1994) 『人間の言語情報処理—言語理解の認知科学—』サイエンス社
- 池田玲子・舘岡洋子 (2007) 『ピア・ラーニング入門 創造的な学びのデザインのために』ひつじ書房
- 石川祥一・西田正・斉田智里 (2011) 『大学英語教育学会英語教育学大系第13巻 テスティングと評価—4技能の測定から大学入試まで—』大修館書店
- 石田敏子 (1992) 『入門日本語テスト法』大修館書店
- 伊東祐郎 (2008a) 『日本語教師のためのテスト作成マニュアル』アルク
- 伊東祐郎 (代) (2008b) 『留学生の日本語能力測定のためのテスト項目プールの構築』平成16年度～平成19年度科学研究費補助金基礎研究(A)研究成果報告書 東京外国語大学留学生日本語教育センター
- 伊東祐郎 (2011) 「項目バンクによって広がるテスト開発の可能性」『日本語教育 148号』日本語教育学会 pp.57-71
- 伊東祐郎・横田淳子・福岡理恵子・高野愛子 (2009) 「アカデミック・ジャパニーズにおける読解力—二つの読解試験分析を通して—」『東京外国語大学留学生日本語教育センター集第35号』東京外国語大学留学生日本語教育センター pp.31-46
- 尹松 (2001) 「聴解ストラテジー使用と聴解力との関係について—日本語を専攻とする中国人大学生の意識調査の結果から—」『言語文化と日本語教育 21』お茶の水女子大学日本言語文化学会研究会 pp.58-70

- 梅村修 (2003) 「日本語の聴解指導—聴き取りを容易にする“知識”とは何か」『帝京大学文学部紀要 教育学 28』帝京大学文学部 pp.117-143
- 大坪一夫・三枝令子 (1992) 「第一部本研究の問題と目的 1.3 まとめ」日本語教育学会調査研究第1小委員会『日本語聴解問題の改善に関する考察—最終報告書—』日本語教育学会 pp.16-23
- 大平未央子 (2009) 「ネイティブスピーカー再考」野呂香代子・山下仁 (編) 『「正しさ」への問い—批判的社会言語学の試み』三元社 pp.85-110
- 岡崎眸・岡崎敏雄 (2001) 『日本語教育における学習の分析とデザイン—言語習得過程の視点から見た日本語教育 (日本語教師のための知識本シリーズ 1)』凡人社
- 尾崎茂 (2008) 『言語テスト学入門—テスト作成の基本理念と研究法—』大学教育出版
- 海保博之・原田悦子 (1993) 『プロトコル分析入門 発話データから何を読むか』新曜社
- 笠原ゆう子・木山登茂子・浜田麻里・古川嘉子・横山紀子 (1994) 「聴解テストの条件」『日本語国際センター紀要 4』国際交流基金日本語国際センター pp.35-60
- 加藤清方 (1992) 「日本語能力試験の聴解—テストで測られる聴解能力とは何か—」『日本語学 vol.11』明治書院 pp.32-42
- 門倉正美 (2003a) 「「アカデミック・ジャパニーズ」とは何か」門倉正美 (代) 『日本留学試験とアカデミック・ジャパニーズ 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』 pp.123-132
- 門倉正美 (代) (2003b) 『日本留学試験とアカデミック・ジャパニーズ 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』平成 14 年度～16 年度 科学研究費補助金基盤研究費(A)(1)研究成果中間報告書
- 門倉正美 (2003c) 「日本留学試験の問題点 (2) —「公開試験問題」の分析」門倉正美 (代) 『日本留学試験とアカデミック・ジャパニーズ 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』 pp.207-216
- 門倉正美 (代) (2005) 『日本留学試験とアカデミック・ジャパニーズ (2) 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』平成 14 年度～16 年度 科学研究費補助金基盤研究費(A)(1)研究成果報告書
- 門倉正美 (2009) 「日本留学試験のプロフィシエンシー—「複テキスト性」という観点の提案—」『アカデミック・ジャパニーズ・ジャーナル 1』アカデミック・ジャパニーズ・グループ pp.1-16
- 門倉正美・筒井洋一・三宅和子 (編) (2006) 『アカデミック・ジャパニーズの挑戦』ひつじ書房
- 金田智子 (2009) 「「日本語教育における学習項目一覧と段階的目標基準の開発」の概要—「生活のための日本語」開発に向けて—」『日本語教育における学習項目一覧と段階的目標基準の開発—報告書—』国立国語研究所 pp.1-23
- 木下康仁 (2007) 『ライブ講義 M-GTA —実践的質的研究法 修正版グラウンデッド・セオリー・アプローチのすべて—』弘文堂
- 倉持益子 (2004) 「アカデミック・ジャパニーズが目指すもの—日本留学試験日本語問題の分析と日本語能力試験との比較—」『神田外語大学紀要 16』神田外語大学 pp.267-285
- 国際交流基金・財団法人日本国際教育協会 (1994) 『日本語能力試験出題基準』凡人社
- 国際交流基金・日本国際教育支援協会 (2009) 『新しい「日本語能力試験」ガイドブック』国際交流基金・日本国際教育支援協会
- 国際交流基金・日本国際教育支援協会 (2009) 『新しい「日本語能力試験」問題例集』国際交流基金・日

本国際教育支援協会

- 西條剛央 (2007) 『ライブ講義・質的研究とは何か SCQRM ベーシック編』新曜社
- 西條剛央 (2008) 『ライブ講義・質的研究とは何か SCQRM アドバンス編』新曜社
- 坂本恵 (1994) 「学部留学生に必要な日本語力—聴解・表現を中心に—」『下関市立大学論集 37(3)』 pp.471-482
- 佐藤郁哉 (2008) 『質的データ分析法—原理・方法・実践—』新曜社
- 静哲人・竹内理・吉澤清美 (2002) 『外国語教育リサーチとテストの基礎概念』関西大学出版部
- 嶋田和子 (2003) 「課題達成能力の育成をめざした日本語教育振興協会における研究活動—日本留学試験に対応した新たな取り組み—」門倉正美 (代) 『日本留学試験とアカデミック・ジャパニーズ 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』 pp.17-30
- 嶋田和子 (2005) 「予備教育におけるアカデミック・ジャパニーズに関する一考察」門倉正美 (代) 『日本留学試験とアカデミック・ジャパニーズ (2) 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』 pp.79-92
- 嶋田和子 (2006) 「日本語学校におけるアカデミック・ジャパニーズ 予備教育の新たな取り組み」門倉正美・筒井洋一・三宅和子 (編) 『アカデミック・ジャパニーズの挑戦』ひつじ書房 pp.55-66
- 島田めぐみ (2006) 「日本語聴解テストにおける選択肢提示形式の影響」博士論文：名古屋大学大学院教育発達科学研究科心理発達科学専攻心理社会行動科学講座
- JACET (大学英語教育学会) SLA 研究会 (2005) 『文献からみる第二言語習得研究』開拓社
- JLTA Language Testing 用語集委員会 (2006) 日本語言語テスト学会運営委員会 (監) 『日本語言語テスト学会 言語テスト用語集 日本語言語テスト学会 テスティングの実施規範』日本語言語テスト学会事務局
- 庄司恵雄・野口裕之・金澤眞智子・青山眞子・伊東祐郎・迫田久美子・春原憲一郎・廣利正代・和田晃子 (2004) 「大規模口頭能力試験における分析的評価の試み」『日本語教育 122 号』日本語教育学会 pp.42-51
- 白畑知彦・富田祐一・村野井仁・若林茂則 (1999) 『英語教育用語辞典』大修館書店
- 菅井英明 (2003) 「言語テスト作成時の諸問題」菅井英明 (代) 『日本語教育における評価法に関する基礎的資料整備とその分析』文部科学省科学研究費補助金報告書 国立国語研究所 pp.17-47
- 菅井英明 (2004) 「日本語能力試験と日本留学試験—到達度テストと熟達度テストの相違を見る—」国立国語研究所 (編) (2004) 『世界の言語テスト I 平成 15 年度「日本語教育の学習環境と学習手段に関する調査研究」報告書』国立国語研究所 pp.39-55
- 菅井英明 (2006a) 「アイディア・ユニットを用いた聴解問題作成時のアイテム分析手法の開発について」『世界の言語テスト』くろしお出版 pp.277-298
- 菅井英明 (2006b) 「測定可能な言語能力とは」『世界の言語テスト』くろしお出版 pp.3-14
- 鈴木美加 (2003) 「学部入学前教育の日本語科目の教育内容と試験で測っているもの—予備教育での試験と日本語能力試験、日本留学試験—」門倉正美 (代) 『日本留学試験とアカデミック・ジャパニーズ 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』 pp.31-46
- 武井昭江 (2002) 『英語リスニング論』河源社
- 竹蓋幸生 (1984) 『ヒアリングの行動科学』研究社

- 館岡洋子 (2007) 『ひとりて読むことからピア・リーディングへ 日本語学習者の読解過程と対話的協同学習』 東海大学出版会
- 田中望・姉齒浩美・河東郁子 (1986) 「外国人の日本語行動—聴き取りのコミュニケーション・ストラテジー」 森本政彦 (編) 『言語生活 9月号 No.418』 筑摩書房 pp.62-71
- 因京子 (2003) 「学部留学生の学習活動の現状と意識—九州大学の場合—」 門倉正美 (代) 『日本留学試験とアカデミック・ジャパニーズ 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』 pp.63-72
- 因京子 (2005) 「「日本留学試験・聴読解問題」分析」 門倉正美 (代) 『日本留学試験とアカデミック・ジャパニーズ (2) 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』 (韓日言語文化研究 4集) pp.265-293
- 當作靖彦 (1988) 「聴解能力開発の方法と教材—聴解のプロセスを考慮した練習—」 『日本語教育 64号』 日本語教育学会 pp.59-73
- 富田かおる・小栗裕子・河内千栄子 (編) (2011) 『大学英語教育学会英語教育学大系第9巻 リスニングとスピーキングの理論と実践 —効果的な授業を目指して—』 大修館書店
- 中居順子 (2003) 「日本留学試験 (聴解・聴読解) における課題達成能力の分析—コミュニケーション指導への第一歩として—」 『東京経営短期大学紀要第11号』 東京経営短期大学 pp.257-268
- 長澤邦紘 (1988) 『コミュニカティブ・アプローチとは何か—その理論と展開—』 三友社出版
- 西川一兼・小牧一裕 (2002) 『コミュニケーションプロセス』 二瓶社
- 西原鈴子 (2011) 「日本留学のあり方と大規模テストの可能性」 『日本語教育 148号』 日本語教育学会 pp.4-12
- 日本学生支援機構 (2010) 『平成 22 年度日本留学試験 (第 1 回) 試験問題』 日本学生支援機構
- 日本国際教育協会 (2003) 『平成 14 年度日本留学試験 (第 2 回) 試験問題』 日本国際教育協会
- 日本語教育学会コース・デザイン研究委員会 日本語教育学会 (編) (1991) 『日本語教育機関におけるコース・デザイン』 凡人社
- 日本語教育振興協会基礎日本語教育研究プロジェクトチーム (2001) 『運用能力獲得のための基礎日本語教育—進学希望者を対象として—』 日本語教育振興協会
- 日本語能力試験実施委員会 日本語能力試験企画小委員会 (2004) 『平成 14 年度日本語能力試験分析評価に関する報告書』 凡人社
- 日本語能力試験実施委員会 日本語能力試験企画小委員会 (2005) 『平成 15 年度日本語能力試験分析評価に関する報告書』 凡人社
- 日本語能力試験実施委員会 日本語能力試験企画小委員会 (2006) 『平成 16 年度日本語能力試験分析評価に関する報告書』 凡人社
- 日本語能力試験実施委員会 日本語能力試験企画小委員会 (2007) 『平成 17 年度日本語能力試験分析評価に関する報告書』 凡人社
- 日本語能力試験実施委員会 日本語能力試験企画小委員会 (2008) 『平成 18 年度日本語能力試験分析評価に関する報告書』 凡人社
- 日本語能力試験実施委員会 日本語能力試験企画小委員会 (2009) 『平成 19 年度日本語能力試験分析評価に関する報告書』 アスク出版
- 「日本留学試験のための新たな試験」調査研究協力者会議 (2000) 『日本留学のための新たな試験について—渡日前入学許可の実現に向けて—』 「日本留学試験のための新たな試験」調査研究協力者会議
- ネウストプニー, J.V. (2003) 「アカデミック・インターアクションの理解にむけて」 門倉正美 (代) 『日

- 本留学試験とアカデミック・ジャパニーズ 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』 pp.139-150
- 早川幸子 (1993) 「聴解における音の知覚と語彙力—金沢大学留学生についての調査から—」『金沢大学留学生教育センター紀要 2』 pp.25-36
- 平尾得子 (1999) 「講義聴解能力に関する一考察—講義聴解の特徴と日本語学習者が抱える問題点」『日本語・日本文化 25』大阪外国語大学研究留学生別科 pp.1-21
- 福田倫子 (2003) 「作動記憶理論を採用した文聴解研究の展望」『広島大学大学院教育学研究科紀要第二部第 52 号』広島大学大学院教育学研究科 pp.237-242
- 福田倫子 (2005) 「第二言語学習者における聴解と記憶—ワーキングメモリ理論を枠組みとして—」博士論文 広島大学大学院教育学研究科文化教育開発専攻
- 藤森弘子 (1998) 「聴解試験のテキスト分析—国外の予備教育課程修了試験をもとに—」『東京外国語大学留学生日本語教育センター論集 24』東京外国語大学 pp.49-64
- 札野寛子・辻村まち子 (2003) 「大学生に期待される日本語能力に関する調査について」国立国語研究所日本語教育センター・日本語教育教材開発室『日本語総合シラバスの構築と教材開発指針の作成論文集第 4 巻日本語学習者・教育方法・学習活動』国立国語研究所 pp.564-624
- 古川ちかし・吉野文・足立章子・下平菜穂・高橋優子 (1992) 「ペア形式による調査のプロトコル分析」日本語教育学会調査研究第 1 小委員会『日本語聴解問題の改善に関する考察 —最終報告—』日本語教育学会 pp.175-367
- 堀井恵子 (2003) 「留学生が大学入学時に必要な日本語力は何か—「アカデミック・ジャパニーズ」と「日本留学試験」の「日本語試験」を整理する」門倉正美 (代)『日本留学試験とアカデミック・ジャパニーズ 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』 pp.113-122
- 堀井恵子 (2005) 「日本留学試験の「日本語」シラバスを再考する—「アカデミック・ジャパニーズ」という概念を教育に埋め込む試みから」門倉正美 (代)『日本留学試験とアカデミック・ジャパニーズ (2) 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』 pp.16-29
- 堀井恵子 (2006) 「留学生初年次 (日本語) 教育をデザインする」門倉正美・筒井洋一・三宅和子 (編)『アカデミック・ジャパニーズの挑戦』ひつじ書房 pp.67-78
- 堀井恵子 (2009) 「アカデミック・ジャパニーズ・グループ (AJG) は何を目指し、何を積み重ねることができたのか—今までの会の活動を振り返って—」『アカデミック・ジャパニーズ・ジャーナル 1』アカデミック・ジャパニーズ・グループ pp.75-86
- 松沢伸二 (2002) 『英語教師のための新しい評価法』大修館書店
- 三國純子・小森和子・近藤安月子 (2005) 「聴解における語彙知識の量的側面が内容理解に及ぼす影響—読解との比較から—」『日本語教育 125 号』日本語教育学会 pp.76-85
- 水田澄子 (1996) 「独話聞き取りにみられる問題処理のストラテジー」『世界の日本語教育 6』 pp.49-64
- 宮城幸枝 (2006) 「日本語能力試験 (聴解) の項目分析から見えるもの—2001 年度、2002 年度日本語能力試験 1 級聴解問題の分析を通して—」土岐哲先生還暦記念論文集編集委員会『日本語の教育から研究へ』くろしお出版 pp.47-60
- 三宅和子 (2006) 「「ことばの教育」は何をめざすのか—アカデミック・ジャパニーズの地平から見えてきたもの」門倉正美・筒井洋一・三宅和子 (編)『アカデミック・ジャパニーズの挑戦』ひつじ書

- 房 pp.189-204
- 宮本典以子 (2004) 「日本語聴解試験における諸問題」 修士論文 東京学芸大学教育学研究科 総合教育
開発専攻多言語多文化教育コース
- 村上京子 (2003) 「大学教育と日本留学試験 (1) —学部留学生の大学生活における日本語運用上の困難」
門倉正美 (代) 『日本留学試験とアカデミック・ジャパニーズ 日本留学試験が日本語教育に及ぼ
す影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに
—』 pp.47-62
- 森朋子 (2005) 「大学教育における「アカデミック・ジャパニーズ」を考える」 『東京家政学院大学紀要
第45号』 東京家政学院大学 pp.117-122
- 山下直子 (2000) 「外国人留学生の講義理解—理解に影響を与える要因とストラテジーに関する意識調査
から—」 『日本語教育 107号』 日本語教育学会 pp.95-104
- 山本富美子 (2003) 「留学生に求められる日本語能力と大学学部教学体制の国際化」 門倉正美 (代) 『日
本留学試験とアカデミック・ジャパニーズ 日本留学試験が日本語教育に及ぼす影響に関する調
査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育の連携のもとに—』 pp.73-88
- 山本富美子 (2005) 「アカデミック・ジャパニーズに求められる語彙知識とは—2.4 級語彙・文法事項の
重要性」 門倉正美 (代) 『日本留学試験とアカデミック・ジャパニーズ (2) 日本留学試験が日
本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語呼び教育と大学日本語教育
の連携のもとに—』 pp.110-126
- 横山紀子 (2004) 「第2言語における聴解ストラテジー研究—概観と今後の展望—」 『言語文化と日本語
教育 2004年増刊特集号』 お茶の水女子大学日本言語文化学会研究会 pp.184-201
- 横山紀子 (2005a) 「「過程」重視の聴解指導の効果—対面場面における聴解過程の分析から—」 『第二言
語としての日本語の習得研究 8号』 第二言語習得研究会 pp.44-63
- 横山紀子 (2005b) 「対面場面における聴解過程の分析—「モニター」の適用範囲を指標として—」 お茶
の水女子大学言語文化学会研究会 『共生時代を生きる日本語教育—言語学博士上野田鶴子先生古稀記
念論集—』 凡人社 pp.262-289
- Ryan, G.W. & Bernard, H.R. (2000) 「データ処理と分析方法」 *Handbook of qualitative research (second edition)*,
Sage Publication 平山満義 (監訳) 大谷尚・伊藤勇 (編訳) (2006) 『質的研究ハンドブック 3巻質
的研究の収集と解釈』 北大路書房 pp.165-190
- Aitken, K.G. (1978) *Measuring Listening Comprehension. English as a second language, TEAL Occasional Papers 2*,
Vancouver: British Columbia Association of Teachers of English as an Additional Language, ERIC Document
No.ED 155 945
- Alderson, J.C., Clapham, C. & Wall, D. (1995) *Language Test Construction and Evaluation*, Cambridge University
Press 邦訳: 渡部良典 (2010) 『言語テストの作成と評価—あたらしい外国語教育のために』 春風社
- Bachman, L.F. (1990) *Fundamental Considerations in Language Testing*, Oxford University Press 邦訳: 池田央・
大友賢二 (監) (1997) 『言語テスト法の基礎』 C.S.L.学習評価研究所
- Bachman, L.F. & Palmer, A.S. (1996) *Language Testing in Practice: Designing and Developing Useful Language
Tests*, Oxford University Press 邦訳: 大友賢二・Thrasher, Randolph, H. (2000) 『実践言語テスト作成
法』 大修館

- Benjamin, J.B. (1986) *Communication: Concepts and Contexts*, Harper & Row Publishers 邦訳：西川一廉 (1990) 『コミュニケーション—話すことと聞くことを中心に—』 二瓶社
- British Council, IDP, University of Cambridge ESOL Examinations (2003) *IELTS handbook*, British Council
- Brown, J.D. (1996) *Testing in Language Programs*, Prentice-Hall Inc. 邦訳：和田稔 (1999) 『言語テストの基礎知識—正しい問題作成・評価のために』 大修館書店
- Brown, J.D. (1999) The roles and responsibilities of assessment in foreign language education, 『外国語教育評価学会 研究紀要 第2号』 (JLTA Journal No. 2), pp.1-21
- Brown, J.D. & Hudson, T. (2002) *Criterion-referenced Language Testing*, Cambridge University Press
- Buck, G. (2001) *Assessing Listening*, Cambridge University Press
- Buck, G. and K. Tatsuoka (1998) Application of the rule-space procedure to language testing: examining attributes of a free response listening test, *Language Testing vol.15*, SAGE, pp.119-157
- Canale, M. & Swain, M. (1980) Theoretical bases of communicative approaches to second language teaching and testing, *Applied Linguistics 1*, pp.1-47
- Carroll, B.J. (1961) *Fundamental considerations in testing for English language proficiency of foreign students*, *Testing the English Proficiency of Foreign Students*, Washington, D.C.: Center for Applied Linguistics
- Downing, S.M. & Haladyna, T.M. (eds.) (2006) *Handbook of Test Development*, Lawrence Erlbaum Associates 邦訳：池田央 (監) (2008) 『テスト作成ハンドブック』 教育測定研究所
- Ericsson, A.K. & Simon, A.H. (1984) *Protocol Analysis: verbal reports as data*, Massachusetts Institute of Technology
- Field, J. (2008) The L2 Listener: Type or Individual?, *Working Papers vol.12*, Research Centre for English and Applied Linguistics, University of Cambridge, pp.11-32
- Flowerdew, J. & Miller, L. (2005) *Second Language Listening: Theory and Practice*, Cambridge University Press
- Flowerdew, J.(ed.) (1994) *Academic Listening: Research Perspectives*, Cambridge University Press
- Goh, C. (2002) Exploring listening comprehension tactics and their interaction patterns, *System 30*, Elsevier Science Ltd., pp.185-206
- Green, A. (1998) *Verbal Protocol Analysis in Language Testing Research: A handbook* (Series; Studies in Language Testing 5), Cambridge University Press
- Heaton, J.B. (1988) *Writing English Language Tests* (New Edition), Longman Group UK Limited 邦訳：(財)語学教育研究所テスト研究グループ (1992) 『コミュニケーション・テストィング—英語テストの作り方—』 研究社
- Hughes, A. (2003) *Testing for Language Teachers* (Second Edition) (First Edition: 1989), Cambridge University Press 邦訳：静哲人 (2003) 『英語のテストはこう作る』 研究社
- Hymes, D.H. (1972) On Communicative Competence, J.B.Pride and J.Holmes(eds.), *Sociolinguistics*, Penguin, pp.269-293
- Lado, R. (1961) *Language Testing: The Construction and Use of Foreign Language Tests*, Longman Group Limited, London 邦訳：門司勝・本田漠・吉田一衛・松畑熙一 (1971) 『言語テスト—外国語テストの作成とその利用—』 大修館書店
- McNamara, T. (2000) *Language Testing*, Oxford University Press 邦訳：伊東祐郎・三枝令子・島田めぐみ・野口裕之 (2004) 『言語テストィング概論』 スリーエーネットワーク
- Morrow, K. (1979) Communicative language testing; revolution or evolution, In C. Brumfit & K. Johnson (eds.),

- The communicative approach to language teaching*, London: Oxford University Press, pp.143-157
- Munby, J. (1978) *Communicative syllabus design*, Cambridge University Press
- O'Malley, J.M. & Chamot, A.U. (1990) *Learning Strategies in Second Language Acquisition*, Cambridge University Press
- Oller, J.W.Jr. (1979) *Language Tests at School*, Longman Group Ltd. 邦訳：堀口俊一 (1994) 『言語テスト』 秀文インターナショナル
- Richards, J., Platt, J. & Weber, H. (1985) *Longman Dictionary of Applied Linguistics*, Longman Group Limited 邦訳：山崎真稔・高橋貞雄・佐藤久美子・日野信行 (1988) 『ロングマン応用言語学用語辞典』 南雲堂
- Richards, J.C. (1983) Listening comprehension: Approach, Design, Procedure, *TESOL Quarterly* 17(2), pp.219-240
- Richards, J.C. (1985) *The Context of Language Teaching*, Cambridge University Press
- Rost, M. (1990) *Listening in Language Learning*, Longman
- Ryan, K.E. & Bachman, L.F. (1992) Differential item functioning on two tests of EFL proficiency, *Language Testing* vol. 9, SAGE, pp.12-29
- Savignon, S.J. (1997) *Communication Competence: Theory and Classroom practice (second edition)*, The McGraw-Hill Companies, Inc. 邦訳：草野ハベル清子・佐藤一嘉・田中春美 (2009) 『コミュニケーション能力—理論と実践—』 (原書第2版) 法政大学出版局
- Sherman, J. (1997) The effect of question preview in listening comprehension tests, *Language Testing*, vol.14 No.2, SAGE, pp.185-213
- Shohamy, E. & Inbar, O. (1991) Validation of listening comprehension tests: the effect of text and question type, *Language Testing* vol. 8, SAGE, pp.23-40
- Ur, P. (1984) *Teaching Listening Comprehension*, Cambridge University Press
- Vandergrift, L. (2003) Orchestrating Strategy Use: Toward a Model of the Skilled Second Language Listener, *Language Learning* 53-3, pp.463-496
- Weir, C.J. (1993) *Understanding & Developing Language Tests*, Prentice Hall
- Wu, Y. (1998) What do tests of listening comprehension test? - A restrospection study of EFL test-takers performing a multiple-choice task, *Language Testing* vol.15, SAGE, pp.21-44

謝辞

本論文の執筆にあたっては、多くの先生方のご指導と多くの方々のご支援を賜った。

まず、指導教授ならびに主査であられる桜美林大学 大学院国際学研究科 専任教授の宮副ウオン裕子先生には、論文の構想から最終稿まで、論理的思考、論理的構造、客観的判断など、精緻なご指導をいただいた。また、精神面では終始一貫して激励して下さった。そして、博士課程後期の期限内に執筆が終わらず、論文博士の道を選択する際にも励ましていただいた。宮副先生の暖かいお言葉がなければこの論文を書き上げられなかったであろう。宮副先生からは、将来自分が指導者となったときにどのようにあるべきかという姿勢も教えていただいた。心より感謝申し上げます。

次に、副査であられる桜美林大学 国際学研究科 専任教授の佐々木倫子先生と堀口純子先生、そして東京外国語大学 留学生日本語教育センター センター長の伊東祐郎先生に感謝申し上げたい。論文の試問時にお三人からご鞭撻いただき、次のステップへと進むことができた。ご指摘いただいた点、ご教示いただいた点は、本論文の論理性を高め、完成度をあげることに繋がった。伊東先生からは、テスト専門家としての貴重なご意見とご教示をいただき大変参考になった。二回の試問でご指導いただいたお三人に深謝申し上げます。論文博士の予備審査の際には、堀口先生から詳細なご指摘を受け、読み手の視点の大切さを教えていただいた。論文の問題点のご指摘だけではなく、良い点も必ず挙げていただき、それが励みとなり書き上げることができた。伊東先生からは研究の発信先を明確にすること、読み手を明確に誘導して読みやすくすることなどの注意をいただいた。最終審査を前に、副査になられた桜美林大学言語教育研究科 専任教授の齋藤伸子先生からは、研究者として興味深い内容であると言っていただき最後の追い込みに力をいただいた。

そして、論文をご精読いただいた清泉女学院短期大学教授の中村洋一先生、元香港理工大学副教授の菅井英明氏、桜美林大学大学院博士後期課程で宮副先生のゼミに同期入学した金英実博士に感謝申し上げます。中村先生はテスト専門家であられ、論文のテストに関する適切なご指摘、参考文献のご紹介、現在のテスト研究の視点などを詳しくコメントしていただいた。菅井英明氏には、論文の全体像、明快さについてご指摘いただき、ある部分では評価していただいた点もあり、非常に励みになった。金英実博士には二度にわたり論文を精読していただき、論理性、明快さ、表現について詳細にご指摘いただいた。金英実博士とはクラスメイトとして、辛い時苦しい時には励まし合い、お互いに鼓舞し続けた事で、常に未来に希望を持ち続けることができた。

本研究のデータ提供をしてくれた留学生、日本人大学生に感謝申し上げます。彼らの協力により聴解テスト研究が一步でも進められたと願いたい。

最後に、元東京女子大学現代文化学部教授 西原鈴子先生に感謝申し上げます。もう 20 年ほど前のことだが、博士号を取得したご友人のことを話して下さった。そのお言葉がずっと心に残り、今回論文を書き上げる力となっていた。そして、論文執筆中に迷った際にも、快く手を差し伸べて下さった。心からお礼を申し上げます。

非常に長い時間がかかった論文であったが、多くの方々の支えのおかげで最後まで書くことができた。私一人では書けなかった論文である。すべての方々に、再度感謝を申し上げます。今後は、研究者として皆様から賜ったお言葉を大切に精進していく予定である。