

文化審議会国語分科会日本語教育小委員会
「日本語能力の判定基準」等に関するワーキンググループ（第4回）議事録

令和2年11月26日（木）
10時00分～12時00分
W E B 会 議

〔出席者〕

（委員）櫻井委員，長沼委員，根岸委員，野口委員，南風原委員，眞嶋委員（計6名）

（文化庁）柳澤国語課長，津田地域日本語教育推進室室長補佐，増田日本語教育調査官，北村日本語教育専門職，松井日本語教育専門職ほか関係官

〔配布資料〕

資料1 第3回日本語能力の判定基準に関するワーキンググループ議事録（案）

資料2 「日本語教育の参照枠」二次報告（案）－日本語能力評価の考え方について－

資料3 日本語の能力判定に係る試験等一覧に掲載する情報について（案）

参考資料1 日本語能力の判定基準に関するワーキンググループの進め方について

参考資料2 平成30年度日本語教育総合調査－日本語の能力評価の仕組みについて－ 概要（一部抜粋）

〔経過概要〕

- 1 事務局から配布資料の確認があった。
- 2 議事1の日本語能力の判定基準について，事務局から資料2「「日本語教育の参照枠」二次報告（案）－日本語能力評価の考え方について－」及び資料3「日本語の能力判定に係る試験等一覧に掲載する情報について（案）」について説明を行った後，意見交換を行った。
- 3 次回「日本語能力の判定基準」等に関するワーキンググループの実施予定について確認した。
- 4 質疑応答及び意見交換における各委員の発言は次の通りである。

○根岸座長

ただいまから第4回「日本語能力の判定基準」等に関するワーキンググループを開会いたします。本日も遠隔による審議となります。本日の会議では，カメラは定点設置となっており，発言者の顔が写りません。よって，御発言いただく際には，お名前を言ってから御発言くださいますよう，委員の皆様の御協力をよろしくお願いいたします。それでは，配布資料の確認をお願いいたします。

○松井日本語教育専門職

配布資料等について説明いたします。まず，本日のワーキンググループについてですが，本日，定員数を満たしているということを御報告いたします。次に，配布資料について御説明いたします。まず配布資料1は，「第3回日本語能力の判定基準に関するワーキンググループ議事録（案）」です。配布資料2は，「日本語教育の参照枠」二次報告（案）」です。配布資料3は，「日本語の能力判定に係る試験等一覧に掲載する情報について（案）」です。参考資料1として「日本語能力の判定基準に関するワーキンググループの進め方について」，参考資料2「平成30年度日本語教育総合調査－日本語の能力評価の仕組みについて－概要（一部抜粋）」です。以上が本日の配布資料となっております。

○根岸座長

ありがとうございます。配布資料1の前回議事録（案）については，御確認いただき，修正の必要がある箇所がありましたら，本日より1週間を目途に事務局まで御指摘をお願いいたします。なお，最終的な議事

録の確定については、座長一任とさせていただきたいと思いますが、よろしいでしょうか。

それでは、本日の審議に入ります。本日の議事は、日本語能力の判定基準についてです。主に配布資料2と3について検討を行います。まず、配布資料2「日本語教育の参照枠」二次報告－日本語教育の参照枠における評価の考え方について－を御覧ください。配布資料2については、第3回のワーキンググループ開催以降に委員の皆様から御意見をお寄せいただき、加筆・修正を加えております。特に眞嶋委員、長沼委員におかれましては、多大なる御協力を頂き、感謝いたします。それでは、事務局から資料の説明をお願いいたします。

○松井日本語教育専門職

事務局より配布資料2「日本語教育の参照枠」二次報告－日本語教育の参照枠における評価の考え方について－の加筆事項について説明いたします。委員の委員方からコメントや加筆案等を頂き、事務局の方で加筆いたしました。主な加筆事項について説明いたします。

まず、「1.日本語能力評価の現状と課題」です。丸が四つございますが、一番下の丸、「地域の日本語教室等で日本語を学ぶ学習者の中には、試験による評価を必要としない者も少なくない。そのため、個々の現場で実施できるポートフォリオによる評価などの代替的評価の方法と事例についても、幅広く示していく必要がある。」という一文を加筆いたしました。

続きまして、「3.日本語教育の参照枠」における言語能力観と評価（何を測るのか）」というところ、ここについては、「①一般的能力」の部分について、「叙述的知識」以降、用語について少し整えました。

次、少しページが飛びますが、「5.日本語能力の判定試験と「日本語教育の参照枠」の対応関係を示す方法」の1ページ前の部分、「(3)筆記試験によらない評価の事例」というところにつきましては、一番初めの説明、「筆記試験によらない評価は、代替的評価とも呼ばれ、」というところに1段落文章を加えました。加えまして、ポートフォリオ評価の事例として、「地域の日本語教室では、吹田市国際交流協会、磐田国際交流協会等のポートフォリオの作成という事例を加えました。続きまして、国際交流基金のJFスタンダードについても書き加えました。

続きまして、「③自己評価の例、相互（ピア）評価の例」につきましても、「CEFRスイスプロジェクトの自己評価チェック表」というところについても説明を加筆いたしました。5章、「日本語能力の判定試験と「日本語教育の参照枠」の対応関係を示す方法」というところにつきましては、「(2)マニュアルにおける対応付けの手続き」という五つの段階で行う具体的な手順、作業等々を加筆しました。①習熟化の事前課題という四角の囲みの部分からは全て加筆いたしました。五つの段階で具体的にどのような作業、活動を行っていくかということについて書き加えております。

次ページの「ワークショップ」、「②Specification」明確化という部分についても大幅に書き加えました。

「③標準化トレーニング/レベル設定」についても、具体的な定義について書き加えました。「④Standard Setting Procedures（基準設定手順）」についても、どのような方法があるかについて書き加えました。

「6.社会で活用される日本語能力の判定試験に求められる要素」という章でございますが、ここにつきましては、「(1)試験開発に関する基本的な考え方」、①から⑥の説明の下に、それぞれチェックリストとして、何を具体的に整えたらいいのかということに関して、チェック項目を①から⑥まで書き加えました。

続きまして、「(2)社会的ニーズに応える日本語の能力判定の在り方について」、ここにつきましても、丸が5個ありますが、三つ目と四つ目、「日本語教育の参照枠」に基づき、今後、日本語能力の判定が必要となる外国人材の活動分野や業種等による分野別の日本語能力の評価が行われるようになると想定されるが、試験実施機関側が試験により判定する日本語の分野やレベルを社会に広く明示するようになると良い。」ということ。その次、「話すこと」「聞くこと」などの言語活動別に求められる能力レベルが示されることにより、日本語学習及び日本語の能力判定に有効に活用されることが望まれる。」ということ、この2項を書き加えました。

加筆事項については、以上でございます。

○根岸座長

ありがとうございました。全体的に加筆修正が加えられておりますので、改めて章ごとに検討していきたいと思います。「1. 日本語能力評価の現状と課題」です。これまで日本語教育小委員会等で挙げた主な意見を記載しています。小委員会、現状と課題を書き分けて示してはどうかという意見を頂いております。この点について委員の皆様から御意見を頂きたいと思いますが、いかがでしょうか。

○長沼委員

現状と課題を書き分けるという御意見をお伺いしたのですが、今、一緒になっていて、それによって、現状と課題のつながりは読みやすくなっているかとは思いますが、ですので、現状と課題がそれぞれきちんと書かれているかという点からもう一度見直した上で、そのままにするのはありかと思えます。もしくは、表のような形にしてもよいかと思えますが、その一方で対応が分からなくなり読みづらさが増してしまう危険性があるかと感じましたが、いかがでしょうか。

○根岸座長

この点はどうでしょうか。今の御提案は、表にしてもいいかということでしょうか。

○長沼委員

表にできるのか自信がないので、原案の方が分かりやすければ、そのままにするか、全体の統一感ですとか見栄えにもよるのかと思えます。

○根岸座長

現状の記述が課題である場合もあります。例えば三つ目の「話すこと、書くことなどの産出に関する能力を評価するものが少ない」というのは現状で、同時に課題とも言えるかもしれません。そこをはっきり書くかどうかということでしょうか。事務局はこの点はいかがですか。

○松井日本語教育専門職

確かに、現状と課題を分けてしまいますと、対応関係等々が分かりにくくなるということと、現状であり、課題でありということが重なっている部分がございますので、現状と課題を明確に整理して示すというのが難しいということもございます。できれば原案どおりで、進めさせていただき、今後の加筆修正の過程で、整理が必要であるということであれば、改めて整理を検討してもよろしいのではないかと思います。

○根岸座長

分かりました。今、3点目のところで私が言及したところに絡んでですが、日本語能力を判定する試験で、CEFRのレベルに総合的な判断をする場合のロジックが、テストによって違ったりするという現状もあるかもしれません。これは日本語能力の判定試験だけではなくて、ほかの言語の場合でも起こっていることです。例えば言語活動ごとにレベル付けしたあと、全体のレベルを総括したときに違う判定になるということも起こり得るので、この部分については個人的には課題と考えています。統一見解を持っていた方がいいのか、少なくとも現状がどうなっているかということもどこかの段階で調べる必要があるのかと思いました。ほかに何かございますか。よろしいでしょうか。

○野口委員

各試験によって様々な基準、CEFRの中の能力記述文との関係も違ったりしていることについては、それでいいのではないかという気がします。根底には確かにCEFRという参照枠がありそれに則る形ではありますが、各試験にはそれぞれに特徴があるはずで、だから、必ずしもCEFRそのものの準拠であるという言い方をしなくてもいいのではないのでしょうか。基準というのが同じにはならないでいいのではないか

という気がします。例えばある試験があつて、ある言語活動においては、例えばB 2 レベルで対応が付くけれども、この試験で、例えば各技能などでは、B 1 プラスぐらいで少しずれるとかというようなことがあつてもいいのではないかという気がいたします。それはその試験の目的等に応じて違いがあると思います。

この3番目の文言で言うと、「評価のための基準も明確に示されている」とは言い難い。」という部分ですが、そのとおりなのかもしれませんが、これは別に広く共通の基準というものを決めてしまつて、それにもつていけという話ではないと思います。試験一つ一つがどういう基準を決めているのか、その根拠は何なのか、しかし根底には、CEFRを中心とした日本語の能力判定に関わる（基準ではなく）参照枠のようなものがあるとうことを明確にしていけばよいのではないかという気がいたします。

○根岸座長

ありがとうございます。この点はどうでしょうか。

○眞嶋委員

この4番目、バランスが随分よくなったのではないかと感じております。3点目について、「言い難い」と書いてあるので、各試験の評価のこともしなければならぬと思いますが、学習者の方の評価の必要性というのは、一律にあるわけではありません。4技能全部やらなければいけないということもないと思います。ここの書き方ですと、基準が明確にされているとは言い難いということなので、明確にされることは構わないとは思いますが、全ての評価方法を確立せねばならないというのと、誰にでも使えるようなものを作らねばならないというふうにとられると違うと思いますので、誤解がなければよいかと思ひます。

○根岸座長

ありがとうございます。念のために、私が先ほど言おうとしたことは、例えばテストによって、リーディングがA 1で、ほかがB 1とかとなったときに、ならしてしまつて、それでこの学習者はB 1となる場合と、一番下のところにそろえてA 1という判断が出る場合と、あるいはテストの平均点で出す場合等、様々あるということです。ヨーロッパの言語の場合、既にかなり歴史があるのでよいのですが、日本語の場合、皆さんが総合的なCEFRレベルを出すというロジックのところはまだ始まったばかりですので、利用者の側が同じようなレベルの組合せなのに、自分はA 1で友達はなぜB 1なのかみたいなことが起こりうるのではないかと思ひ発言しました。もちろん、テストの特徴を生かすということで、そこに反対しているわけではないのですが、ロジックが様々あることについては見ていて思ひました。

○眞嶋委員

CEFRを作つた方たちは、最初から個人の能力というのは平均化するものではなく、4技能全部同じようになっていくものでもなく、プロフィールと言われるように、鼻が高い人もいれば、額が広い人もいて、それで構わない。それは事実だし、その学習者の目的に合つていけばそれでいいわけですので、プロフィールがみんな同じになるようにするのが目的ではないですから、使い方かと思ひます。

○根岸座長

そうですね、プロフィールの考え方は、Companion Volumeでも強調されているにもかかわらず、最終的にはある学習者をA 1などというラベル付けを周りが求めてしまうのですが、利用する側もプロフィールの方をきちんと見ることを奨励、強調することが必要なのかと思ひました。

○長沼委員

今の趣旨に賛成します。今、英語だと4技能型試験がかなり強調されてきています。大学入試との関係もあり、全ての試験が4技能であるべきみたいな形もあるわけですが、しかし、もっと考えると5領域のテストにはなつていないわけですが。眞嶋委員、根岸委員がおっしゃつたように、それぞれの必要性に応じて、この

技能についての評価は出していて、ここは出していないということを明確に示すために共通参照枠があるのだと思いますので、この試験ではここを測っている、このレベルまで測っているということが、全体の見取り図をもとに渡せるという価値をもう少し強調した書き方にしてはどうかと感じました。

関連したことですが、どこまで厳密にやるかは難しいと思いますが、1ページの現状と課題に続く理念の部分では、現状と課題に対応して答えを出しているのだと思います。その関係をどこまで明示するかは別としても、このような現状があるが、今回は解答できていませんとするのか、ここで解決策を提案することによって、それがある程度解決されるという形になるのかとういことについて、不明瞭な部分があると感じました。理念のところでももう少し解決策などについて明示的に提案することで、回答していることが示されるといいのかと感じました。

○根岸座長

ありがとうございます。それでは、頂いた御意見を踏まえて、更に整理・修文をさせていただきたいと思います。続きまして、2の評価の理念に移ります。ここでは理念を三つ挙げました。ここは、本ワーキンググループとしては大切にしたいところかと思えます。委員の皆様から示し方に関する御提案など積極的に頂きたいと思いますが、いかがでしょうか。

○長沼委員

それでは続きまして、例えば、1ポツのところの、「外国人等が適切な日本語教育を受けられ」というところが、例えばその前の課題と現状の部分の二つ目の丸の最後の文の「留学生とは異なる目的、場面」等として示している課題や、四つ目に全体的に書き加えていただいたところとかに対応しているかとは思いますが、その辺りがこの短い文で「適切な」でまとめられることによって本当に解決されているのかがやや不安です。もう少しどういった必要性があってこの理念を示したのかということを示すことで、現状・課題と理念との対応関係が分かりやすくなるかと思いました。

○根岸座長

分かりました。ありがとうございます。その点はどうでしょうか。表現の修正をした方が分かりやすいということであれば、後で事務局にさせていただこうかと思いますが、いかがでしょうか。

○長沼委員

もう一点よろしいでしょうか。理念の方の2ポツのところでは代替的評価というのが最後の方に出てくると思えます。その前の現状と課題のところでは、代替的評価に触れているのは、四つ目の加筆されたところです。気になったのは、1行目の「試験による評価を必要としない者も少なくない。」の「試験による」というのは大規模試験という意味でよいのでしょうか。試験というのがパフォーマンステストなどを含まないということであれば正しいかと思えます。ただし、逆に言うと、教室内でパフォーマンステストをきちんと設けて、生活者の日本語についても、大規模試験ではカバーされていないような能力で、学習者の必然性が高い領域がカバーされていなかったら、それは代替できちんと教室の中で評価するということが重要ですので、そのようなポートフォリオ評価だけではないものについても加筆できるのであれば、対応関係がより明確になるかと思いました。

○根岸座長

それは、1現状と課題の方を修正するというのでしょうか。

○長沼委員

はい。あとは課題について加筆した四つ目のところに、パフォーマンス評価など、理念のところでも示されているような代替的評価のほかの必要性に関しても加筆するとバランスが取れるかと思いました。

○根岸座長

ではその方向で取りまとめの段階では事務局に修正いただくということによろしいでしょうか。

それから、私の方で申し訳ないですが、3ページ目のところの頭に「この三つの言語教育観の柱に基づき」とあり、話が先に進んでしまうのですが、次のページのところも「言語能力観と評価」というのがあって、(1)に言語能力観で、そこの一般的能力とか四つ挙がっているのと、こちらの理念について三つの言語能力観が並んでいるので、分かりづらいという気がしました。

○眞嶋委員

今御指摘くださった部分は、最初の方は「言語教育観」ではないかと思います。そして、後の方は「言語能力観」ではないでしょうか。使い分けているようです。

○根岸座長

分かりました。

○眞嶋委員

しかし、どちらも大事なことです。そして、このような教育観に基づいて言語能力をどう把握していくかということが大切かと思えます。

○根岸座長

ありがとうございます。あともう一点、気になったのが、言語教育観の③のところ、黄色になっている部分ですが、「基準（尺度）」というのがあります。これはテストをやっている人には、言いたいことが分かるという感じがしますが、この基準（尺度）というものは、どうでしょうか。一般の方というか、日本教育に関わる方たちが読んだときに、意図しているものを伝えることができているでしょうか。

○長沼委員

確かに、専門用語かとは思いますが、例えば段階的尺度など、少し尺度の概念を補うような言葉を足すのではいかがでしょうか。

○根岸座長

それは基準の括弧の中に段階的尺度ということになりますか。

○長沼委員

基準の説明ということであれば、尺度とほぼ同義だとは思いますが、段階性というのが尺度の特徴だとは思っているので、それが数値的な段階であれ、質的な段階であれ、両方とも尺度になるとは思います。この場合、特に質的な段階かとは思いますが、いかがでしょうか。

○根岸座長

テスト関係の方などいかがでしょうか。

○眞嶋委員

C E F Rの古い方か新しい方にあっただかはっきりとは覚えていないのですが、人間の言語能力を物差しにしたときに、それを六つに切るか幾つに切るかというのは、虹を見て7色と思うかどうかということも人や文化によって違うので、絶対的なものではないということと、それから段階と段階の間がきちりと線が引けるものであるという大前提に立たない方がいいのではないかと、言語能力はもう少し柔軟で曖昧な部分があ

るのではないかという含みを持たせるような書き方をどこかにしてあったと思います。ですので、この部分についても杓子定規のような印象を与えるとすれば、それは違うかもしれないと思いました。

○根岸座長

Companion Volumeの中にカラーの虹の写真が確かにあります。その一方で、テストとCEFRを関連付けるStandard Settingのための厳密なマニュアルも幾つか用意されています。ここでの記述はStandard Settingの部分で出てくる話をイメージしながら、基準という言葉が使われたのかと思ったのですが、眞嶋委員おっしゃるとおりに、虹のように切り方は様々だという考え方は、CEFRとしても重要な考え方だと思います。その虹については、6色だと仮定してみるのがStandard Settingの話なのかと思うと、どのように書き込んだらいいのか思いました。この書き方についてはいかがでしょうか。

○野口委員

これは、CEFRを作るスイスプロジェクトのときからになるのかもしれませんが、段階をどう切っていくかということかと思えます。当時作った能力記述文がACTFLやAustralian Second Language Proficiency Ratingsなど様々なものから取ってきて、並べて尺度化していったときに、連続的に並べられるものではなく、あるところで開くところがあるというような記述があったように思えます。虹の場合ですと細かく言えばどこかで切ることにはできるのですが、開いているところを強調して、そこに線を引くというのがいいのか悪いのかということだと思えます。例えば今私は千葉県にいますけど、すぐ近くに江戸川があり、それを越えたら東京都です。大して違いがないわけです。そういうところをどのように表現していくか。つまり、段階をきちんと設定して、位置付けていくことも大事です。しかし、段階と段階の境界線はある程度ボーダーレスです。両方必要だと思います。何も言わないと線があって、それを越えたか越えないかという議論になっていくと思うので、段階と段階の境界線の話はありますが、それはボーダーレスであることはしっかり示しておくことが必要で、正に今の虹の話強調していく形になると思います。

○根岸座長

ありがとうございます。能力というのは連続的なものではあるけれども、CEFRでいったときに六つの段階に切ることになる。そこで、透明性を持ってやるようにしようということを書き込む感じでしょうか。あともう一つは、ここに出すかどうかということもあると思いますが。時間の関係もあるので、事務局と相談したいと思しますので、御意見ありましたらお送りいただければと思います。

○長沼委員

言語能力が連続体であるという考え方と同時に、CEFRの場合、能力段階を枝分かれさせ、細分化するという考え方を取っているの、そこをどう行っていくのかということについては、かなり大きなところからドリルダウンしているというイメージは共有しておいてもいいか感じました。もう一つは、先ほどの段階のときに質的と量的の話をしました。今話を聞いていると、質的な段階である必要はあるのかと思えますし、逆に、テストのようなスケーリングの一つの心理的尺度には、等間隔でという感覚では必ずしもないと思えます。ですので、その尺度のニュアンスがテストの厳密な尺度とは違って来るかもしれないと思いました。特にCEFRは、レベル間の弁別的な基準特性ということをやっていると思えます。特にプロファイル・プロジェクトというのはそういうところを強調していて、それぞれの価値付けや意義付け、意味などの質的なところをきちんと出しているというような部分と、基準特性をきちんと踏まえているのがCEFRのよいところかと思えますので、そのように明示的に書き分けて、意味付けをしていることが分かれば大丈夫であろうと思いました。

○根岸座長

ありがとうございます。南風原委員、お願いします。

○南風原委員

表現だけですが、基準（尺度）という書き方をしていると、基準という言葉だと何か足りないので足したような感じがあって、一つに決まらない曖昧さがあると思います。二つ並べると、又別の考えが出てくるので、一つに決めた方がいいと思いますが、③が基準（尺度）で、その後が評価手法になっています。これは結局、評価の基準と評価の手法ということだと思いますので、基準（尺度）というのを、例えば評価基準ぐらゐの表現でまとめられるのかと思いました。ここで段階か、連続かみたいなことを出すのも細か過ぎると思いますので、照らし合わせる基準とその方法という整理でいいかと思いました。

○根岸座長

明確な案だと思います。それを参考に、修正させていただきたいと思います。よろしいでしょうか。

○野口委員

これは難しいのですが、「モトジュン」（基準）か「ノリジュン」（規準）かということですね。ここの「キジュン」は、「モトジュン」（基準）でしょうか。

○根岸座長

（尺度）と括弧しているところはモトジュン（基準）を意識して言おうとしたのではないのでしょうか。

○野口委員

南風原委員のおっしゃるとおりだとは思いますが、ここで尺度といったときには、連続性をイメージしたのではないかというような気がします。

○根岸座長

先ほど南風原委員がおっしゃったように、「キジュン」という言葉、基準尺度、それから評価手法ということにしておいて、後で又Standard Settingの話が詳細に出てくるので、そこのところで尺度的な議論になるのかという気がします、いかがでしょうか。

○長沼委員

ここで一貫性について書かれているので、私も野口委員に近い考えです尺度は混乱するのですが、単なる基準ではなく、段階性を持った基準であるということがC E F Rの連続性を言うにしても、そこに意味のある区切りがあり、明示されることで次のステップが学習者に示されるというのが意味のある情報になると思いますので、生涯に渡る発達ということを踏まえると、その段階性や見通しが入るとうれしいと思います。

○根岸座長

分かりました。この1行に入れるか、下の方に加筆するかという案もあると思いますがどうでしょうか。

○長沼委員

「一貫した」の部分に段階性についての内容を加筆することについては、南風原委員の趣旨を踏まえますと賛成です。

○根岸座長

分かりました。取りあえずここまでにして、あとで修正案をお示しするような形になると思います。

それでは、3の「日本語教育の参照枠」における言語能力観と評価」です。「(1) 言語能力観について」と「(2) 評価について」の二つに分かれています、記載はC E F Rからの引用が中心になっています。

御指摘などありましたらお願いいたします。

○長沼委員

長沼です。③のコミュニケーション言語活動を改めて読んでみて、やや分かりづらいと思いました。これは以前発言したので、もしこれでいいということであればいいと思いますが、この四つの言語活動の受容、産出、やり取り、仲介というのと、その後の五つの言語活動別というところで、同じ言語活動ということが使われています。コミュニケーションモードとスキルというのは、CEFRだけでなく、ACTFLでもこの二つは切り分けて話をしています。ですので、ここも書き分けをきちんとできればと思います。我々は背景を知っているので読めますが、普通に読むと、ここは何が違うのだろうか、これと同じなのだろうかとか疑問を持つと思いますので、もう少し加筆してきちんと明文化できたら読みやすくなると感じました。

○根岸座長

ここは二つのことが連動して出てきてしまっているのも、初めてこの考え方に触れる人には分かりにくいかもしれません。ですので、モードというのはいいかもかもしれませんが、いかがでしょうか。

次に5ページで、前から議論になっている「総括・広義の評価 (evaluation)」ですが、注では「総括」は一般的であるとは言えず」ということで「言葉を補った」という説明になっていますが、どうでしょうか。

○長沼委員

中黒がよいのか、括弧に総括を入れるのがよいのか、こういうものは併記するようなものなのかというのは、性質や観点が異なるので、中黒でつなぐのがふさわしいかどうか気がになりました。

○根岸座長

もともと総括という言葉は、このような文脈の中ではあまり使われません。総括というと、ある種の幾つかのデータを総括したり、観点別評価の総括をしたりする評定などという感じで使われたりすると思います。このような意味で使われることは私が知る限りないと思いますが、いかがでしょうか。何の説明なく総括といったときに、evaluationの意味で取ってもらえるということはないと思います。もう一歩もし踏み出すなら、「総括・」を取ってしまって、「広義の評価」だけで (evaluation) としておくのが分かりやすいかと思いますが、いかがでしょうか。

○眞嶋委員

同感です。総括という言葉がここで使うのには違和感がありますが、吉島委員たちの和訳の本をしっかりと読んでいらっしゃる方からすれば、総括はどこへ行ったというふうに、無視されているみたいに思われるのもいがかかと思いますが、ですので、ここでこのように参考として別枠に取ってあることで、つなげられるのではないかとも思いますので、このままでも名案かもしれないと思いました。

○根岸座長

ここは訳があえて出ているというところなので、その意味では妙案であるかもしれません。

○南風原委員

一つ前のページになりますが、言語能力観として①から④まで並んでいて、①と②は能力と書いてあります。④は方略なので、これも方略的な能力というふうに統合できると思いますが、③が活動となっていて、能力の整理として合わない感じがするのですが、ここはどういう理解をしたらいいのでしょうか。

○根岸座長

私もここは気になったので、元の英語版の方を見ました。先ほども申し上げたように、言語能力観という

ような書き方ではなかったと思います。私の記憶では、Action-oriented approachの説明の下にこれが出てきますが、明確に言語能力観と書いてありましたでしょうか。特に③が能力ではないという特徴が一番象徴的に出ていると思います。

○松井日本語教育専門職

事務局から説明させていただきます。確かに、CEFRの訳文でも原文でも、ここは言語能力観という言葉は使われておりません。どういう説明になっているかといいますと、ここは、言語使用者・学習者の言語熟達度を構成する能力及び言語能力を四つ挙げているということになっています。主にこれは①から④を構成する上位概念として、言語の熟達度についてはこういうものが関わっているという説明であったと理解しております。それを今回の二次報告で言語能力観というふうにまとめたわけですが、そこに概念の飛躍等があるがゆえに、このような御指摘を頂いたのではないかというふうに思っています。前の案で出していた図では、一番上の熟達度からこの①から④が下りているという感じでした。ここは「日本語参照枠」における言語能力観と評価」というタイトルが適切かどうかということに関わってこようかと思しますので、もしこの3のタイトルを変えるとすれば、「日本語教育の参照枠」における言語熟達度の整理」という書き方になるのかもしれませんが、このことに関しては、事務局で再度原検討させていただきたいと思っております。

○根岸座長

言語能力観というのをもし残すとすれば、ここに対応している章のAction-oriented approachが言語能力観かと思えます。それをもしここに入れるのであれば、次のところに(2)などとして、「言語使用者・学習者の熟達度を構成する能力及び言語能力」というような形で立てて四つ挙げるといふにすれば、上のタイトルはこのままでも大丈夫でしょうか。

○南風原委員

「3. 「日本語教育の参照枠」における言語能力観と評価」の後に括弧で(何を測るのか)とあるので、この見出しだけ見ると、何を測るかの答えがそこに並んでいると思えますが、そこに活動とかがあるとびんときません。見出しの問いと下の答えが対応していないという感じがします。並んでいる言語活動のところを見ると、受容や産出などということが多面的に見ていくということだと思えます。ですので、その能力をどう見ていくか、その諸側面や構成要素、そのようなことがここに書いてあるのだと思えます。そうすると、それに対応した見出し、Action-orientedであれば、正にそれが言語観だと思えます。その場合、そのAction-orientedにはこういう側面があり、こういう一般的な能力もあるという整理なら分かると思えます。

○根岸座長

次のページの「(2) 評価について」と一緒に様々な内容を全部一つのところに入れてあるので、この辺りも混乱している原因だとは思えます。

○南風原委員

(2)の評価についても曖昧です。内容が妥当性、信頼性なので、評価についての評価になるのでしょうか。評価の質的なことであるなら、ここも見出しとは合わないかと思えます。

○長沼委員

この部分は熟達度(プロフィシエンシー)を説明するような図式であるのだらうと思えます。根岸委員がおっしゃったように、ここはAction-orientedという考え方と、その背後にあるコンピテンスを捉えていることを示す部分だと思うので、そこがもう少し明示化されれば誤解が少なくなるかと思えました。

○眞嶋委員

もう一つ追加といたしますか、ここの言語能力観を考えると、①、②は学習者の内面的といたしますか、内在しているものだと思います。CEFRの考え方で、今回の日本語教育の参照枠の最初の方の理念のところ、学習者を社会的行為者と捉えるという文言があったと思います。社会的に活動できる人なのかどうかということを見ないといけない。社会的な行為としてコミュニケーション言語活動を評価するというのではないかと思いますので、①、②、③は並列に並んでいるのではないと思います。しかし、このような示し方だと平板に見えるので、確かに分かりにくいですね。

○長沼委員

CEFRでは、言語的な能力についてはCommunicative Language CompetencesをLinguistic, Sociolinguistic, Pragmaticの3つに分けています。その辺りを少しきちんと明示して、どういう考え方で成り立っているのかというところが示されるといいのかと思いました。凡人社から出ている『プロフィシエンシーを育てる』には、根岸座長も関わられています。そこでは背景の議論があるので、プロフィシエンシーの概念が分かりやすくなっています。あまり理論的にならない方がいいとは思いますが、背景的な考え方を示しておくとも読みやすくなると感じました。

○根岸座長

今、原典を見ましたが、コミュニケーション言語活動というのは、英語ではlanguage activitiesという部分でしょうか。事務局、ここはどこから来たのでしょうか。

○松井日本語教育専門職

Language activityのところを言語活動というふうな言葉で当てています。

○根岸座長

コミュニケーション言語活動という、そのコミュニケーションと付けたのには理由がありますか。

○松井日本語教育専門職

ここは、これを参照するに当たって、吉島訳のものをそのまま持ってきているという整理です。

○根岸座長

そうすると、原典にない「コミュニケーション」がタイトルに補われているということでしょうか。

○松井日本語教育専門職

はい。

○長沼委員

Companion Volumeの方を見ていたのですが、方略はcommunicative language strategyというふう捉え直しがされています。activityはcommunicative language activityで、その前はcommunicative language competencesとなっているので、これが比較的参照性が高いものであるとすると、そちらにそろえて注釈を入れて書き直すのもありかと思えます。方略も様々な方略があり、学習方略ではないと思いますので。

○根岸座長

分かりました。様々な論点が出て、時間内にはまとめられないかもしれないので、タイトルと今の四つ、特に③について整理したいと思います。それから(2)の評価については、いかがでしょうか。

○長沼委員

先ほど根岸座長がおっしゃっていた総括的評価のところ、原典に戻って見ていたのですが、**evaluation**と言っているだけで、**summative**は入れていません。もし原典に忠実にということであれば、現在注釈で既に「一般的であるとは言えず」というような書き方をしているので、広義の評価にしたなどと注釈のところでフォローすれば、眞嶋委員の言われた、総括はどこ行ったのだという議論はなくなるかと思えます。訳文をきちんと提出して、このようになっているので、ここはこのようにしたと書けば、解決できるのかと考えました。

○眞嶋委員

いいと思います。

○根岸座長

ありがとうございます。御賛同を得られましたので、そのようにしたいと思います。

それでは次に入ります。「4. 「日本語教育の参照枠」における多様な評価の在り方と事例」は、(1) 主な熟達度評価の在り方、(2) 評価の種類、(3) 筆記試験によらない評価の事例、の3項目に分けて示しています。かなり加筆修正をいただきましたが、これらの点について御意見やコメントをお願いいたします。

1点いいですか。ルーブリック評価、これは前も発言させていただいたのですが、ルーブリックという単語がCEFRの中には1回しか用いられていなくて、違う意味で用いられています。ルーブリック評価というのは、アメリカ英語ではこういう意味で使います。しかし、CEFRの中ではこのような使い方はないと思います。CEFRのどの部分をルーブリック評価と呼んだのかについて言及しておいていただけると、オリジナルを知っている人が読んだときに、どれのことを言っているのかというのが分かりやすいのかと思います。その点について、事務局、これはどこから来ているのでしょうか。

○松井日本語教育専門職

ルーブリック評価は、CEFRというよりも、例えば**JF Standard**であるとか、「とよた日本語能力判定」もそうかもしれませんが、CEFRからヒントを得て日本で作られたもので行われている評価の中に、文中に示したような形でルーブリック評価が取り入れられていますので、例として挙げています。根岸座長がおっしゃったとおり、CEFRでのルーブリックとは性質が違いますが、日本語教育の実践の文脈から取ってきているという実情がございます。その点は書き方を整理する必要があるかとは思いますが。

○根岸座長

CEFRの中から取ってきている指針（基準でしょうか？）に基づいた判断や分析的評価などは、このルーブリック評価が重なる部分が多いと思います。ここでこのルーブリック評価を出しておいてから、後ろの方はCEFRのものに基づいて書くとなると、関連する概念を違う表現しているという事態が生じているように思います。CEFRに基づいて参照枠を作りましたというところに、CEFRにないものが入っていたときに、日本語教育の方はあまり違和感がないのかもしれませんが、この点はどうでしょうか。

○南風原委員

まず大きなところとしては、(1)の主な熟達度評価の在り方としていろんな評価が並んでいて、その3ページ後に又大きな(2)評価の事例として、ここにも自己評価などが並んでいるので、(1)の分類と(2)の分類の関係が分からないというのが一つあります。

それから、今おっしゃったルーブリック評価、ここも違和感があるのですが、ルーブリック評価のところを見ると、例えばやり取り能力の熟達度をパフォーマンス評価によって測る場合にルーブリックを使うとあります。その前の前のページにはパフォーマンス評価というのが、今度はこれが試験として実施できる評価に組み込まれているわけです。もう少し整理すると、①試験として評価できるものに筆記試験とパフォーマンス評価がある。ルーブリックというのは、そのパフォーマンス評価に使われるものであるとしながら、今

度は②の方の試験によらない評価のところ、ポートフォリオ、自己評価や何かと並んでルーブリック評価とあるので、そこは混乱しているのかと思います。ルーブリックは評価の際の参照の基準で、採点表のようなものですので、ここに他の評価と並べるのは違和感があります。

○根岸座長

今おっしゃっていただいたように、試験によらない評価で使うこともあるかもしれませんが、試験の方でも使われることがあるにもかかわらず、後者の試験によらない評価の方にだけルーブリック評価があり、なおかつ、CEFRの中でこれが出てきていないので、整理が必要かという気がしますが、いかがでしょうか。

○長沼委員

今のところに関連しているのですが、試験による評価のところでは、筆記試験とパフォーマンス評価となっていると思いますが、パフォーマンス評価とパフォーマンステストは、厳密には異なりますが、場合によっては試験の中のパフォーマンステストと捉えて、ペーパーテストとパフォーマンステストとしてもよいでしょうか。試験と評価が並んでいるのも違和感があります。特にルーブリックの作り方によってはより広い評価にも捉えられますし、教室でやる活動でもパフォーマンステストは作れるので、書き方が難しいとは思いつつ、パフォーマンス評価が筆記試験と並んでいることに違和感があります。

○根岸座長

パフォーマンス評価は試験によらない評価でもできるということかもしれませんが、しかし、パフォーマンス試験という言い方になじみがあるかということ、どちらかということとパフォーマンステストの方がなじみがあるように思います。そうすると、筆記試験は筆記テストにするのか、様々影響が出てくるかと思えます。ただ、趣旨としては、筆記試験はよく行われるけど、パフォーマンスを見るテストというのがありますよねということを最初で紹介したくて、試験というものによらないものとして、ルーブリックを入れないとしても三つここでは紹介しています。このルーブリック評価という用語そのものの問題もありますし、ここに入れると、あまり相性がよくないのではないかという感じがしますが、いかがでしょうか。

○長沼委員

ルーブリック評価のところ、パフォーマンス評価とポートフォリオ評価の両方に用いられるとなっていて、このルーブリック評価の3行目のところに評価表と書いてあるところは、評価表より評価基準表とした方がいいのかとも思いますが、CEFRでは、Assessment Criteria Grid、新版ではCriteriaが抜け、Written Assessment Gridのようになっていると思いますが、Criteria Gridという言い方をしていると思います。主にCommunicative Language Competenceの基準を集め、WrittenとOralで整理をされて別途Gridとして設けられていたので、CEFRを見るとこういうルーブリックがあると補ってもいいのかと思いました。

○根岸座長

最初に南風原委員がおっしゃってくださったような、ルーブリック評価と今ここに書いてあるものの、その情報はどこに置くのかという話と、中身としては、このルーブリック評価は日本語教育の方は分かりやすいとすれば、もう少しCEFRとCompanion Volumeなどを参照して補い、CEFRとの関連が見えるように、記述を補足したりする必要はあるかと思いました。評価の種類の方はいかがでしょうか。

○櫻井委員

評価の種類ではないのですが、熟達評価のやり方に挙がっている評価のやり方のところと、その前の議論で出ていたCEFRで評価の議論に挙がっている妥当性、信頼性、実行可能性の三つの概念の関係性についてですが、特に信頼性という点に関して、例えば代替的評価や、ポートフォリオによる評価など、そういったものの信頼性というものは、ここに書かれている信頼性では担保が難しい部分があるのかと思ひまして、

その辺の関係性はどうか捉えられているのかと疑問に思いました。

○根岸座長

ありがとうございます。ここでの信頼性の書き方というのは、信頼性のある側面であると言えます。実はこの文章の構造的には、もっと後ろの方に信頼性についてのチェックリストが出ていて、そこで詳しく書いてあります。ただ、ここまではこれしか出てこないの、あの観点からするとこれらはどうなのだという話になるかと思えます。それはどうしたらいいのでしょうか。

○櫻井委員

今の書き方だと、その後挙がっている評価の在り方というものが三つの評価の概念というものを補助しているかのような印象を与えてしまうので、書き方を変えるなり、工夫する必要があるのかと思いました。

○根岸座長

例えば評価回数を増やしたりして、試験によらない評価であっても信頼性は担保すべきだということを書くかどうかでしょうか。そもそも信頼性の確保が難しいタイプのものもなくはないかもしれませんが、少し情報を補う書き方ができればよいのでしょうか。

○櫻井委員

ここで言われている信頼性を、例えばランゲージポートフォリオなどがそれほど重要視しているかといったときに、そうではない側面もあるのかと思うので、強調し過ぎるのもどうかとは思っています。

○根岸座長

ももとの三つ、妥当性、信頼性、実行可能性が挙がっているところで、もしかしたら、これら全ての要件を高いレベルで実現するのは難しく、現実的な選択をしなければならぬということには触れておいて、評価によっては犠牲にしたりする必要があるというようなことにも触れてみてはどうでしょうか。

○眞嶋委員

受験勉強などで試験といえば答えが一つのものだと刷り込まれているといいますか、そういう人が見たら、量的に筆記試験で答えが1個しかない、それが分かるようになった、記憶したということを確認する試験で、それはもちろん妥当性も信頼性も担保できる、すべきものだということは言えると思います。一方で、先ほどから櫻井委員もおっしゃったような、1人の学習者の習得の状況を記録したものとか、ポートフォリオとか、丸かバツか答えが1個しかないものではないものについても一緒に示そうとしているので、評価にはこういうのが様々ありますよということを羅列してしまっているのだと思います。一般の人が見たら、すっきりしない、それぞれの関係が分からないと思われるかもしれません。試験といったときに、答えが一個しかない筆記試験の話をしている場合と、試験という大きな傘の下に口頭の会話試験もあれば、筆記試験もあるというようなことを考えて試験と使っている場合もあるかと思えます。それらを列挙すると、全部同じような様々な種類というようにしか見えないところが、苦しいと思いました。

○櫻井委員

もう一つだけ、その信頼性のところに、特に「学習者の序列が同じになるかどうかの蓋然性」という記述があります。それは集団筆記テストでは非常に重要な点であって、担保はなされるべきですが、生活者としての外国人の自己評価と言ったときには、そもそも序列を重要視しないために、この日本語教育の参照枠を作ろうという部分があったのではないかと思うので、それに相反するような書き方がなされて、羅列されるのはどうなのだろうかと思いました。

○根岸座長

信頼性の書き方が、今おっしゃったように、大規模テストの要件みたいな場合の書き方になっているので、もう少し引いて、全体がカバーされるような書き方に変えるのか、あるいはこの表現のところに注釈を入れるという感じでしょうか。

○野口委員

この信頼性はかなり厳密な定義になっていますが、ここで言いたいのは、教師の単なる主観であってはいけないのだという話だろうと思います。何かしらの根拠をもとに同じ人が同じ学習者を見たときに、習得が進んでいたら、評価は変わりますし、同じ程度だったら、同じ評価になるようにということであって、2回やって完全に得点が一致するという話ではない。しかし、ある程度の客観性は持たなければいけないのではないかという表現をここでは盛り込んだほうがよいのではないかと思います。

○南風原委員

どれぐらいの抽象性が適切かという話だと思いますが、2回実施とかというのは非常に特定のです。なので、もっと曖昧というか、もっと広く、例えば評価の結果がぶれずに安定しているという程度のことでいいのではないかと思います。

○根岸座長

ありがとうございます。具体的な御提案だったので、文言を書き直してみて、皆さんにもう一回諮るということでいかがでしょうか。この段階では、南風原委員がおっしゃったぐらいの方がいいのかと思います。次に、評価の種類のところはかなり記述が多いですが、これはこんな感じでよろしいでしょうか。

○長沼委員

先ほどからの議論を聞いていますと、重複する情報がどこに出てくるかが難しいという感覚は持っています。そして、この評価の種類を本当にみんなに読んでほしいのか疑問に思ってきて、これは参考資料としてもいいような性質のものではないのでしょうか。その一つ前のところの評価の三つ種類を、先ほど御指摘もあったように、きちんと理解してもらえれば、大枠としてはいいのかという気がします。

○根岸座長

細かい情報にもかかわらず記述の量が多いので、この部分を読んでいると、全体がぼけてしまうという感じがします。外形的には大きな変化ではありますが、参考資料としてもよいのではないかと思います。

○眞嶋委員

私も初めてCEFRの原著を読んだときに、何でここにこのような話がだらだらとあるのか、評価法の概論の教科書のようなうだと思ひ、違和感がありました。今御提案あったように、参考資料としては、意味がないことはないですが、これを本論として出さなくてもいいのではないかと思います。

○根岸座長

では、今御提案いただいた方向で処理して、事務局とも相談しますが、いかがでしょうか。

○長沼委員

全般的に先ほど根岸委員から御指摘があったところだと思いますが、信頼性については信頼性についての記述の部分はCEFRの定義なのか、我々の定義なのか曖昧です。書いてあることがCEFRからの引用だとすると、囲みへ入れるなどして、レファレンス性を高めておけば、もう少し何がどこで議論されているのかという混乱は少なくなるかと感じました。

○根岸座長

CEFRを参照しながら作ったということだとすると、CEFRのどの部分なのかということは、用語的にも明らかな場合はいいのですが、それ以外の部分については参照性があった方がよいと思います。先ほどの部分もそうですが、ここで全部説明し切れなくなると、最終的には元の方に参照していただくしかないと思います。そうしたときに、参照しやすいようにできていた方がいいのかと私も思います。

次に、4の多様な評価の在り方と事例ということで、(1)主な熟達度評価の在り方、(2)評価の種類、(3)筆記試験によらない評価の事例の3項目に分けて示しています。この部分についてはかなり加筆修正をいただきましたが、これらの点について御意見、コメントをお願いします。

○長沼委員

(3)筆記試験によらない評価の事例のところ、パフォーマンス評価の例や、ポートフォリオ評価の例が幾つか加わってきたと思います。この部分のポートフォリオ評価の例にヨーロッパ言語ポートフォリオはきちんと出した方がいいのかと思いました。あと、その上にACTFLが出ているので、そうであるならば、アメリカのLinguaFolioも2020年度版でアップデートされて使われているようですので、そういったものも整理して書いた方がいいかという気はします。

○根岸座長

私もポートフォリオ評価の例でCEFRの流れで来ている割には、ヨーロッパ言語ポートフォリオの紹介がなかったので、紹介しておいてもいいのかと私も思いました。

○長沼委員

プロジェクトとしては、アイルランドの事例なども見せるとするのであれば幾つかあるとは思いますが。

○櫻井委員

同じところの意見ですが、ポートフォリオ評価の例で、CEFRの方を入れていただいた方がいいのかと思ったのと、ここで子供のものが挙がっています。内容を拝見しましたら、日本語のレベルチェックシートなども入っていたりして、そのポートフォリオが何なのかといったときに、学習記録ではない、作品集だけではないレベルに関するものがあり、そのレベルが何を意味するのかということになってきたりすると思いますので、誤解を与えないような事例がいいのかと思いました。

○長沼委員

個別のかなり小さい団体の例はたくさんあると思います。英語教育でも大学レベルでやっているポートフォリオプロジェクトなど、様々なものがある中、このかなり限定した小さい例が挙がるのに違和感があります。特別に代表的なものとして日本語教育で認知されているのであればそれでいいのかと思いつつ、私は日本語教育の専門でないので、見たときに、ややここにそぐうのかという疑問があります。

○根岸座長

この点について、事務局いかがでしょうか。

○松井日本語教育専門職

ここではあくまでも一例として挙げました。なぜこれを挙げたかといいますと、文化庁国語課の補助事業であるとか、生活者としての外国人に対する日本語教育支援の施策等々で作っていただいたものですので、それらの事例については幅広く示しておきたいとは思いますが、御指摘はもっともだと思います。いきなり

個別のものが挙がっている一方でE L Pが挙がっていないこともございますので、ここは全体を整えていけたらと思っております。

○根岸座長

E L Pを入れるとして、広く使われているものからローカルな形みたいなふうに並んでいると、そのように並べたというのが分かると思いますが、真ん中にローカルなものが出ていて、そして次は基金の事例なので、広がったり狭まったりするような感じなので、見せ方を修正してはどうかと思います。

○長沼委員

文化庁にひも付いているのであれば、脚注で示すというやり方もあるかと思います。

○根岸座長

ほかにもございますか。大丈夫ですか。時間の関係もあるので、進みたいと思います。5の日本語の大規模試験と日本語教育の参照枠の対応関係を示す方法ということで、ここから大規模試験に関する章に移ります。今回、(2) マニュアルにおける対応付けの手続についての部分、五つの手続について大幅な加筆を行いました。更に加筆して示すべき観点などありましたら、コメントをお願いしたいと思います。

○長沼委員

こちらは事前に宿題で頂いて、原典に当たってかなり見させていただいて、そのときにはコメントしなかったのですが、(2)のマニュアルの対応付けの手続の中の④で統計的な手法が列挙されていると思います。ここは原典ではすごく細かいことが書いてあって、それを受けてこの章では方法だけが列挙されています。しかし、方法が列挙されていても理解はできないので、統計的な手続が必要な場合は、例えばC E F Rのこの原典へ行くことと示されているので、参照すべしと説明するにとどめるか、参考資料に追い込むかした方がよいと思います。ここに列挙されていてもあまり助けにならないという気がしました。その前の部分は原典のマニュアル本文をもとに加筆したりして、読みやすいようにはなっているかとは思いますが、ここに関してはどう手を入れていいか分からなかったもので、是非議論していただきたいと思っていました。

○根岸座長

私もここは気になっていて、原典を見ました。そこで、この部分に関しては受験者で見ると、テスト項目で見ると、I R Tを使って統計的にやる場合の三つの方法があるが、より詳細には原典を確認せよという形がいいように思います。もし本格的にこれらを全部説明し出したら、何十ページもかかってしまうということだと思いますが、いかがでしょうか。

もう一つ気になったのは、このStandard Settingを行うといったときに、誰がどんな目的ですることが日本語教育の場合生じるのかということだと思います。大規模テストをやっている人たちのStandard Settingだとすると、主に9とか10というI R Tのデータがあるので、それで受験者を見ていたり、テスト項目を見ていたりすることになると思います。しかし、大規模テストを行っている以外の人たちがここで挙げたようなStandard Settingをやるというのはどれくらい想定したらいいのでしょうか。

○松井日本語教育専門職

Standard Settingは、誰がどのような目的で行うかということに関しては、次の資料で御検討いただきます。現状ある日本語試験の中には大規模なものもありますし、比較的小規模でも一般に向けて行っている試験もあります。ですので、ここでは主に現状実施されている試験団体向けに示しているということでございます。

○根岸座長

そうすると、先ほど言った三つ程度の情報は必要で、IRTで分析できるデータがある場合は、ある団体さんは9、10を参照してください、そうでない場合は、タイプによって1から8のどれかを使ってやってみてくださいというような情報提供になりますか。

○松井日本語教育専門職

はい、そのようになろうかと思います。

○根岸座長

確認ですが、マニュアルは公的に翻訳されているわけではないので、Standard Settingをやる場合には、英語版を見るということになりますか。

○松井日本語教育専門職

マニュアルに関しては、公式に日本語で翻訳され一般に公開されているものはございませんので、現状では英語のものを参照いただくことになるかと思います。

○根岸座長

そうしますと、ここでは全体像を示して、詳しくは英語版を参照してくださいという形になりますね。

○松井日本語教育専門職

現状ではそのような状態になっています。

○根岸座長

分かりました。

○南風原委員

先ほど根岸委員が言われたような大きなくくり、test-centeredというようなことについての説明があれば、こんなふうにするのだという情報になります。ですので、一つ目が、例えばテスト項目を専門家が評定して判定する方法というような説明。もう一つは受験者集団になるとは思いますが、その三つ目のIRTというのがこれと並ぶと中身が分からなくなるので、ぴんとこなかったのですが、つまり、テスト中心、受験者焦点、そして三つ目にIRTが来るということになるのでしょうか。

○根岸座長

これはマニュアルの中にIRTと書いてあるのですが、私なりの解釈で言うと、IRTだと、テストアイテムと受験者が同じ尺度の上に乗っかるのではないかと思います。それで、私の経験で言うと、産出技能の方のものは、並べた受験者の方を見ながらパフォーマンスの違いを見てStandard Settingをするというようなことはイメージされているのかと思います。受容技能の方は、テストアイテムを並べておいて、アイテムの質的な変化があるところで切っていくということか思います。ですので、IRTという名前は、並びとしては違和感があって、examineeとtest-centeredではないですが、その両方参照しているということだと思います。

○南風原委員

分かりました。

○長沼委員

原典を見ていたのですが、このテストの2番目のカテゴリーは、**Without any empirical test taking data** みたいに書いてあるので、データなしでテストそのものを見ている感じかと思います。

○根岸座長

それは2番目というか、それを修飾しているのが、多分同じところで、**in the sense that it can be applied without any empirical test taking data**というのは、多分最初の二つ両方ともじゃないかと思います。問題の性質を見ながらやるのと、**test taker**の、例えばここで言う**Body of Work**というのは、ある学習者が書いたり話したりした様々なものをまとめたものを見て、こういう感じだとこのレベルだろうということをするということなので、多分どちらも**empirical data**ではありません。先ほどのIRT観点から言うと、上の8番までは**empirical data**ではないということになります。以上、複雑な議論ですので、南風原委員にも御賛同いただいたので、こういう三つのアプローチがあるという説明で、詳細は後ろか注かという形で処理するのがいいかと思いました。

○野口委員

そのまとめで結構だと思いますが、結局これはCEFRに付けるという話ですよね。我々が今考えているのは、日本語教育の参照枠ということでよいでしょうか。

○根岸座長

はい。

○野口委員

そこですけれども、日本語教育の参照枠が、CEFRとかなり類似したもの、ほぼ等しいとまでは言えないと思いますが、違うところが出てくるはずで、このCEFRに付けるということ、マニュアルに従って書くのではなくて、大枠でどんなやり方でできるのか。もしかしたら、日本語のテストの開発主体が、こんな方法を考えたというのがあったら、それはそれでいいのだと思います。しかもそれは、今挙がっていた三つカテゴリーの中のこれに相当するのだという新しいものがあったとしても、**Council of Europe**のカテゴリーにはないものでもいいような気がいたします。そういう意味で、ここを細かく書くことはやめた方がいいかという気がいたします。

○根岸座長

ありがとうございます。ハイブリッド型のようなものも実際には使われている方たちもいると思うので、これはCEFRの分類を紹介した、**Council of Europe**のウェブサイトにあるこのマニュアルの例を載せたということを書いておけば、これを基に何か考えてくださいということでもいいかと思います。

先ほど言及があった6の方に移りたいのですが、社会で活用されている日本語能力を判定する大規模試験に求められる要素ということで、まず、「社会で活用される日本語能力の判定試験に求められる要素」として、(1)試験開発に関する基本的な考え方、(2)社会的ニーズに応える試験の在り方についての2点を挙げています。(1)試験開発に関する基本的な考え方としては、①から⑥までの六つの大規模試験において最低限必要となる要素を並べ、それぞれチェックできるよう項目例を挙げています。本日はこのようなチェック項目の示し方及びその項目について検討していきたいと思いますが、いかがでしょうか。先ほど信頼性についての議論が既に出ましたが、こちらの方に信頼性などについて詳細に書かれています。

○長沼委員

チェックリストの部分は我々の方で様々コメントしたものを集約された形になっているので、ある程度網羅性があるとは思いますが、これが「ザ・チェックリスト」であるように思われてしまうと、危険性があるかと思うので、CEFRと一緒にですが、これはあくまで例示ですという形にした方がよいかと思います。

○根岸座長

様々なドキュメントから事務局が取りまとめて、チェックリストにしてくださったということでしょうか。

○松井日本語教育専門職

はい。

○根岸座長

ということですが、いかがでしょう。

○眞嶋委員

すごく分かりやすく、構造として取っつきやすいと思いますが、先ほど長沼委員がおっしゃったように、これが「ザ・チェックリスト」で、これさえあればというようなところに行かないようにという注意は必要ではあるものの、一般の方にも分かりやすいのではないかと思います。

○根岸座長

ありがとうございます。ここはアクセシブルになりました。先ほどのStandard settingのところはわかりにくいような感じでしたが、ここは皆さん使いやすいかと思います。よろしいでしょうか。

それでは次に、(2)の社会的ニーズに応える日本語能力の判定の在り方についてですが、このほか、将来的に求められるであろうという観点がありましたら挙げていただきたいと思いますが、いかがでしょうか。ほかにも細かい部分で御指摘あるかと思しますので、お気づきの点があれば、事務局までお願いします。提出していただいた指摘、課題を踏まえて、再度資料を練り直していくこととしたいと思えます。事務局、よろしいでしょうか。

○松井日本語教育専門職

はい、よろしくをお願いします。

○眞嶋委員

時間がないところで恐縮ですけれども、この社会的ニーズに応えるというもので、皆さんと情報共有しておこうかと思うことがあります。留学生ではない、例えば生活者とか就労の方の日本語能力をどうやって現場で測ることができるのかと考えたときに、ドイツの移民難民への会話、ドイツ語能力の測定をどのようにしていращやるのかというのを、ドイツで移民難民へのドイツ語教育を担当したことのある日本人に聞いたところ、ある資料をいただきました。ドイツ語なので和訳することをやってみて、それがほぼできましたが、ドイツの移民局とゲーテ・インスティトゥートの許可がないと公開してはいけないということなので、参考までに皆様に表紙だけ共有させてもらってもいいですか。でも公開は無理ですかね。和訳は、ドイツにおける統合コースのためのプレイスメント・システムの試験官のためのハンドブックというものです。CEFRのA1, A2, B1までを測るものです。1対1で、15分以内でやる会話の試験のやり方、どういう項目をどんなふうに聞くとか、その評価方法についてもマニュアルで分かりやすく示してあるので、参考にはなるのではないかと思います。会話試験と、それと筆記試験と両方あります。こちらは2007年にできていて、ドイツでは何万人、何十万人もの人がこれを受けているので、実用性もあり、もちろん問題もあるようですが、そこも含めて調査といいますか、勉強する意味はあるのかと思っています。今後、許可が取れたら、委員の委員方はじめ、一般公開できたらと思っています。

○根岸座長

ありがとうございました。ヨーロッパの方は様々進んでいる国も多々ありますので、有益な情報がもしあ

れば、紹介していただければありがたいと思います。

それでは続いて、配布資料2の参考資料として現在準備中である配布資料3「日本語の能力判定に係る試験等一覧に掲載する情報について（案）」について検討したいと思います。事務局から資料の説明をお願いします。

○長沼委員

その前に1点、先ほどの社会的ニーズのところ一言コメントを言わせてください。順番の並び替えをした方がいいかと思います。今、話すこと、書くことなど、項目をまたいで書いてあったりするので、順序性は後で又整理するのかなと思いますが、順番を並べ替えた方がいいかと思いました。

○根岸座長

そこも含めて事務局の方で整理いただければと思います。それでは配布資料3です。

○増田日本語教育調査官

事務局から説明させていただきます。配布資料3「日本語の能力判定に係る試験等一覧に掲載する情報について（案）」を御覧ください。こちらに挙げられた項目について試験実施団体から情報を集約し、本二次報告の巻末に参考資料として収録を予定しております。

一覧のイメージとしましては、参考資料2「平成30年度日本語教育総合調査－日本語の能力評価の仕組みについて－ 概要（一部抜粋）」を御覧いただきたいと思います。これは、今期の審議の基礎資料として、平成30年度に国内外で実施されている日本語能力の評価の仕組みの実態について調査を行ったものです。1ページおめくりいただきますと、試験の一覧というべき概要が現在このような形で公開されております。平成31年以降の新たな試験も加え、現状行われている試験の一覧を二次報告と共にお示しできたらと考えております。平成30年度調査では、16の試験、評価を実施する機関に御協力を頂きましたが、現在は既に20を超えておりますので、新たな試験機関にも御協力をお願いできればと考えております。

配布資料3に戻っていただき、今期の調査項目を御確認いただけますでしょうか。「1. 試験の概要」はほぼ前回の調査と同じ試験の基礎情報となっております。今回御意見を頂きたいのが、「2. 社会で活用される試験に望まれる要素」として挙げている10項目です。テストスペック、試験の細目表が整備されているか、サンプル問題を公表しているか、得点配分、合否の判定方法の公開がされているか、試験の開発・実施に際しIRTに基づく得点等化が行われているかといった試験の内実について、現行の体制の有無を確認する一覧表を考えております。

次に、「3. 日本語教育の参照枠・CEFRのレベル尺度との対付け」となっております。これらの項目について、ご審議を頂いた上で、御協力いただける試験団体に当たって調査を行いたいと考えております。資料の説明は以上です。

○根岸座長

確認ですが、現時点の案としては、試験の基礎情報及び社会で求められる要素は有無を問うものであって、今回は立証資料などについて全て公開を前提として求めるものではないという認識でよろしいですね。

○増田日本語教育調査官

はい、そうです。

○根岸座長

分かりました。それでは、この案について委員の皆様にお諮りしたいと思います。

○長沼委員

3の日本語教育の参照枠のところですが、括弧してCEFRと書いてあって、同義と読めるような書き方がされていますが、これは大丈夫でしょうか。先ほど野口委員からの発言の趣旨を踏まえると、少し慎重な書き方をした方がいいかと思いました。

○根岸座長

本日は問題点の指摘だけするという形でもよろしいですかね。

○増田日本語教育調査官

はい、結構です。

○野口委員

今の長沼委員の御発言を受けてですが、3のところに、例えば、日本語の独自性ということについてどういう配慮をしておられるかというようなことが入るといいのかという気がいたします。CEFRだけではなくということです。難しい問題が様々あるから、一言で答えられないかもしれないのですが、日本語教育の参照枠イコールCEFRであると固定してしまうのは怖い感じがするというので、発言しました。

○根岸座長

ありがとうございます。ほかにもございますか。

一つ、1の試験の概要で、(14)に結果の通知方法というのがあると思いますが、これは、本人への結果通知と、例えば留学試験の場合は、大学への通知であったり、仕事だったら会社への通知であったりなどの方法も含んでいるのかもしれませんが、いかがでしょうか。

○増田日本語教育調査官

結果の通知方法につきましては、参考資料2の平成30年度調査の一覧にも掲載していますが、ほとんどの試験は、受験者本人に対する通知の方法を記載しております。企業や大学等に対する得点の公開・結果通知についても実情に応じて記載いただけるように配慮したいと思います。

○根岸座長

機関を通してしか先方に送れないパターンや自分でコピーを取って送るというようなパターンなど、様々あるかもしれないと思います。

○長沼委員

2に関してですが、今見ていたところ、先ほどのルーブリック評価をどうするかにもよるのですが、ルーブリック評価のような、特にパフォーマンスの部分に関する採点基準の公開や、どのように判定されているかというような質的な基準について、TOEIC SWなど公開しているものもあると思うので、そのような判定基準の公開の有無などを含めてもいいかと思いました。パフォーマンステストを実施している場合は重要な情報かと思いました。

○根岸座長

それに絡んでですが、テストの評価者のトレーニングというのはどこかに入りますか。

○増田日本語教育調査官

はい、今までは聞いてはおりません。項目には入れておりませんでした。

○根岸座長

どういう人が採点するのか、どういう人がどんなトレーニングを受けているのかなど、長沼委員のおっしゃったことについての結果のモニターをしているのかは、観点としてはあり得ると思います。

ほかには何かございますか。ありがとうございます。それでは、頂いた御意見を踏まえて修正し、調査を実施いただくようお願いいたします。

それでは、最後、全体的なことについて発言できなかったことなどがありましたら、最後にお伺いしたいと思いますが、いかがでしょうか。

それでは、本日のワーキンググループの議論はここまでといたします。最後に事務局より連絡事項等あれば、お願いいたします。

○松井日本語教育専門職

次回、第5回ワーキンググループは、1月15日金曜日、午後3時からの開催を予定しております。詳細な時間及び会場等につきましては、確定いたしましたら、改めて御連絡をさせていただきます。なお、開催通知は、別途1週間前にホームページで公開するとともに、委員の皆様にはメール又は書類にて御連絡をいたします。以上です。

○根岸座長

ありがとうございます。これで第4回日本語能力の判定基準に関するワーキンググループを閉会いたします。