

小泉利恵科研プロジェクト第2回例会（共催：日本語テスト学会）

第52回日本語テスト学会研究例会

■全体のまとめ

全体のまとめを行う時間が、全体討議で取れませんでしたので、ここにまとめます。（文責：小泉）

全体の構成

	講演 1	講演 2	講演 3	報告 1	報告 2
テストの種類	大規模テスト	授業内テスト	授業内テスト	授業内テスト	授業内テスト
主な対象層	一般	高校	高校・大学	高校	小中高大
スピーキングテストの種類	独立型・技能統合型	独立型	技能統合型	独立型が主	独立型・技能統合型

講演 1：9:45～11:00 大規模スピーキングテストにおける採点の運用と課題

講演者：井上千尋（University of Bedfordshire）

で大規模テストの運用や課題から、それをある程度解決できるものとしての校内スピーキングテストの役割が述べられました。実際に大規模テストで採点とそのトレーニングや認定がどのようにされているか、自動採点がどのような観点でなされているかの説明が詳細にあり、大規模テストの利点と弱点が浮き彫りになりました。

講演 2：11:10～12:25 信州英語プロジェクトにおける教室内スピーキングテストと採点者信頼性

講演者：酒井英樹（信州大学）

では、信州英語プロジェクトの概要と、スピーキングテストの発表とやり取りをどのように行ったか（DVDを流すことで一斉に実施。ペアを組ませて、一人ひとりにICレコーダーを持たせて録音等）、また採点者トレーニングと採点の一致度などがどうだったかの話があり、授業内でテストを45分で行うための工夫や、採点トレーニングの重要性がよく分かる内容となっていました。

講演 3：13:15～14:30 教室内技能統合型スピーキングテストにおけるループリックと採点

講演者：平井明代（筑波大学）・横内裕一郎（弘前大学）

では、前半は、教室内技能統合型スピーキングテストとして、Story Retelling Speaking Testでの再話を使った際に、採点の一致度を高めるために必要は方策が、研究結果に基づいて語られました。後半では、要約でのルーブリックや採点時の工夫が例とともにまとめられました。「読んだり聞いたりして得た情報を話す」という形式の教室内技能統合型スピーキングテストを通して、「読む・聞く行為がない状態で話す」独立型のテストにも一般化して使えるような示唆をご提示いただきました。

プロジェクト報告 1：14:40～15:10 高校の教室内スピーキングテストにおける採点に関する課題

小泉利恵（順天堂大学）・渡邊聡代（栃木県立宇都宮南高校）・初澤晋（宮城県石巻高校）・磯部礼奈（宮城県石巻高校）・松岡京一（宮城県石巻高校）

では、2つの高校での、授業内に実施し、授業内に採点を終える形のスピーキングテストの例を紹介し、その状況での採点者間の信頼性（採点の安定性）の報告がありました。詳細なトレーニングができない状況でも、ある一定の条件を満たせば信頼性がある程度は確保できるという結果は、実践時に安心感をもたらすのではないかと思います。

プロジェクト報告 2：15:10～15:40 高校の英語授業内スピーキング評価における「信頼性確保のための採点指針」の作成に向けて

小泉利恵（順天堂大学）・深澤真（琉球大学）・横内裕一郎（弘前大学）・井上千尋（University of Bedfordshire）

では、授業内スピーキングテストの実施と採点に困難をおぼえる教員が多いため、それを改善する一つの方法として、スピーキング評価に特化したウェブサイトを作る計画が語られました。既存の発話データをどのように使い、新たに何を作るかが今後の課題となります。

■当日のチャットやアンケートで寄せられたテーマに関わる感想や質疑応答をまとめます。

講演 1：9:45～11:00 大規模スピーキングテストにおける採点の運用と課題

講演者：井上千尋（University of Bedfordshire）

Q: 井上先生、大規模テストを大学入試として使うことについては、どのように考えられますか。

A: 一からスピーキングテストを作ると膨大な時間と労力、費用がかかります。その点では

すでに開発されている大規模テストを使おうという発想は理解できます。ただ、学習指導要領との関連や、受験機会の公平性などで困難になったというのも理解できます。何を重視するかの議論が必要だと思います。テスト団体側としても、それぞれのテストがどんな能力に焦点を当てているのか、という透明性を高めることも重要です。

Q: 英国には Ofqual <https://www.gov.uk/government/organisations/ofqual> がありますが、どんな役割をしているのでしょうか。

A: Ofqual (The Office of Qualifications and Examinations Regulation) は、英国内のさまざまな認定試験の統制をおこなっている政府団体です。義務教育修了試験や大学入試においては、テスト細目(設計図)を Ofqual が作成し、それに基づき複数のテスト機関がテストを開発しており、どれを選ぶかは各学校に任されています。タスクの大枠や結果の級などはテスト細目で決められているものの、具体的なトピックや配点も違うので、テスト機関によって難易度に差があると言われていました。しかし、それらが等価のものとして扱われており、日本で言う公平性とは異なる原理が動いているように見えます。

Q: このコロナ禍で、大規模実施の際の、採点者の集合研修や集合採点がかかなり難しい状況になっているかと思うのですが、もし、現状、これらのテストがどう対応されているのか、ご存知でしたら教えてください。

A: スピーキングテストの受験者自体が減っているので、すでにトレーニングを受けている採点者で十分な面はあるかもしれません。必要があれば Zoom などを使ってトレーニングをしていると思います。

Q: TOEFL iBT のベストスコアの反応はどうでしょうか? 私の教える大学では、最初に受けた時に自分が納得がいかなかった技能の点について、次に受ける前にそこだけ力を入れて勉強して、点を挙げて、Best score を高めるという学生も出てきており、Best score がその学生の全体的な力を示すかという懐疑的な意見も見られます。

A: ベストスコアの導入の背景には、「1技能だけ点が少し足りない場合、また全部受け直すことになる。受け直した試験で別の技能が低かったら、また全部受け直さなくてはならない。これを何度繰り返せというのか」という悲痛な受験者の声がありました。ですから、受験者からは歓迎されています。受け入れる大学側からしても、一番重要なのは有効期限内のスコアであることなので、今までと比べて非常に問題視されている…というわけではないようです。

Q: 大規模テストでは幅広い能力や、様々な母語の学習者を対象にしています。日本の学校での授業内評価だと、能力の幅は限られ、日本人特有の苦手な面などがあり、その意味では焦点がしやすいので、大規模テストで必要な採点トレーニングの時間が、授業内テストでは短くて済むということは考えられますか？

A: それは大いにあると思います。そこも校内スピーキングテストの強みの一つですね。

Q: Writing と Speaking の採点トレーニングで4日とあり、2技能を両方採点するというのが前提になっているということでしょうか？

A: そうですね。IELTS examiner になるとき、どちらかの技能だけに特化してトレーニングを受けるということはないとのことでした。

コメント：ルーブリックの共有は必要だが、詳細を伝えすぎると教育に悪影響 (teaching to the test, learning to the test、全力を出して頑張らない等) ということがあります。大規模テストでも、採点の詳細は採点者にしか伝えないという部分があり、同じだと思いました。

コメント：例えば Level 2 と 4 の間の記述をぼかして書いてあるルーブリックの形は参考になりました。少しだけプラスというのはよくあるため、ルーブリックにその形を使ってみるのもありだなと思いました。

コメント：内容が大規模テストでは入らず、言語に偏っていたというのは、ある程度仕方がないこととも考えられる。また自動採点も制約を認めつつ導入する方向で考えるのが良いのではないか。

コメント：採点者が2人いるときに、2人とも同じルーブリックで採点する方法が一般的だと思いますが、Cambridge 英検のように、1人が全体的な holistic 評価をして、2人目が詳細に analytic 評価をするという形もよいと思いました。授業担当者が教えた表現を使ったかの観点を見て、別な教員が他の観点を見るなど、応用もききそうです。

Q: そのような場合、トレーニングはどのように行うのでしょうか？

面接官+holistic rating の採点者：面接の方法と、採点を一緒に学ぶのでしょうか？

Analytic rating の採点者：面接の方法は学ばずに、採点のみのトレーニングを受けるのでしょうか？

A: 詳細は公表されていないので、これは私の推察にすぎないのですが、おそらく評価官

(Analytic rating の採点者) も面接の方法を学んでいると思います。というのも、評価官が「渡したイラストはあれだな、だからきっと使われる語句表現はこのあたりだな」とか「今面接官がこのくらいの難易度の質問をしているな」といったことを認識できないと、受験者の回答のレベルや扱えるトピックの抽象度などを正しく判断できないからです。また、トレーニングを行うならば、どちらの役割もできる人を育てた方が効率もいいですし、試験・採点の実施も安定して行えます。ですから、きっと両方の役割ができるように訓練を受けていると思います。

コメント：大変興味深く拝聴させていただきました。私立の小学校に勤務しているものですが、来年度は、学期に1回全学年でのスピーキングテスト(インタビューテスト)を予定しています。外国人とのインタビューテストを中心に、児童同士の会話もできるな、ということも考え、今後に活かしていきたいと考えました。また、ループリックを提示し、フィードバックも返すことなどの大切さ、複数の採点者がいることなども再認識できて、来年度への参考になりました。ありがとうございました。

コメント：大規模テストのスコアレポートは教育に活かせる情報が少ないというのはそうですが、テスト作成団体としては、その中でも教育に活かせる部分を増やそうと learning-oriented assessment という考えで進めていると思います。そこでも限界がありますが、努力はしていることは認めておきたいと思いました。

コメント：テストができることには限りがあるので、そこは認めたくえで、テストができることを測っていけばいいのではないのでしょうか。その意味では、自動採点も役立つし、それぞれのテストも役立つのではないのでしょうか。

講演 2: 11:10~12:25 信州英語プロジェクトにおける教室内スピーキングテストと採点者信頼性

講演者：酒井英樹 (信州大学)

Q: ある面では(機器の問題をクリアできれば) COVID-19 状況下でも実施しやすい形態のスピーキング・テストであるわけですが、採点者の引継ぎ・拡充も含め、2020年度の展開について報告は入っているのでしょうか。拠点校で引き続きをやっている、できていないなど。

A: 一部の中高では tablet を使って実施していると聞いています。

コメント：静岡では、信州のようなことはやらなかったのですが、各学校で作成したパフォーマンス課題の収集・分析は行っていて、そちらは県の総合教育センターに担ってもらったので、委託事業終了後も継続性があり、現在も収集・分析・共有を続けることができます。ただ、院生や学生のアルバイトに頼っていた部分がこうした委託事業の終了とともに続かなくなってしまうので、トレーニングを要する複数評価者の部分が、学校での継続的な実施にとっては課題になるかなと思いました。信州は、先生向けのワークショップも複数回開催していたので、ノウハウは残って、附属を中心に長野県の各学校で引き継いで行けるといいのかなと思います。

大学を超えて教職を目指す院生・学生の評価者組織を別に作り、おっしゃったような ALT の派遣企業などの助力も得て、アセスメントについて協同を求める学校や教委と繋ぐというような非営利団体があればいいかなと思ったりします。

Q: やり取りのテストと発表のテストだと、採点の一致が難しいのはどちらというのありましたか？私の研究では、やり取りの方が難しいという場合があります。

A: どちらもほとんど同じ一致度でした。

Q: それは、一人一人に IC レコーダーを持たせて、採点者が一人の生徒に集中して採点できたことは理由としてあると思いますか？また、ロールプレイなどを使ってどんな発話がくるかが予想しやすかったためもあるでしょうか？

A: それはあると思います。

Q: あるレベルで B になって、次の難しいレベルでも B というのは問題とおっしゃっていましたが、言語能力はそのようなものではないでしょうか？ある程度の能力の幅では、あるタスクはできたり、できなかったりで、かなり上のレベルまで伸びたところで A が取れるような感じなのは、自分のプロジェクトでも見られる傾向でした。

A: それもあると思います。

コメント：酒井先生、とても示唆に富む実践例をありがとうございました。できるだけ男女ペアで組ませて録音で区別が付きやすいようにする、というのはとても重要だなと思いながら聞いていました。

Q: どうしても男女が同数いない場合、また女子校男子校だとそれが難しくなりますね。信

州プロジェクトで、やむなく同性ペアになった生徒たちの声は、区別がつかなくて採点に難儀した、ということはありませんか？

A: ダイアログ課題は、例えば、Aさん断る役割、Bさん誘う役割というように役割を比較的明確にしたので、似たような声でも比較的問題なかったようです。ただ、声が重なってしまうと、判断しづらいことはありました。

Q: 大学生、大学院生が採点する場合と、現職教員が採点する場合には、違いはありましたか？

A: 現職派遣の院生に採点してもらいましたが、教員も大学生・院生も違いはあまりありませんでした。両者とも文法的に判断することが多かったように思います。採点トレーニングする際には、最初に、これは伝わるか、理解できるか、その後表現が分かるかという点を強調して判断してもらいました。指導にあっている先生たちが採点をする際には、指導した内容が影響したのか、大学生・大学院生よりも甘く採点してしまう傾向がありました。

コメント: 様々な制約、限界をクリアするためにたくさん工夫されているところが、とても参考になりました。

コメント: 貴重なお話をありがとうございました。クラスで一斉に行うスピーキングテストの方法とその指示の仕方までなどが具体的にわかり、大変参考になりました。

Q: 少し気になったのが、そのテストの結果を受けて学校現場では指導にどのように生かしているのかな？ということです。一括して別の機関や団体、専門家の方々に評価していただけることは現場としてはありがたいか、と思いましたが、一度手を離れてしまうと、それが現場の指導にきちんと生かされているのかな？と不安になりました。

A: 前の質問・コメントからご指摘受けました点とも重なります。初年度の2016年の時に、事業のスピーキングテストを体験していただき、生徒が楽しそうにやっていたということで、スピーキング・テストを取り入れるようになった学校（その後、iPadを使うようになった2校）があります。現場の先生方に取り入れていただくには、静岡大学で試みたような仕掛けが必要かもしれないと思いました。長野県の人事異動は、3年から4年でころころと変わりますので、その点も学校での実践が長く続かないという課題があります。

Q: 周りの声に合わせて真似したと分かった場合に何か対処をしましたか？私の研究では、そのような人を入れると他のデータと異なる動きをして（ミスフィットになって）テストとしておかしくなるので問題なければ分析から外すこともあります。

A: そのような生徒は確かにいましたが、今回は特に処理をしませんでした。特に、簡単なタスクにおいては、「真似」の扱いをどうするか、検討する必要があるかもしれません。

講演 3 : 13:15～14:30 教室内技能統合型スピーキングテストにおけるルーブリックと採点

講演者：平井明代（筑波大学）・横内裕一郎（弘前大学）

Q: 内容の評価の詳しい基準が知りたいです。分析から内容は3レベルで良さそうですね。

A: 例えば

話の筋をほとんどカバーしている

話の筋を3つ以上述べている

話の詳細を含み、その内容に誤りがほとんどない

感想が十分で適切である

を見ていく形のルーブリックを使っています。

別なルーブリックでは、

与えられた文章の必要な情報をすべて話しており、話の内容に一貫性がある

その要素が部分的に満たしている

などの形も使っています。

基本的に、主題やトピックセンテンスなどの重要なポイントを控えておいて、それをどの程度満たしているかを見ます。

Q: 要約の観点に、言い換えが入っていない理由は何ですか？ライティング評価だとこれが重要になってくると思います。

A: ライティングと違って、考える時間があまりないため、言い換えはなかなか出てこないと思います。また対象者によってはそこも見られるかもしれませんが、自分の研究の対象者では、言い換えの力までは見られないためです。

コメント：紹介されていた PeerEval の論文とアプリの URL は以下です。

http://www.juce.jp/archives/kaizen_2017/c-10.pdf

<https://apps.apple.com/jp/app/peereval/id1228030869>

コメント：生徒に書き起こしをさせたり、ピア評価をさせたりするなど、指導の中の位置づけが良くわかり、参考になりました。

Q: 技能統合型のタスクで、横内先生も言及されていた、CS（学習指導要領）に述べられているような「聞いたり読んだりしたことに基づいて、自分の考えや気持ちなどを伝える（伝え合う）」タスクは、やはり評価は難しいでしょうか。

A: 「聞いたり読んだりしたことに基づいて、自分の考えや気持ちなどを伝える（伝え合う）」ということは中学校のやりとりにある記述ですが、正直に申し上げて中学生の場合には難しいタスクだと思います。一例ですが、地球温暖化についての映像を見せる→「温暖化についてどう思いますか？話してみてください」という流れのタスクを行う…という感じになると思います。そのうえで、自分の考えや気持ちを伝える活動を行い、評価するというのであれば、比較的簡単なのではないかなと思っています。「(1) 地球温暖化について自身のスタンスを明らかにして、(2)意見や解決策を述べている」というような観点＋言語面の評価を使って、評価の段階を減らせば評価できるのかなと思います。伝え合う、となると中学生にはハードルが高くなると思いますが…

コメント：確かに、タスクそのものが複雑になってしまいますね。経験的にも、中学生にとって難しい課題なのだと理解しました。ご提案の観点も参考になります。意見を述べるようなタスクの場合、技能統合型のタスクで無理して測定するのではなく、独立型で評価したほうが良いのかもしれないと感じました。

コメント：採点の厳密な一致が大切という話は、特に採点者が 1 名の時にはそうだと思います。勉強になりました。

コメント：Opinion で高校生の評価が教員と一致する傾向があったため、Opinion など言語面でないところを生徒に自己評価・相互評価させるとよいという話でした。そのように、観点を分ける方法は有意義だと思います。ただ、言語面も自己評価・相互評価させた方が、Reflection をさせて今後の習得につなげるという意味では大切なので、それはさせておいて、採点に使うのは一部という方法もあるのではないかと思います。

コメント：発音の評価は、自由発話のスピーキングテストではあまり違いがでずに観点として入れにくいと思っていました。今回のお話で、音読、再話、要約と広げていき、そこでの発音を評価するという話があり、その形はできそうな感想を持ちました。再話・要約だとある程度自由発話的な要素がありますが、モデルがあるため、それに近づけて話せるため、そ

れができるかどうかで幅ができて、力の差も見やすいかもしれません。

プロジェクト報告 1 : 14:40~15:10 高校の教室内スピーキングテストにおける採点に関する課題

小泉利恵（順天堂大学）・渡邊聡代（栃木県立宇都宮南高校）・初澤晋（宮城県石巻高校）・磯部礼奈（宮城県石巻高校）・松岡京一（宮城県石巻高校）

Q: Group discussion での Topic はどのようなものでしたか？

A: 学校 1 では、1 回目には一般的で話しやすいトピックを選び、2 回目には教科書のトピックに関連付けたものにしました。例えば地震の話題の章の場合、地震に備えるには何をするかという感じです。

Q: グループディスカッションに参加をした生徒たちの反応はどうだったのでしょうか。参加直後の反応や採点返却後の反応を知ることができれば有難いです。

A: 生徒の反応は良かったです。テストだけれどやってよかった、充実していたという反応でした。採点返却後に行ったアンケートでも、数回のスピーキングの授業よりも 1 回の緊張感を持って行うスピーキングテストの方が成長につながったというコメントがありました。また反応をアンケートで調査しましたが、良い反応でした。

コメント: 教育的効果が高いということがわかりました。私も形成的評価の一つとして活用していきたいです。

Q: 生徒の対話型のタスクの使用が少ないのはそうだと思う。小中では特に。相手となる生徒の影響がだいぶ出てしまうということで先生との面接型になってしまうという面があるだろう。例えば、2 人での会話で 1 人の発話は決まっています、もう一人が自由回答という形（タイプ A）は、この形に入るか。質問が決まっている場合など（What sport do you like? Why? という路線が決まっている場合）

A: 広く捉えた場合は入る。生徒のやり取り型の利点は、生徒の自然な発話が入ってくること（会話を始める、終える、質問する、続かないときに助ける等）で、それはあまり見られなくはなるが、それでも生徒同士で会話する意味はたくさんあり、そこで見えるスピーキング力もあるため、その形も生徒のやり取り型として行った方がよいと考える。この形だと、採点者の信頼性は高くなり、実行可能性が高くなるのが利点で、一方妥当性は下がるというところは認識しておきたい。

Q: 制限が強いものから、弱く自由産出が多いものから、タスクは様々なものがありえる。中高だと、表現の幅がある場合もあるが、小中だと知っているものが制限されてしまうので、**What sport do you like?**でも **Tennis. Basketball.**と答えるだけでも自由度が高いということになる。自由度の取り方も、そもそもの範囲によって変わってくる。自由度と制限の理解の仕方が難しいのではないか。

コメント：決まっている表現を出すところからスピーキングテストに入れて行っていくのがよいだろう。そうでないと、あまり話せずにやる気を失うなども出てきてしまうだろう。

A: 長期的に考えて、最初は全部学んだ表現が使えるかのテストで、例えば年末のテストで少し自由度を高いものを入れるなどもできる。

コメント：生徒は「正確さ」にこだわるあまり口数が少なくなってしまうおそれがあるので、「タスク達成度」でスピーキング力を測ることができるということを教えていただいたのは私にとっては大きな学びでした。

コメント：タスク達成度だけだと正確さ的なものが大きく欠けてしまい、正確さだけだと他が欠ける。両方バランスよく入れて、重要な方に重みを大きくする方がよいのではないでしょうか。

コメント：年間計画に年数回パフォーマンステストをしっかりと位置づけていて、実際に運用していることが分かった。

プロジェクト報告 2 : 15:10~15:40 高校の英語授業内スピーキング評価における「信頼性確保のための採点指針」の作成に向けて

小泉利恵 (順天堂大学)・深澤真 (琉球大学)・横内裕一郎 (弘前大学)・井上千尋 (University of Bedfordshire)

参考にした TEAL の情報

Tools to Enhance Assessment Literacy (TEAL). (2020). TEAL for teachers of English as an additional language. <https://teal.global2.vic.edu.au/>

コメント：これに近い形でのサーバーでの共有を、委託事業の 5 年目に（第一段階としては）静岡県の中高教員向けにやろうと思っていたのですが、委託事業が 4 年で終わってしまいました。

静岡県内の全ての高校から集めたタスクについては、高等学校における英語運用能力アセスメントの現状と課題：静岡県立公立高校のパフォーマンス・タスク分析 https://shizuoka.repo.nii.ac.jp/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=11397&item_no=1&page_id=13&block_id=21 で分析しています。

コメント：既に、生徒のビデオなど集めたものはありますが、マスキングをする手間等を考えると、難しいかもしれません。肖像権、個人情報尊重したうえで行うことが必要でしょう。それもあって、ビデオ等の資料は、静岡県では教員のみがアクセスできる場所で、県内高校教員のみでの活用を基本に構想していました。

Q: 日本の先生方に非常に有用なサイトになると思って拝聴しておりました。日本では、小中に関しては、ある程度教科書の種類が限定されておりますが、教室内テストを主眼に置くと教科書に沿った形のテストのサンプルというのも先生としては利用しやすいのかと思ったのですが、いかがでしょうか？

A: 教科書に基づく利点は、その教科書を使っている先生方にはすぐに使っていただける点、指導と評価の関連が見えやすくなる点です。ただ、欠点もあり、その教科書が変わると使えなくなることがあります。今回のプロジェクトでは、20 年後も使えるウェブサイト作成を考えているため、逆に教科書に基づかない方がよいという考えもあると思います。

コメント 1: 確かに教科書は変わっていくので、おっしゃる通りですね。

コメント 2: 著作権の関係もあるので、それも考慮に入れた方がよいでしょうね。

コメント 3: 公立小学校は検定教科書を採択し、私立小学校では、その他の教材を用いていますので、教科書に基づくのは問題もありそうです。

コメント：とても興味深い研究の構想を立てられていらっしゃるな、と考えました。このような活動を教室で行おうとすると、どうしても個人で手探りで行うしかない状況だと思いますが、このようなサイトがあれば、大きな後ろ盾を得ながら自信を持って教室で取り組めるかと考えます。

コメント：高校の先生にとっては適切なタスクづくり自体が壁であると同時に、評定上の B 評価と A 評価の例のニーズがあります。

発表者：C 評価も含めてサンプルを複数提示できたらいいなと思います。

コメント：学校でスピーキングテストを実施したら、それをもっと分析したりフィードバックを準備したりすることができる時間が必要だな、と思います。

コメント：学校現場での実施状況やデータとリンクさせた取組みでよくわかった。ウェブサイトができるといいですね。

コメント：ご紹介してくださった TEAL の日本版のようなものこれから作られるというお話に、大変興味があります。大学の教員も活用できるものがあると嬉しく思います。ぜひ活用させていただきたいと思います。

■全体に向けたコメント アンケートより

- ・ speaking の評価はいつも難しいと感じていたので具体的に切り分けて考える良い機会だった。
- ・研究と実践が結びつく内容で、今回の内容を参考にさせていただき、早速来年度試してみたいと思います。
- ・教員が実際にどのようにルーブリックを評価をしているか勉強することができた。
- ・やり取りのパフォーマンステスト、授業内テストのタスク、ルーブリック、採点、フィードバック、日々の授業や教科書との関係、学校種による異同等整理していくことが多い課題だから（テーマがよかった）。
- ・英語教育の現場で指導と評価のさらなる一体化が望まれる中、とても良いテーマだったと思います。
- ・中学・高校での定期テストレベルでのスピーキングテストの実施方法とその採点方法は、中学・高校の英語の先生方からのニーズの最大の一つだと思うので、とてもよいと思います。
- ・スピーキング採点は喫緊の課題だと思います。
- ・自分の取り組みを振り返る機会となった。