

2. 主催者挨拶

梅原秀哲

EGST教育実施委員会委員長

(教育担当副学長)



ただいまご紹介いただきました教育担当の副学長 梅原です。本日は年度末のお忙しい中、当セミナーにご出席いただきましてまことにありがとうございます。名古屋工業大学では平成17年度から文部科学省現代GPにおいて採択されました『発信型国際技術者育成のための工学英語教育』に取り組んでおります。このプログラムは過去5年間で着実に効果を上げてまいりましたEGST教育(English for General Science and Technology教育)を核としまして、これを1年次、2年次の共通科目から3年次、4年次の専門科目まで体系的に発展させることでより実用性の高い英語運用能力を習得させ、そして国際社会で活躍できる技術者の育成を図ろうとする取組でございます。すなわち入学時から卒業まで工学技術英語を教育し、海外において即戦力として十分に活躍できる人材を育成しようとするものです。この取組を実施するにあたりまして、まず「名古屋工業大学EGST教育実施委員会」を立ち上げました。この委員会を通じて取組を順次検討しながら行ってまいりました。平成18年度はプログラム2年目ということもあり、昨年度は十分に取り組みなかった事業を実施しました。代表的なものが工学系プログラムを含んだ「海外語学研修」の実施、夏期休暇中と春期休暇中のネイティブ教員による「集中クラス」の開講、工学専門科目への英語化推進を目指した3年次学生対象の「実験・演習科目」における英語による指導などであります。その具体的な取組内容に関しましては本日のセミナープログラム第1部にあります「平成18年度事業成果報告会」においてそれぞれの担当者が詳しく報告しますのでそれをお聞きください。セミナー第2部としまして今回は筑波大学名誉教授の大友賢二先生に基調講演をお願いいたしました。それは現代GP取組の成果をどう評価するのかという問題意識を踏まえつつ、そもそも「評価とは何か」という、いわば基本的な視点からお話したいと願ったからです。セミナー第3部では立命館大学の清水裕子先生をお招きし、英語教育に特化する形での「評価」の問題をパネルディスカッションという形で考えることにいたしました。これは国立大学が法人化以降に認証評価を含めさまざまな形で評価にさらされているという状況を考慮にいれても重要な問題であると思っております。本日のセミナーによって本学の英語教育プログラムの推進によりいっそう弾みがつくことを大いに期待しております。簡単ではございますが挨拶とさせていただきます。よろしくお願いたします。

3 18年度EGST教育主要事業報告

I 「実験・演習科目」

英語化推進実施部会報告

奥 淳 一

名古屋工業大学院助教授



生命・物質工学教育類のEGST教育への取組を報告させていただきます。まずは17年度末のEGST教育実施委員会において「『実験・演習科目』の英語化を下さい。特に今年度は試行ですが、来年度から本格実施となります」ということが決定いたしました。そのためにどういう条件で実施するかということが各教育類に委ねられました。その際、この科目は3年次学生対象で、しかも一部の学生だけが選択する科目ではなく全員が履修する必修科目であるということ。もうひとつは、専門の内容について質疑応答さらにプレゼンテーションなどを行うこと、言い換えれば教員が一方的にしゃべるのではなく、教員と学生との間で双方向的に英語を使用するという。この2つの条件が設定されておりました。ただ問題点もございまして、3年次学生が受けている現在のカリキュラム

生命・物質工学類の3つのプログラムと 実験・演習科目（3年次）

- ・物質化学系プログラム
（前期）物質化学実験 I-III
（後期）物質化学実験 IV-VI、 卒業研究
- ・生物生命系プログラム
（前期）生物生命化学実験 I-III
（後期）生物生命化学実験 IV-VI、 卒業研究
- ・生体材料系プログラム
（前期）生体材料化学演習 I、 生体材料化学実験 I
（後期）生体材料化学演習 II、 生体材料化学実験 II、 卒業研究

実験については、安全確保の点で英語化は困難

は平成16年度からすでに実施されているものです、しかしEGST教育は平成17年度から、いわば途中から急に入ってきたプログラムです。そのため、現行カリキュラムはEGST教育について考慮した上で科目設定がなされたわけではございません。こうした問題があることを最初に指摘して

おきたいと思います。

本日は学外の方が大勢いらっしゃいますので一応簡単に説明させていただきますが、生命・物質工学教育類というのは3つのプログラムからなっております。まず「物質化学系プログラム」があり、次に「生物生命系プログラム」、もうひとつが「生体材料系プログラム」でございます。そのため、実験科目としては「物質化学系プログラム」では「物質化学実験Ⅰ～Ⅵ」、「生物生命系プログラム」では「生物生命化学実験Ⅰ～Ⅵ」、「生体材料系プログラム」では「生体材料工学実験Ⅰ、Ⅱ」があり、これが前期と後期に設定されております。しかしながらこの生命・物質工学教育類というのは化学を基本にした学科でございますので、万が一意思疎通が不十分であると非常に危険であるということがありまして、実験については英語化は困難であろうというのが学科内の議論の結論です。したがって、残りますのが「プレ卒業研究」という科目です。これは卒業研究に入る前に卒研指導の教員のうち、2名の教員の指導を受けるというものであります。この科目が「物質化学系プログラム」と「生物生命系プログラム」にございます。「生体材料系プログラム」には「プレ卒業研究」に加えて「生体材料工学演習Ⅰ、Ⅱ」というのが前期と後期にございます。したがって、この中から実際に英語化をするという科目を選ぶわけですが、最終的に選びましたのは「物質化学系プログラム」と「生物生命系プログラム」では「プレ卒業研究」を、「生体材料系プログラム」では「生体材料工学演習Ⅰ、Ⅱ」を選びました。

「プレ卒業研究」を対象科目としました物質化学系及び生物生命系のプログラムは、これを合同で実施いたしました。内容としましては英語の論文を読んで内容を発表させるというものです。3年の後期を前半と後半の2回に分けて2人の教員の指導を受けるという形で実施いたしました。実施方法を簡単に説明いたしますと、物質化学と生物生命を担当する教員を分野によって2つに分けます。学生は「物質化学プログラム」と「生物生命プログラム」のそれぞれの所属の学生がおりますので、その所属によって学生を2つに分けます。そして教員からそれぞれ5編ずつ英語論文を出してもらってそのリストを学生に提示し、その中から学生が論文をひとつ選び、それについて学生が後期のさらに半分の期間で理解し発表するというわけです。

発表方法としては7つの研究分野に分けて、それらのうちの所属の教員の前でプレゼンテーションソフトを使って作成したスライドを用いて発表することにいたしました。では何が英語化されたかと言いますと、最低限の要求として「スライドを英語で作成する」ということにいたしました。可能ならば、学生の発表と聴講者との質疑応答を英語で行うということにもいたしました。その際、私たちが重視したのは、英語が自分たちの研究生活に欠かせないものであるということを実感させるということです。そのため、細かい文法とか単語とかではなく、内容理解と表現を重視すること、言い換えれば実践的な内容を中心とすることにいたしました。一方、「生体材料工学演習Ⅰ、Ⅱ」ですが、これは前期と後期の両方で行ったわけですが、実施方法につきましては、学生を約9人のグループに分けてそのグループを教員が3週間3回ずつローテーションして指導しました。その際、学生は2ページ分ずつ担当して英語で内容紹介をした上で教員や他の学生との質疑応答をすべて英語で実施しました。担当学生が答えられない場合には他の学生のうち答えられる者が英語で回答を出すというやり方をおこないました。この際、注意したのは、とにかく最初から英語が嫌いになってくれないように困るということです。もちろん最初か

ら英語が苦手な学生もいますが、とにかく苦手意識を取り払ってもらえるように注意して、教員も分かりやすく説明し、できるかぎり敷居を低くして耳を慣れさせるということに重点をおきました。さらに回答が間違っていたとしてもとにかく話す、積極的に発言できるということの評価対象にすることにいたしました。

今回の成果としましては3つのプログラムを通して次のように考えられます。まず英語に対する抵抗が減ったということ。さらには何をやったら良いのかということがはっきりしておりますので学生が意欲的に参加したということ。これは先ほどの「生体材料系プログラム」での英語のディスカッションでもどんどん手が上がるようになって来たということですし、他の2つのプログラムでも、教員に英語の書き方などを訊きに来た学生が非常に多くございました。もうひとつは「生体材料系プログラム」でのディスカッションの際、徐々に積極的に発言できるようになってきたということです。これは学生のレベルにも依存するわけですが、それなりにだんだんできるようになってきたという成果があります。

一方、改善すべきと思われる点もございます。まず必ずしも全学生のレベルが高いレベルではない点です。基本的な英語力のある学生もいますし、授業以外ではこれまでまったく英語をしゃべったり聞いたりしたことのない学生もおります。そういったレベルの異なった学生と一緒に指導するのは非常に困難ということがございます。それからレベル分けして指導することも考えられますが、その場合レベルの高い、何でもできるような学生はいいのですが、英語はもう嫌だという学生ばかりを集めると、これは授業になりません。したがって、低学年からそういうレベルの低い学生をある程度まで持ち上げる、より一貫した英語教育がぜひ必要と思われます。学生のことばかり言いましたが、じゃあ教員はどうか、必ずしも全員が積極的に関与できたとはいえないと思います。すなわち人前でしゃべるのは苦手、英語を学生相手にしゃべるのは恥ずかしいという方が結構ありまして、先ほどの「生体材料系プログラム」でも全員が英語で指導したわけではございませんし、生物生命・物質化学の方でも全員英語でディスカッションしたわけではなく、一部の教員が学生に英語で発表させて英語でディスカッションしたということです。したがって、私の考えますのはこういった教育に関して今後専門教員と英語教員の間で密接に協力して指導していくことを考える必要があると思っております。これが1年間の試行で得られた成果の報告でございます。

I 「実験・演習科目」

英語化推進実施部会報告

佐藤 淳

名古屋工業大学大学院教授



情報工学教育類のEGST教育への取組ということでお話をさせていただきます。まず情報工学科の全体図ですが、情報工学科というのは3つの系プログラムからできておりまして、「ネットワーク系プログラム」「知能系プログラム」「メディア系プログラム」とあり、それぞれの系が演習科目を英語化するという取組を行いました。3年次学生対象の必修科目ということで、ネットワーク系とメディア系が後期に、知能系は前期に実施しました。必修科目ですので、3年生全員が必ず参加したということになります。私は「メディア系プログラム」の教員ですのでメディア系を中心に説明させていただきます。

メディア系は「メディア系演習3」ということになっておりますが、この科目がどのような科目かということを紹介しておきたいと思っております。この科目は、問題解決の過程を総合的に学習するという目的のため、1つのプロジェクトを15週間かけて実施する中で問題を発見したり、研究を企画したり、問題解決したり、それからディスカッションしたり、プレゼンテーションしたりというように、いわば問題解決の方法を総合的に磨く訓練をしたいと考えて新しく設けた科目です。それともうひとつ、チーム問題解決能力を涵養したいという意図もあります。これは、コンピュータに向かっていてもひとりひとりが個別にやっていって、協力するという姿勢がなかなか育たないということがあります。私たちはこれを非常に問題視しています。そこで学生3人でグループを形成し、その中でひとつのプロジェクトを完遂するというような、チーム問題解決能力を育てたいということです。それから少人数教育をしたいということで、教員ひとりが高々2グループしか担当しないということにいたしました。

こうした条件の中で「メディア系プログラム」ではどういう取組をしたかと言いますと、まず3本柱を建てました。ひとつは「資料の英語化」、それから「ディスカッションの英語化」、そして3番目に「プレゼンテーションの英語化」、これらを実現しようと考えました。まず「資料の英語化」ということから説明していきますと、これは学生が調査するいろいろな資料をできる限り英語化するということです。次に「ディスカッションの英語化」ということでTAになるべく英語が達者な者あるいは留学生を起用して英語によるディスカッションをスムーズに行えるようにしました。そして「プレゼンテーションの英語化」ということで、「メディア系プログラム」所属の学生60人全員が集まってそれぞれのプロジェクトを発表する合同発表会において完全に英語化することにいたしました。学生の発表資料は全部英語で作らせ、それを英語で口頭発表させた上で最後の質疑応答も全部英語で行う、こういうことを課したわけです。学生にとってはほとんど初めてみんな

の前で発表する、しかもそれが英語だということでもかなりプレッシャーではあったと思いますが、皆かなりうまくやってくれたという感触です。

メディア系演習IIIの企画

- 問題解決の過程を総合的に学習
 - 1 プロジェクトを16週間かけて実施する中で、問題発掘、企画立案、問題解決、ディスカッション、プレゼンテーション等を総合的に学習
- チーム問題解決能力の涵養
 - 学生3人で1グループを形成し、グループ内で協力して1プロジェクトを完成
- 少人数教育による密度の高い指導
 - 1教員は高々2グループ(6人)を担当
- 専門を英語で学ぶ英語空間作り
 1. 資料の英語化
 2. ディスカッションの英語化
 3. プレゼンテーションの英語化

| | | | | | |
|----------|---------|--|-------------------|---------|----------|
| ガイダンス(1) | 知財教育(1) | 資料調査 テーマ検討 計画立案 役割分担 (2-4) | プロジェクト実施 (8-9) | 発表準備(2) | 合同発表会(1) |
|----------|---------|--|-------------------|---------|----------|

英語化推進にあたって工夫した点をあげますと、学生と教員との間で密な英語空間を築くようにしましょう。そのため1人の教員にあたり学生を最大6人とした、先ほど紹介した通り2グループ6人ですね、これしか担当しないということで学生と教員が面と向かって英語で話し合うそういう空間を作ったということですね。それから英語による説明能力や議論能力の関与ということで、英語能力といってもいろいろあるわけですが、ここで関与すべきなのは説明できたりする能力を磨くべきだろうということで、ディスカッションやプレゼンテーションというところでの英語化を重視したということです。それから専門教育レベルの低下や安全性に対する配慮ということで、当初、教員は、英語化することで専門レベルが落ちるのではないかあるいは高度なことが教えられなくなるのではないか、それから演習中の安全性確保ができないのではないか、この2つを全員が気にしました。そのため、こういうことがないようにと、演習中の重要な説明等に関しては、なるべく英語化は避けました。したがって、教員の説明するところよりはむしろディスカッションやプレゼンテーションの方で英語化を行うという工夫をしたということです。それから学生ごとの偏りがないように注意しました。チームで発表するとなるとどうしても英語が達者な学生だけが発表するということになりがちですが、3人が均等に発表することを義務付けました。それから英語化推進に対する取組の促進ということでちょっとした工夫をしました。それはプレゼンテーションのときに、賞を出すということにしまして、優秀賞を3点出しました。そのお蔭か、学生の目の色もだいぶ変わった感じでかなり真剣に取り

組んでいました。この科目は半期 15 週の科目ですが皆 18 週、19 週はやっていたように思います。教員もそれに付き合いました。全体的にかなり熱心に取り組んでいました。

資料はできる限り英語化しましたと、それから学生が調査に使う資料もできる限り英語化しました。それからディスカッションの英語化に関しては、3人でひとつのプロジェクトを形成するのですが、毎回ある一定時間、20分か30分程度、英語によるディスカッションタイムというのを実施しました。ディスカッションタイムでは英語しか使えませんよとした上で、プロジェクトの進捗状況や今後の方向性等に関し教員との間で、あるいは学生同士でディスカッションを行いました。学生は自分から何も言わないかなと思ったら意外と発言していました。それからプレゼンテーションの英語化ということですが、先ほどの言いましたように、最終回に学生 60 人 19 チームがすべて集まって合同発表会を実施しました。各グループは、発表時間が 5 分間、質疑応答の時間が 3 分間、合計して 8 分間という時間枠の中で発表いたしました。その際、グループ内の 3 人が分担して発表することを義務付けました。ということは、全員が 2 分程度は英語でしゃべるということになります。時には教員から質疑応答の時間に厳しく突っ込まれますので、汗を掻きながら必死に回答する場面もありました。最後に表彰式をして 3 グループに優秀賞を出しました。

英語化推進で成果だと感じたことをまとめてみますと、ひとつは、学生それぞれが不十分であっても曲がりなりに英語でディスカッションし、英語で発表しては質疑応答をしたということです。こういう経験をしたことで英語で専門分野に立ち向かうという姿勢が芽生えたように思います。これは非常に大きな成果だと思います。これに関してはすべての教員がそう感じたはずです。次に、英語による発表という大舞台を経験したことで学生が自信を持ったということがあります。また教員側の英語化に対するアレルギーも少なからず改善されたと思われます。取り組む前は本当にうまくいくのかなと思っていたところもあったし、実際そう口にする教員も結構いましたけれどもやってみたところ予想以上にうまくいったというのがほとんどの教員の感想です。こうしたことから英語化に対するアレルギーはかなり減ったと思っています。

最後に、今回の取組のまとめと今度の課題ということですが、今回実行してみて予想以上にうまくいったというのが私の感想ですし、「メディア系プログラム」のほとんどの教員の感想です。部分的に英語空間を構築したことによりかなり無理なく入ったという感じがあります。無理なく英語空間が設定できて英語で専門を学ぶ、そういうことに対する精神面での教育効果が非常に高かったと皆感じております。当初は英語化による専門学習効果の低下とか安全性の問題という点が心配だったのですが、部分的に英語化することによってそれらの問題は結果的には生じなかったと思われます。課題としては、今回少人数で実施しました。少人数でやると英語空間構築のためには非常に良いのですが、ひとつの欠点は英語化のレベルが教員の中でかなり違っているという点です。これは多少感じられました。今後は教員間の英語化のレベルをいかに揃えていくかということが課題だと思っています。

Ⅱ e-Learning 教育実施部会報告

中野智文

名古屋工業大学情報基盤センター助手



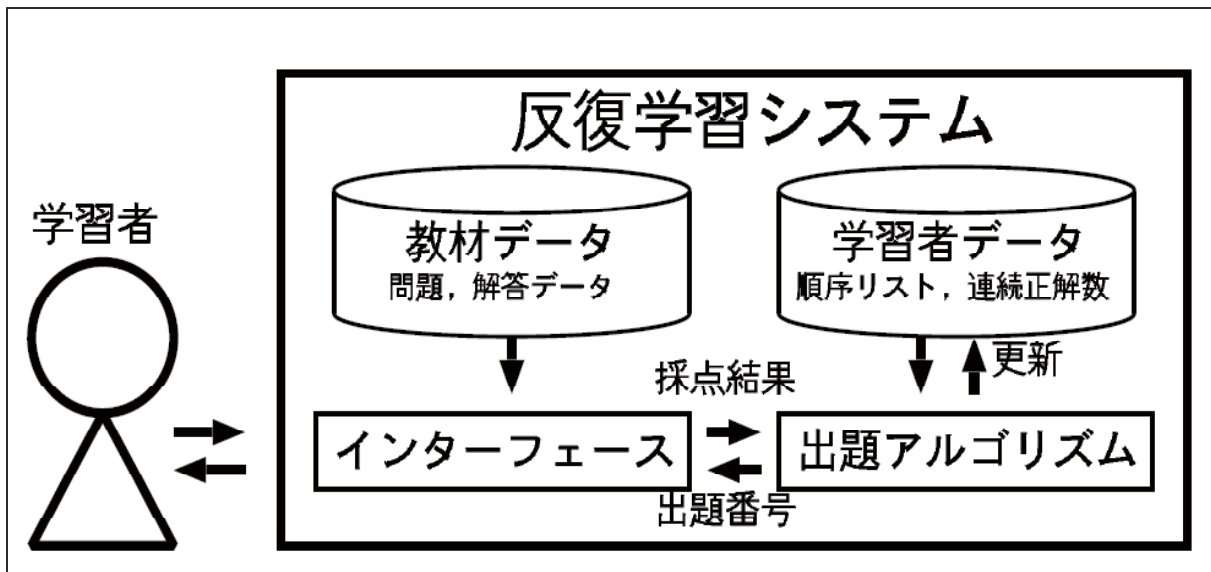
本学の e-learning 教育には大きく三つの取組があります。今回はそれらをご紹介したいと思います。ひとつは単語学習、次に読解支援、最後に作文支援です。最後の作文支援は実際にはまだ学生諸君には直接やってもらってはいないのですが、来年度から実施したいと考えておりますので今回ご紹介したいと思います。

従来の一般的な単語の学習のシステムとしては、たとえばグループ 10 問とか 20 問くらいを一度に出題し、それを順番に解答して行き、その際誤答があればその誤ったところだけ復習するというものだったと思います。これは、学習間隔が長い方から短い方に移っていく形になっています。しかし学習というのは短期記憶から長期記憶へもっていくのが理想ではないかと思えます。そのため本学の取組ではこうした形でのグループ分けがしてあります。その際、分散学習アルゴリズムというものを提案しました。このアルゴリズムでは正解すればするほど同一単語の出題間隔が長くなって理想的な短期記憶から長期記憶へ学習ができるのではないかと考えております。ただ実際にやってもらいますとすでに知っている単語が何度でも出てくるのが非常にしつこいと言われました。そこでいろいろな改良点を加えました、またうっかり間違えた場合でも連続正解数が 0 になってしまい、これも困るということですので、過去の平均に基づいて連続正解数を 2 にするといった改良を加えました。現在このシステムは Web と携帯電話のほとんどの機種でうまく動くことを確認しております。また、コンピュータでやった時もモバイルでやった時も同じ記録を使っておりますので、大学のコンピュータで学習し途中で家に帰って携帯電話でやってそしてまた翌日大学で勉強するといった使い方もできます。実際いまは 800 人の方が使っております、約 25 万回のアクセスがありました。さらには学生アンケートを採ったところ 90% の人が効果ありと回答していました。今後の課題としましては、問題を作成するというのが非常に大変でして、システムができてはもう一度はどうかしたら良い英語の問題を作れるのか、今このところを頑張っているところなんです。

次に読解支援システムです。読解支援システムでは品詞表示のシステムを作りました。そこでまず初めに品詞表示の効果について説明したいと思います。調査によって、工学系の学生はまず読むための英語の学習が必要だということが分かっています。しかし長文や複雑な文章の理解が困難であるという結果も出ています。そこで品詞がもし表示されたらより容易に理解できるのではないかと考え、そのようなシステムを考案しました。現在本学の 1 年次、2 年次学生対象に使われている独自の教科書が二つあるのですが、それらの教科書に掲載しているリーディング用英文すべてをこのようなシステムの中で見ることが

できます。

次に作文支援システムの話を行います。ただ時間があまりありませんので今回は新たに提案した補完システムについてのみ説明させていただきます。これは、文章作成する途中で、ある単語にはこのような繋がりがあるということを示してくれるシステム、それをここでは「共起」と呼びますが、その共起を自動的に出してくれるシステムです。たとえば今これから文章作成しようとしています。まず普通は We くらいから始まりますね、その後、仮に suggest というボタンを押すとこの隣に来そうな単語というものを自動的に選んでくれるというシステムです。このような形で単語をどんどん選んでいけば適当に文章は作成出来るというわけです。ただ今はまだ精度があんまり良くないので使い物になりませんが、今後このようなインターフェースを考えて作文支援のシステムを作りたいと考えています。



単純分散学習アルゴリズム

Input and Output:

- O # 出題順リスト
- p # 連続正解数の配列

begin

$q \leftarrow O$ の先頭要素を取り出す (シフト)

q 番目の問題を出題

if 正解 then

$$p[q] = p[q] + 1$$

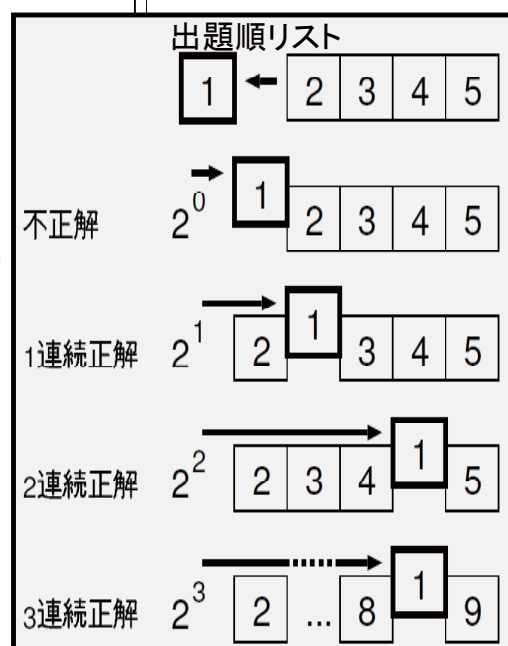
else

$$p[q] = 0$$

endif

O の $2^{p[q]}$ 番目の前に q を挿入

end



Ⅲ 海外語学研修実施部会報告

永井正司

名古屋工業大学大学院助教授



海外語学研修に関し報告させていただきます。この取組には下の【図1】にありますように特色が四つあります。

【図1】

海外語学研修

1. 専門性(工学)を重視した語学研修
2. 1年次・2年次のEGST教育を踏まえた海外体験
3. 企業視察
4. 単位認定(選択科目)

次に平成18年度に試行的に行った研修のご報告をいたします。【図2】をご覧ください。

【図2】

平成18年度の試行

1. 平成18年8月11日ー9月12日
2. マンチェスター大学(連合王国)・語学センターにおける4週間のプログラム
3. 帰途のシンガポール滞在・企業視察

私たちはこの取組のため「海外語学研修実施部会」を設置し、この部会の教員と研修先のマンチェスター大学・語学センター教員およびマンチェスター大学・工学部教員、この三者で綿密に打ち合わせ

せて今回の研修を実施いたしました。さらに今回シンガポール航空を使用しましたので、シンガポールにて企業視察を行うということにいたしました。さらに今年度の研修プログラムでこういった工学的内容の授業を行ったと言いますと、いずれも現地のマンチェスター大学の先生にお願いして、アイスクリームを作る実験とか3Dの実験をしたとか、あるい

はレーザー機器で触れてみるとか、さらにはマンチェスターには「化学産業博物館」という世界的に評価の高い博物館があるのですがそこを視察する、そういったことを行いました。【図3】をご参照下さい。

【図3】

研修プログラムの工学的内容

- Making Ice-cream
- 3D-Visualisation
- Fun with Lasers
- Visit to Museum of Science and Industry, etc.

19年度からこの取組は本格的に実施するということです。日程その他をすでに決定し、学内説明会もすでに一回行っております。日程的には8月2日から9月2日まで、研修先は同じです。ちなみにマンチェス

ターというのは産業革命の発祥地でもありますので非常に産業の集積度の高い都市です。また現地の企業の視察に関しましては今のところ交渉中ではありますがもある大きな製作所の現地子会社を予定しております。

なお来年度からの本格実施に際し、いくつかの変更点があります。それは参加資格の厳密化ということです。今年度の場合、試行ということですので参加資格を特に厳密に設けておりませんでした。そのため、学部の1年生でも参加できますし大学院生博士課程の後期3年生でも参加できました。しかしこれでは先ほど特色として申し上げた1年、2年次のEGST教育を踏まえた海外語学研修ということとちょっと齟齬がありますので、来年は学部3年生を対象とし、しかもある程度の英語力を身につけた学生に行ってもらおうということで、TOEIC500点以上を参加資格としました。さらに事前研修を充実させようと考えております。それから来年度はこの研修を単位認定することができるようになりましたので、3年次学生対象の選択科目「科学技術英語Ⅲ」を単位認定いたします。もちろん、自動的に認定されるわけではなく、研修中にしっかり学習を行ったという証明書をもった上で帰国後に本学英語教員と面接等を行って単位の認定をするかどうかを決定するという手順です。

IV 集中クラス実施部会報告

永井正司

名古屋工業大学大学院助教授

集中クラス実施に関し報告させていただきます。集中クラスについては【図1】にあります通り、いくつかの特色があります。

【図1】

集中クラス

1. ネイティブ・スピーカーによる少人数の集中型クラス
2. 外国語演習 I (1年次選択科目)の特別クラスとして設定
3. オリジナルの教科書で英語4技能の実践的訓練

これらに関し、もう少し詳しく説明いたします。この集中クラスは、英語力の高い者をさらに伸ばすという趣旨でやっておりますので、2つのレベルがありまして、レベルAがTOEICで550点以上獲得した学生を対象に3クラス25名程度

のクラスの定員で行っております。レベルBはTOEIC500点以上獲得した学生4クラス25名程度のクラスで行っております。今年度は、去年の夏8月7日から11日までの一週間、毎日午前10時半から12時まで、その後1時間昼休みが入った後、再び午後1時から2時半までと2時40分まで4時10分までの3コマを連日行いました。春季も同じスケジュールで、3月上旬に一週間実施しました。その内容に関しては【図2】をご参照下さい。

集中クラスは全クラス、ネイティブスピーカーの先生が担当し、1日1トピックについて、この【図2】にあります通り、たとえば午前がリーディングとライティングとし、午後がリスニングとスピーキング、その後、ライティングと文法を行います。

次に1年次に集中クラスを受講した学生を後でどのようにケアするかということですが、【図3】にあります通り、1年次に集中クラスを受講した学生を対象として2年次生対象の必修科目「科学技術英語Ⅱ」にも特別クラスを設けまして、1年次の集中クラス

の受講と併せて英語力のいっそうの伸長を図ることにいたしました。

【図2】

集中クラスの内容

- 1日1トピック

Architecture

Natural Disasters

Transportation, etc.

Reading/Writing, Listening/Speaking.

Writing/Grammar 各90分

【図3】

2年次以降

- ◆ 1年次集中クラスを受講した者を対象として科学技術英語Ⅱ(必修科目)の特別クラス

英語力の一層の伸長

課題学習(英文多読)

3. 基調講演

＜目標に準拠した＞評価の課題

大友賢二

筑波大学名誉教授



発表要旨

わが国の学校教育における英語学力の測定と評価の現状を検討し、その課題を見いだそう、というのがこの講演の目的です。そのために、言語テストの視点から「目標に準拠した評価」に関連する以下の3つの事項について考えてみることにします。

(1) 学校における「評価」の現状：わが国においては、「指導要録」、「観点別学習状況の評価」、「規準と基準」など、評価について検討すべきことが多岐にわたっています。そのうち、測定結果の解釈の方法で分類することができる2つの考え方を取り上げます。「目標基準準拠テスト」(criterion-referenced test: CRT)と呼ばれているものがその1つです。それに対し、「集団規準準拠テスト」(norm-referenced test: NRT)というものがあります。これを検討することから始めます。また、現に起こっているいくつかの現象をとりあげます。

(2) テスト・データと言語能力：「英語の何を測定するのか？」に関しては、学校教育では、4つの観点が示されています。しかし、この示し方は妥当であるかという議論も少なくありません。測定すべき言語能力の枠組みをどう捕らえるのがよいのでしょうか。最近、Can-Do Statementなどを通して、テスト・データと言語能力の関係を考える動きもあります。これに関連したテスト問題作成に必要な「テスト細目」(test specifications)の課題も見逃すことはできません。

(3) 分割点設定と項目応答理論：CRTにおける重要な課題のひとつは、設定した目標の到達・未到達はどうすれば分かるかということです。いわゆる、「分割点の設定」(standard setting)に関する課題です。たとえば、目標到達基準が80点と設定されているとしても、その恣意的な設定手順に対する疑問は多いようです。この分野の研究で、目下、注目を浴びている「しおり推定法」(Bookmark Method)に関して、「項目応答理論」(item response theory)の利用も含めて考えてみることにします。

参考文献

大友賢二 (1996) 『項目応答理論入門：言語テストデータの新しい分析法』 大修館書店

池田 央・大友賢二監修 (1997) 『言語テスト法の基礎』 みくに出版

大友賢二・R. スラッシャー監訳 (2000) 『＜実践＞言語テスト作成法』 大修館書店

Cizek, Bunch and Koons (2004) Setting Performance Standards: Contemporary

Methods, Educational Measurement: Issues and Practice, Vol.23, No.4, NCME

皆さん、こんにちは。わたくしの話は言語テストに関する2, 3の問題提起から始めることとします。

第1は、「英語35点」ということについてです。英語が35点というのは、非常に悪い点数ですが、それを取っている学生がもしいるとすれば、それはその学生の能力が低いので35点なのか、それともテスト問題が難しいので35点となったのでしょうか？どちらが原因なのでしょうか。お分かりでしょうか？難しい質問です。この課題はどうすれば解決できるのでしょうか？

第2は、「成績の伸び」に関する問題提起です。成績が伸びるとか伸びないというのがあります。たとえば同じテスト問題を学年始めと学年末に実施し、その成績を見ると、向上しているのが普通です。しかし、最初に良い成績を取っている学生は成績があまり伸びないようです。なぜ伸びないのでしょうか？油断しているからでしょうか？いや、成績はほんとうに伸びていないのでしょうか？これが第2の問題提起です。

第3は「テスト項目困難度」についてです。試験問題が難しいとか易しいというのがよくありますが、テスト項目の難しい度合は何で決まるのでしょうか？受験者集団の能力によるのでしょうか？受験者集団が優れた学生で構成されていれば、そのテストに正解するものが多く、そのテストは易しいと見なされます。これに対し、同じテスト問題を、あまり優れていない受験者集団に対して与えれば、正解者は少なく、そのテストは難しいと見なされます。これでは、そのテストの本当の項目困難度は求められないのではないのでしょうか？

第4番目は、「英語能力」ということです。受験するテストによってその学生の点数が変わってくるのが普通です。テストが易しい場合には、ほとんどの受験者は高い得点を取り、したがって、その受験者は高い能力の持ち主となります。しかし、テストが難しい場合には、ほとんどの受験者は低い得点を取り、したがって、その受験者は低い能力の持ち主となります。そこで、問題は、受験者の真の能力は、どのテストの場合の能力を指すのでしょうか？

以上のいくつかの問題提起に対する解答を考えながら、これからの私の話をお聞きください。今回の講演題目「目標に準拠した評価」では、柱としては以下の3本を考えました。

- (1) 学校における「評価」の現状
- (2) テストデータと言語能力
- (3) 分割点設定と項目応答理論

まず、(1)の「学校における「評価」の現状」から話しましょう。このセミナーに中学校あるいは高校の先生もいらっしゃるのですが、お分かりかと思いますが、「観点別学習状況」において何を評価するかという問題については、今のところ、文部科学省では4本の柱としてあげております。(1)コミュニケーションへの関心・意欲・態度、(2)表現の能力、(3)理解の能力、そして(4)言語や文化についての知識・理解です。これらを観点別にA,B,Cで評価し、さらにそれをまとめて5段階評定という形で評価することになっています。それから「規準」と「基準」とはどう違うのかという問題がよくあるのですが、この理解にも、少し、混乱があるようです。簡単に言いますと、「規準」のほうは達成目標であって、「基

準」というのは実現状況を評価するための指数、分割点であると思います。質的な評価目標をクライテリオン、そのクライテリオンの達成度を示す量的なものをスタンダードとし、前者のクライテリオンは「規準」を、後者の量的なスタンダードには「基準」の語をあてて区別することとするのがいいのではないのでしょうか。

それから、用語の問題ですが、相対評価に対して、今度は、絶対評価だといわれています。この中で、相対評価は、教育測定の世界では「集団基準準拠テスト」(NRT: norm-referenced test) といわれているものです。現在UCLA教授のBachman(2004)から引用すれば以下の通りです。

In a norm-referenced interpretation, an individual's score is compared, in terms of its relative standing, with those of other individuals in a group that has taken the test.

これに対して現在学校教育で行われているのが、いわゆる絶対評価です。これを、私は「目標基準準拠テスト」(CRT: criterion-referenced test) と言っております。これも簡単に言うと、

In a criterion-referenced score interpretation, an individual's score is interpreted with reference to a pre-determined criterion that is independent of the way test takers perform on the test.

これらを簡単に比較しますと、項目の型、妥当性・信頼性の重要度、テスト作成の原則(除く困難度)など、共通点も随分あるわけです。その一方で、違っているところを指摘するとすれば、NRTの場合は受験者の中の「弁別」を中心に考えています。しかしCRTの場合には受験者の「記述」を中心とします。それから、問題の難しさに関して言えば、NRTの場合は中ほどの困難度をもった問題を考えるのに対し、CRTの場合には非常にやさしい問題でも難しい問題でも構わないという考え方です。一般的には、NRTの場合は受験者集団の明確な記述が必要であるのに対し、CRTの場合には学習領域の明確な記述が必要になってきます。今、そういう形で実際に絶対評価が行われているのですが、どうもそれがうまく機能していないというような声もたくさんございます。たとえば、AERA No. 22, 2006, 5月15号に示されているように、評定「5」が持っているほんとうの学力は、学校によって、かなりの差があるということなどです。

次に(2)「テストデータと言語能力」ということですが、テストにおける言語能力の見方というのは、今までいろいろあるわけです。先ほど申しました「観点別学習状況の評価」は4つ柱を立てて言語能力を見ようというわけです。「コミュニケーションへの関心・意欲・態度」、「表現の能力」、「理解の能力」、「言語の文化についての知識、理解」というものです。この立て方が本当に良かったのかという声もあります。どうも、外国語とは関係ない他教科との関連で足並みをそろえている感じがあります。したがって、言語テスト分野特有の柱を立てる必要があるのではないかと考えます。

言語テスト特有の柱を考えるための、そのひとつの例としては「技能と要素に基づくモデル」がありますが、これは、当時 Georgetown 大学教授 Robert Lado らの考え方。それから「語用論に基づくモデル」、これは当時 New Mexico 大学教授 John Oller らが推奨したことですが、それはどう捕らえたら良いのでしょうか？あるいは Canale and Swain (1980) の「意思伝達能力に基づくモデル」による文法能力とか社会言語学的能力あるいは談話能力や方略的能力といったようなもの。それからごく最近では Bachman & Palmer (1996) で言いだしている「構造的知識や語用論的知識」、こういったようなものが今までいわゆるテストの中で何を測ったらよいかということについての議論がなされています。こういうものを参考にしてもう一度言語テスト特有の柱を考えておく必要があるのではないかと思います。

最近のテストの動きとして「CAN-DO-STATEMENT」というのがあります。これはテストの点数だけで分からないので、それをいわゆる言語能力と結びつけようといった動きです。これは、非常に意味があることだろうと思いますが、以下のようにさまざまな努力がなされています。

- (1) TOEFL iBT Overall Language Competency Descriptors
- (2) TOEIC Speaking/Writing Test Proficiency Level Descriptors
- (3) Common European Framework of Reference for Language: illustrative descriptors, common reference level, global scale
- (4) 日本英語検定協会： 英検 CAN-DO リスト

測定目標設定を考慮に入れたテスト項目作成で、少し見落としていた事があるのではないかと思います。それは、「テスト細目」(test specifications)の要素の構成に関する事です。これは、テスト問題作成の基盤になるものです。当時 UCLA 教授の Popham、現 Illinois 大学教授の Davidson らが述べている「テスト細目」の構成要素としては、General Description, Prompt attributes, Response attribute, Sample items, Specification supplement をあげております。私たちが見落としていたのは、この中の prompt attribute ということだろうと思います。この prompt attribute というのは「what will be given to the test taker」というものですから、どういう条件があってどういうことが受験者に与えられた場合に何ができるか、この辺を明確に示せる目標設定の下で試験問題の作成を検討すべきではないかと考えます。単に「・・・ができる」というのではなく、「どんな事が受験者に与えられたときに、・・・ができる」という要素を、目標設定や、テスト問題作成の基盤として、考慮に入れることが必要であろうと思います。

さて、3番目の柱に移ります。(3)「分割点設定と項目応答理論」という大きな課題があります。

いったい目標到達や未到達をどこでどう決めるかという点に関する研究がいままで日本ではあまり行われてきていないようです。外国ではかなり進んでいるようです。目標到

達・未到達の方法で最も簡単なことは、当時 Pittsburgh 大学教授の Nitko によって “feet on the desk” procedure といわれていたものです。これは「デタラメな、いい加減な」方法ということです。たとえば 80 点を取ったら合格とするということです。しかし、テストが違ったらどうするのでしょうか？ 80 点というのはどこから出てきたのかわからない。いや、もっとゆるくして 70 点を取ったら合格としましょう。しかしその根拠がよくわからない。勝手に設定していることが多いのではないかと言う批判です。それから、最終的な分割点を決定するときにはどんな研究においても複数の方法を利用する必要があるんじゃないかということを言っている方もおります。このように、分割点の設定は非常に難しいようです。日本ではほとんど立ち入った研究がなされていない分野です。しかも、分割点の設定に関していろいろ進めていく場合には、やはり従来の古典的テスト理論ではとても駄目だろうと私は思います。

ではどうして駄目なのか、その理由を見つけましょう。まず「成績の伸び」ということを考えて見ます。成績が良い学生の伸びは見かけ上小さい。非常にいい点数を取ったものはあまり伸びないようですが、そうでしょうか？ 実際は伸びているだろうと考えられます。ただ、それが目に見えないのです。何故かというと、満点や零点の近くの得点の歪みが生じているのですが、素点ではそれが目に見えないのです。そして、得点の歪みが生ずるいわゆる天井効果や床面効果をなくす必要があります。天井というのは 100 点ということです。100 点という天井があるために 130 点の力を持っていても 100 点にされる。150 点の力をもって 100 点にされる。そうするとその 100 点の近辺の物差しが歪んでしまう現象です。これがいわゆる天井効果と呼ばれるものです。そこで、100 点の近辺でも歪まない物差しを作ろうという考えがあるわけです。それが項目応答理論の重要な視点の 1 つです。具体的には、得点の歪みを取り除く LOGIT CORRECT(略称：LC) と呼ばれているものです。

$LOGIT\ CORRECT = LN(p/(1-p))$ ここで、LN は natural logarithm (自然対数)、p は、(NRS (number-right score: 正答数に基づく得点) / MS (maximum score: 満点))。たとえば、 $LN(3)$ というのは、 $e(2.718)$ を何乗したら 3 になるかを問うものです。その答えは、1.099 となります。

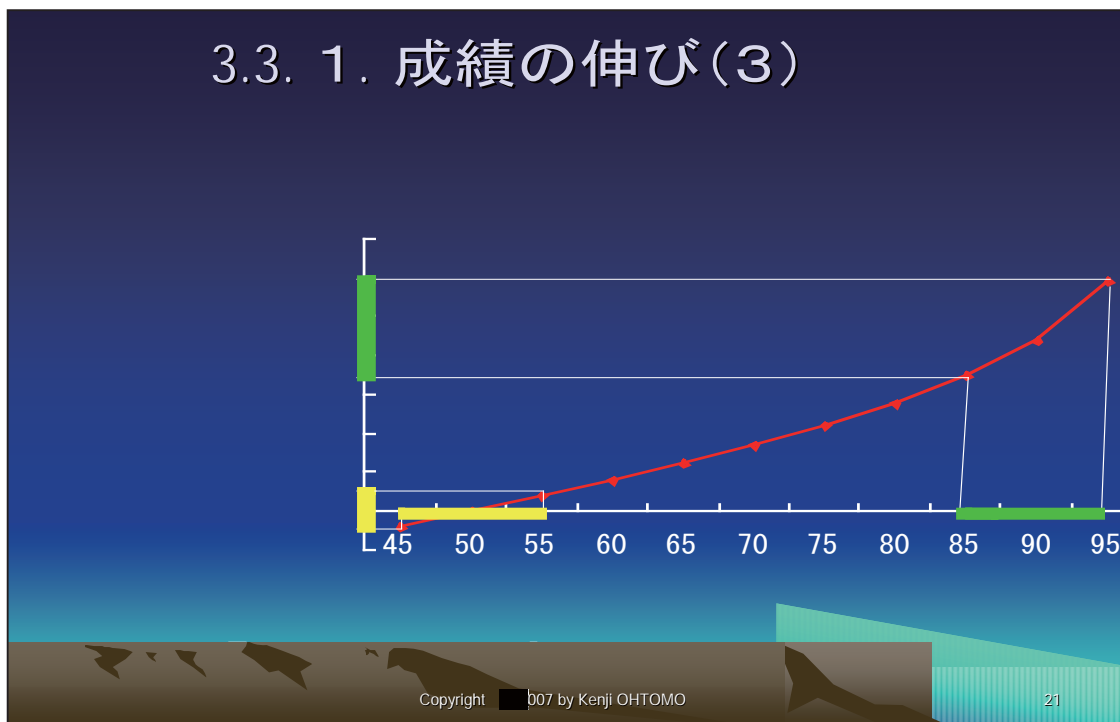
| 素点 | LC | 素点 | LC |
|----|--------|----|-------|
| 55 | 0.201 | 95 | 2.944 |
| 45 | -0.201 | 85 | 1.735 |
| 10 | 0.402 | 10 | 1.209 |

たとえば素点が 45 点から 55 点まで向上した場合には、素点では 10 点の向上ですが、歪まない尺度 (LC) でその向上を見ると 0.402 となります。これに対し、85 点から 95 点まで上がった場合、素点ではこれも 10 点上がっているわけですが、この 10 点は LC では、1.209 となります。つまり、素点では同じ 10 点の伸びでも、満点に行くにしたがって、歪まない尺度 (LC) での価値は高いというわけです。これをグラフで表すと以下の図よ

うになります。

下の図でお分りの通り、横線は素点を、縦線は LOGIT CORRECT を示しています。ここで、素点の 45 点から 55 点までは、LOGIT CORRECT で見直せば、黄色で示されている縦軸のようになります。同じように 10 点でも 85 点から 95 点までだと縦軸の緑色の部分で、黄色の 3 倍くらいの長さになっています。だから 85 点や 95 点の近辺で少し

3.3. 1. 成績の伸び(3)



か上がらないように見えますが、実際には、45 点から 55 点までの間のほぼ 3 倍の向上・価値があることがわかります。たとえば、満点近くでの素点 2 点の上昇は、中頃の学生では、6 点ぐらいの上昇を意味しているわけです。最初に成績がよかったものは、成績があまり向上しないということではありません。ほんとうは、かなり向上しているのですが、素点では、それが見えないということです。

それから「英語 35 点」というのは受験者の能力が低いために 35 点になったのか、テスト問題が難しいから 35 点になったのか、よくわかりません。この解明は、なかなか難しいものです。しかし、この問題が「項目応答理論 (item response theory : IRT)」の重要なポイントの一つになっております。35 点というのは能力が低いから 35 点なのか、それとも問題が難しいから 35 点なのか、これが混同しているのですね。それが分かるように困難度と能力を分離してみるという発想、これも IRT の非常に重要なポイントのひとつです。簡単にその IRT のモデルを書くとしたら、「 $P = \theta - b$ 」となります。P というのは正答確率、 θ というのは受験者の能力、b というのは問題の難しさの度合いです。受験者の能力がテスト問題の難しさよりも大きければ P いわゆる正答確率が高くなります。しかし、反対に b の方が大きければ、受験者能力よりもテスト項目困難度の方が大きいので、正答確率が低くなるというわけです。また古典的テスト理論では受験者集団が異なればテスト項目の困難度が違ってきます。IRT ではそれをなくそうと考えたのです。そして、どんな受験者集団に対して実施したテストでも、一定不変の困難度を求めることができる

ようにしたわけです。さらには、受験者の能力についても同じですね。難しいテストを受験すれば正答数は少なく、能力が低いと判断されるのですが、簡単なテストを受験すれば能力が高いと判断される。このように、古典的テスト理論では、どの場合が、本当の能力かわらなかった。それをIRTではどちらの試験を受けても一定の能力を発見できるように工夫をしたわけです。

「目標に準拠した評価」のための実施で、きわめて重要なことは、以上述べたような、(1) 成績の伸び、(2) 英語35点、(3) 項目困難度、(4) 英語能力などに関する適切な処理が可能な理論と実践を積み上げていくことであろうと考えます。つまり、新しいテスト理論：項目応答理論の活用は、ぜひ、必要になってきています。

それでは、「目標に準拠した評価」において、もっとも重要な課題である「基準」の設定について話をすすめたいと思います。現在、中学校や高等学校で行われている学習目標達成の状況：「十分満足」、「おおむね満足」「努力を要す」という基準の設定は、大学においても、十分考えなければならない課題です。この設定では、いわゆる、目標到達か否かの「分割点 (cut score) の設定」を考えなければなりません。今までこの分割点の設定にはどんな方法が考えられたかと言いますと、ひとつはテスト中心のモデルで分割点を決定しようという方法です。これには、Angoff法、Ebel法、Jaegar法などがあります。もうひとつは受験者中心のモデルでこれを決定しようという考え方です。これには、境界線グループ法、グループ対照法などがあります。しかし、今日は、分割点の設定で、最近、非常に注目をあびてきている「Bookmark Method (しおり推定法)」という方法を簡単に紹介したいと思います。

Bookmark Method (しおり推定法) の概要はこのようなものです。現在、(1) 米国<50州>のうち31州が、このBookmark Methodで、教育の成績に関する分割点 (cut score) の設定を行っているということです。(2) 開発者は、CTB/McGraw-Hillの研究者であるLewis, Mitzel and Green (1996) などです。(3) なぜBookmark Methodと呼ばれているかということ、分割点推定者が、「テスト項目冊子」に「しおり」を置くことによって、成績の分割点を定めることからそう呼ばれるようになったようです。

Bookmark Methodの特徴は、まとめると、5つぐらいになります。(1)は、他の設定法ではこれまで見られなかった新しいテスト理論「項目応答理論」を利用しているということです。第(2)は、複数の分割点を設定することができるということです。第(3)は、多肢選択法でも論文形式法でも利用することができるということです。第(4)は、分割点推定者の作業をきわめて簡素化することができ、第(5)は、テスト問題の内容が分割点設定の作業にも反映できるということです。

これを実施するには、OIB (ordered item booklets) を作るというのが基本になります。最もやさしい項目が最初で最も難しい項目が最後という順番にカードを並べます。この項目困難度は、受験者集団によって左右されない、項目応答理論で求めたものです。それからこのカードには正答確率0.67の場合の受験者の能力値が推定されています。テストに十分解答できる受験者の能力を見つけ出すようなことは、これまでの古典的テスト理論で

は、不可能な事でした。この 0.67 は、R P (response probability) とよばれているもので、特定の水準のボーダーラインにいる学生が、正解するであろう最低の正答確率と考えられています。これには、項目応答理論で求めた、受験者能力、項目困難度、正答確率が必要です。それがテスト項目内容とあわせて示される、そういうカードを作成することになります。

最後に Bookmark Method の実施手順の 1 例をご紹介します。テスト後、項目応答理論の 2 パラメーターロジックモデルで a (弁別力)、 b (項目困難度) を求める。それから 0.67 の正答確率になる場合の受験者の能力 θ を項目ごとに求めます。その後、易しいほうから難しいほうへの困難度順に OIB を作成する。ここで、複数の分割点推定者に対して目標到達者に辛うじて所属すると思われる受験者が 67% の正答確率になると考えられる最後の項目の頁の次の頁に「しおり」を置くよう指示します。これがいわゆる bookmark といわれるものです。複数の推定者、たとえば、10 名程度の方がこれに携わっているようです。そして、分割点推定者の提出した受験者の能力値の平均を求め、それに該当する素点を求め、それを分割点とする。こういった手順で、実施や研究が行われています。

時間がだいぶ過ぎましたが、まとめとして、申し上げたい事のひとつは、やはり目標基準準拠測定の趣旨がどうなっているのかということをきちんと掌握することです。そしてその趣旨が生かされる体制を教育と研究の場で確立する必要があると考えます。また、テスト項目の分析や解釈には、言語能力の視点にたった、より適切な掌握が必要です。その基盤に立って、分割点設定の科学的な手法を開発する必要があるでしょう。80 点をとったら合格だというのはどういうことなのか、なぜ 80 点なのか、その辺をもう少し科学的な手法で開発する必要があるだろうと考えます。教育現場で日夜努力をされている先生方の更なるご活躍を祈っております。ご静聴まことに有難うございました。

以上

5. シンポジウム「英語教育に関する評価の諸相」

評価の有り方に関し

名工大インテンシブコースのデータから

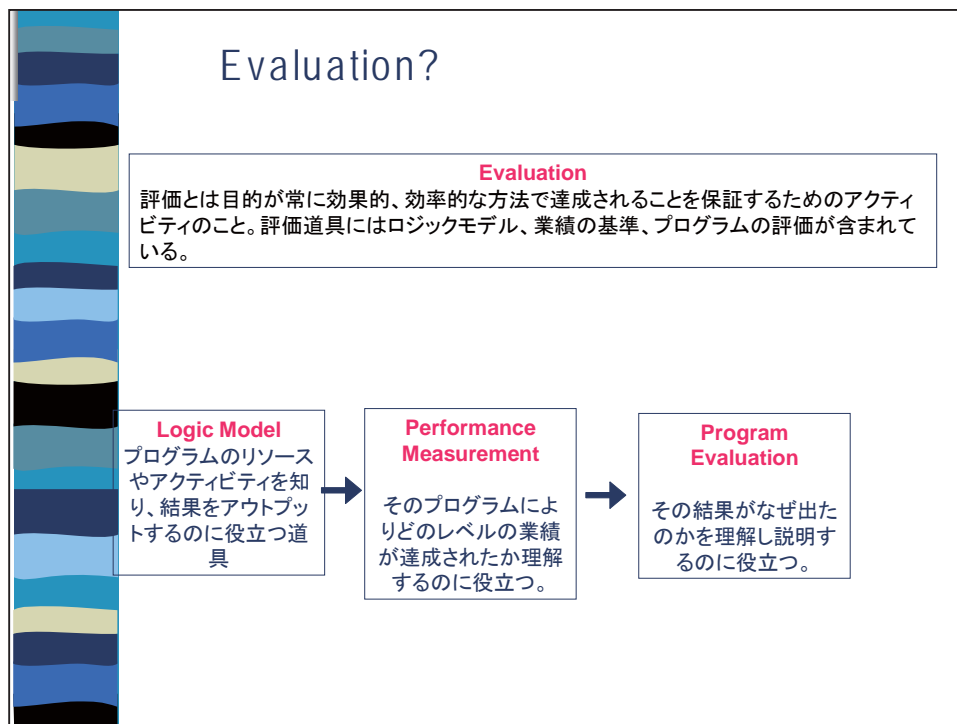
クイン・ケリー

名古屋工業大学大学院助教授



皆さん、こんにちは。名古屋工業大学のケリー・クインです。皆さんに負担になるかと思いますが、今日、私は日本語で発表します。私の課題は「評価のあり方に関し、名工大集中クラスのデータから」です。この発表のため、私は「evaluation」の意味を真剣に考えました。大友先生の前に発表すれば良かったなという気がしています。いろいろな本やresearchを参考にしましたが、まだまだわからないところがあります。しかし、結局、私は「evaluation」を【図1】のように定義しました。

【図1】



この定義を踏まえ、私は今からロジックモデルについて説明したいと思います。Performance Measurementのことについてはこの後に発表する小山先生、石川先生が説明すると思います。Program Evaluationのことは教員でカリ

キュラムの改善のための考察をします。

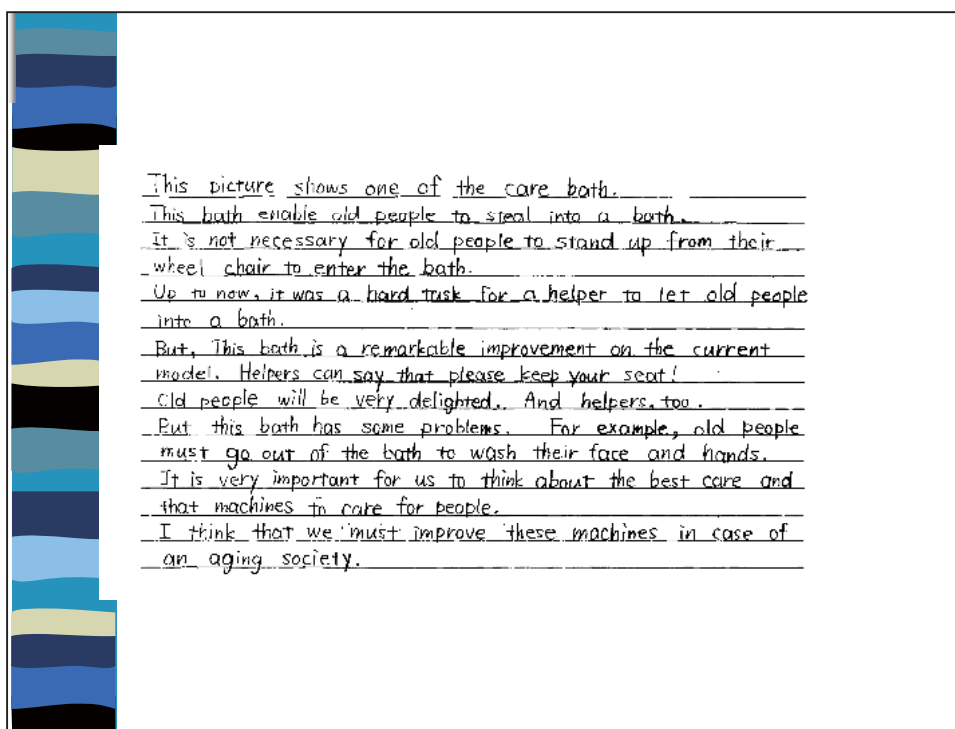
最初に、GPプログラムの目的を紹介します。『「知識としての英語」から「道具としての英語」へ』というものです。これは、英語を使える学生を育てたい、英語を書ける学生、英語を話せる学生、英語を理解できる学生を育てたい、というものです。この目的を満たすために、GPクラスにいる1年生の英語カリキュラムを紹介します。

1年生の英語の講義は一週間で二つ、必修科目の「科学技術英語Ⅰ」と選択科目の「外国

語演習」が1週間1回、1学期で15回。「科学技術英語 I」は一週間おきですが、教科書と期末試験は統一しています。夏休みの間に課題を学生に与えています。「外国語演習」は選択科目です。内容は各教員が決めるのですが一応スキル別で分けています。もう少し後で教材のことを説明しますが、しかし、先ほどご紹介した「英語を使える学生を育てる」という目的を達成するためにはこのコマ数では足りません。とりわけ、リーディングとライティングの課題を増やすべきだと思いました。だから、GPクラスの学生は、一般クラスの学生と同じ教科書を使っていますが、多読プログラムを設けましたし、作文の課題も多く与えるようにプログラムを設計しました。この二つの課題がGPプログラムの特徴です。もうひとつ特徴があります。学生の人数が多すぎると教員の負担が大きくなるので、GPクラスは学生人数を制限しました。一般クラスの場合、必修科目の「科学技術英語 I」で再履修生も含め40名程度、選択科目の「外国語演習」で50名程度ですが、GPクラスは少人数で25名としました。そして夏期休暇中と春期休暇中に「集中クラス」を設けました。先ほど、永井先生がご説明したように、「集中クラス」は1週間毎日3コマの授業を集中して行います。

今から教材を紹介します。「科学技術英語 I」は全クラス同じ教科書を使います。中身はサイエンステクノロジーに関係しています。

【図2】

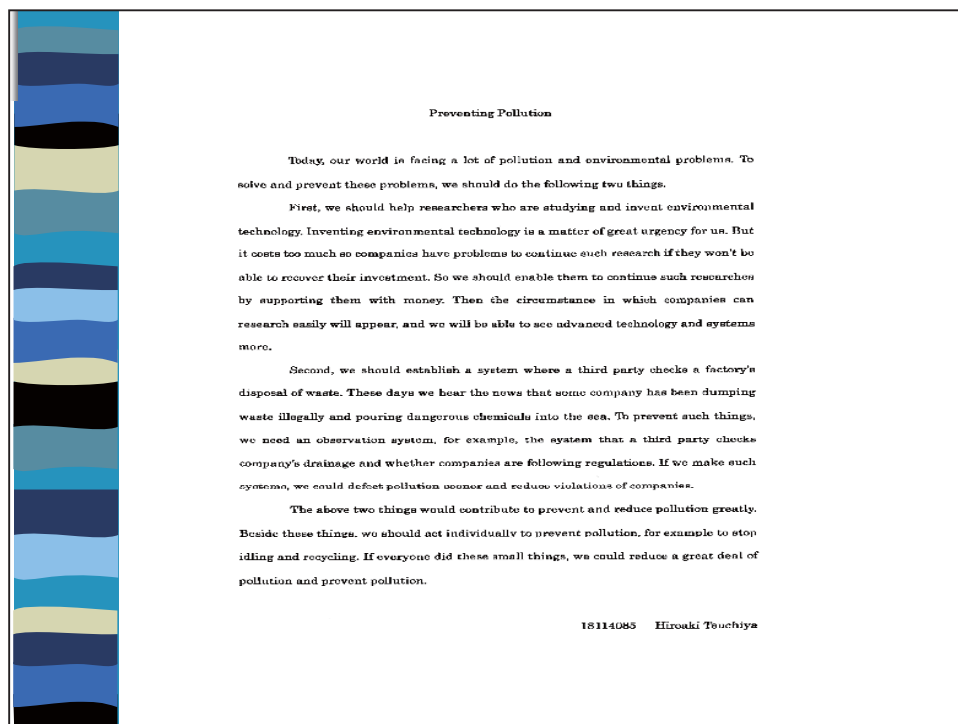


各ユニットは同じアクティビティからなっています。最初に語彙演習、それからカンバセーション、レクチャー、リーディング、リサーチ&プレゼンテーション、ライティングと続きます。簡単に言えば、この教材の目的は、学生に直接教えるのではな

なくて、たくさん練習するチャンスを学生に与えることです。教材として、十分なリスニングとスピーキングチャンスを学生に与えています。ライティングとリーディングの部分はちょっと足りない。たとえばライティングをしないと先生が思っても、何を書くか、どういう風を書くか、どのように情報を整理するか。ということは、教科書自体にはあまり書いていないと思いました。そのためGPプログラムではもうちょっと学生に直接教えたいと考え、そのようなプログラムを設け、学期中に三つの小論文を書かせました。また休暇中に実施した集中クラスでは、先ほど永井先生からも紹介がありましたように、午前中

にリーディングとライティングを勉強し、1時から2時30分まではリスニングとスピーキングを練習、2時40分から4時10分までの授業ではライティングと文法を学生にやらせました。このように、毎日ライティングと文法を入れました。あまり時間がないのですが、学生のエッセイを二つ見せたいと思います。【図2】と【図3】をご覧ください。

【図3】



最初は通常のクラスの学生の書いたものです。まあ、悪くないと思いますが、多少の問題があります。パラグラフのことをあまり理解していない、文章があまりつながっていない、という問題があります。次に、GPクラスの学生のエッセイです。パラグ

ラフはきちんと教えられたのでしょ、ファースト、セカンドと続き、そのつながりもきちんと書いてあります。このエッセイでは情報が整理されています。

GPクラスの2つ目の課題は多読です。たくさんの読み物を集めました。まず、アメリカやイギリスの中学生や高校生向けにリバイズされた英文学や児童文学などです。さらに若者向けのノンフィクションも集めました。この本は、中世時代の城の作り方というものです。本の単語数別とレベル別に点をつけました。簡単な本は1点か2点、やや難しい本は5点。かなり難しい本は10点か15点をつけました。1学期でGPクラスの学生は30点の点数を集めなくてはならないとしました。本を読み終わって短いブックレポートを書かせました。学生は自由に本を選ぶことができます。今回の取組で500冊以上の本が集まりましたから、かなりいろいろな本を揃えることができました。

ブックレポートの書き方も学生に説明しました。条件が5つあります。ブックレポートの中にまず粗筋を書かせました。次に登場人物の中の誰にもっとも共感するか、それはなぜか、このようなことも書かせました。このブックレポートは英語で書きます。これはすごく重要なことです。それを担当教員に提出します。石川先生が今から科学技術英語 GPクラスのアンケートを分析しますが、私からもひと言。夏期集中クラスアンケートの結果を分析しますと、下のレベルであるレベルBの学生からの回答では作文能力が身についたという回答が多くありました。その通りであれば幸いです。以上で発表は終わります。

量的評価の観点

TOEIC、統一テスト、語彙データから

小山由紀江

名古屋工業大学大学院教授



発表要旨

C. Alderson は "Evaluating Second Language Education" (1992) の中で、英語教育の評価のガイドラインについて論じているが、そこでは評価の計画、実施、解釈と報告、評価の利用と評価自体に対する評価が、それぞれさらに詳述されている。Alderson によれば、評価の計画に関して重要なことは、目的 (Why?)、評価の audience (Who for?)、評価する主体 (Who?)、評価の対象 (What?) を出来る限り明確にすることである。(p274-304) ここでは、何を What? 何のために Why? 評価するか、ということ視野に入れつつ、量的データを基に名工大現代 GP 英語プログラムの評価を試みてみたい。しかし、今回使用するデータは本プログラム全体をカバーするものではなく、2006 年度の 1 年生が受けた三種類の試験の結果だけであり、この評価が限定的なものであることを最初にお断りしておく。

本学では 1 年生全員に対して、入学時と学年末に二回の TOEIC テスト、及び、前期後期のそれぞれ最後に「科学技術英語 I」の期末試験として統一テストを課している。さらに、後期の途中で科学技術語彙のパイロットテストを様々な英語習熟度レベルの 283 名の学生を対象に実施したが、このテストの語彙は独自に構築した科学技術論文コーパス (5 分野、約 350 万語) の分析結果を基にした工科系大学生用の教育用語彙表から抽出したものである。名古屋工業大学における現代 GP の英語プログラムも 2 年目を終えようとしている現在、この発表では、これらのテストの結果を分析し、本プログラムにおける学習者の学習状況を明らかにすることにより、今後のプログラム進行に反映させ得る形成的評価として何らかの方向性を提案したいと考えている。

私は英語のテストを分析した結果を皆様にお話したいと思います。先ほどこイン先生も真剣に悩んだと言われましたが、私も「評価とは何か」ということに関して真剣に考えました。やっと昨日の夜くらいに少し自分なりに整理されてきましたので、簡単にもう一度説明させていただきます。

今回、分かったことは、評価というのは切り口によっていろいろな見方ができるということです。Charles Alderson は「Evaluating Second Language Education」(Cambridge Univ. Pr.) (1992) という本の中でいろいろな切り口について言及していますが、評価の目的やオーディエンス、それからいったい誰が評価するのかという評価の主体など、これらできるだけ明確にすることでよい評価ができるという説明をしています。

では「評価の対象は何か」といったときに、これは二つのレベルがあると思いますが、大きく分けると、個々の授業なり、ある単元なりの中での学生のパフォーマンス、それを日々、小テストで測る、あるいはまとめて一ヶ月くらいたったときテストをする。そういう形での学習者のパフォーマンスを対象にする。もうひとつは、たとえば今私たちがやっ

ているような、この現代 GP プログラムを「プログラムとして」評価するというものです。このように、二つの見方ができると思います。また「評価の目的」という観点から見ますと、これは教育評価と言ったときに誰でも思い浮かぶことですが、ひとつが学習者へのフィードバックです。学習者自身がテストスコアを見て、自分が何をどれだけ理解できたか、あるいはできなかったということを自分の学習の指針にするということ。もうひとつは教員自身へのフィードバックです。これら二つを総合して考えていくことによって、カリキュラム等への見直しを行い、より良いカリキュラムを作っていくということになるのだと思います。

次に「評価を行う時期」によって評価の分類がされています。ひとつには「診断的評価」。これは、学習の準備状況を知るためにあるコースや学習の始まる前に行うものです。そして次が「形成的評価」。これは、コースや学習の途中で行って、目標と照らし合わせながら学習のペースを調整する、あるいは問題点を診断してそれ以降の学習を改善するために行うものです。そして「総括的評価」。これはあるプログラムの最後に、その全体が果たしてうまくいったのか、あるいは何らかの問題があったのかというようなことを評価する。これが時期によっていろいろ見方があるという評価の三つの形です。

また「評価 (Evaluation)」とは「測定された結果に何らかの価値判断を加えること」に対し、「測定 (Measurement)」というのとは「学習者のパフォーマンスに対する客観的なデータを得ること」であって、しかも客観的にデータを得るためには、客観的にデータを得るための条件があります。これは、「一定のスケールを用いて数量的に客観的に、正確に捉える」ことが必要条件であるということです。

これらの特定の道具として、教育現場での「テスト」があります。これが客観的データを得るために使われているわけですが、客観的で正確なデータを得るために何が重要かという、テストの一番の基本は「信頼性と妥当性」です。ここでは本当にごく簡単に論じますが、「信頼性」というのは consistency of measurement (Backman & Palmer, 1996) ということです。「妥当性」というのは「測定しようとしている能力を適切に測定しているか」という意味での appropriateness, meaningfulness, and usefulness (AERA, APA, 1985) です。信頼性とは妥当性を得るために必要条件で、信頼性のないテストは妥当性を得ることはできません。その逆は成り立ちません。妥当性の高いテストであるためにはまず信頼できるテストでなければならないということが言えます。

ここで「量的評価と質的評価」という観点から評価のあり方を整理いたします。後から石川先生が質的データに関する発表をさせていただきますが、単純に言うと量的評価とは量的データに基づく評価、質的データとは質的評価に基づく評価ということになります。では量的データは何かというと、皆さんが日ごろ扱うことが多いと思われる、テストのスコアなどの数値です。質的データとは、たとえば観察したり、インタビューをしたり、アンケートをとったり、ポートフォリオをやったり、ログをとったり、数値的評価の難しいもの、これらが質的データと呼ばれるものです。こちらのほうは整理、分類、データ解釈などに非常に手間がかかる種類のデータです。

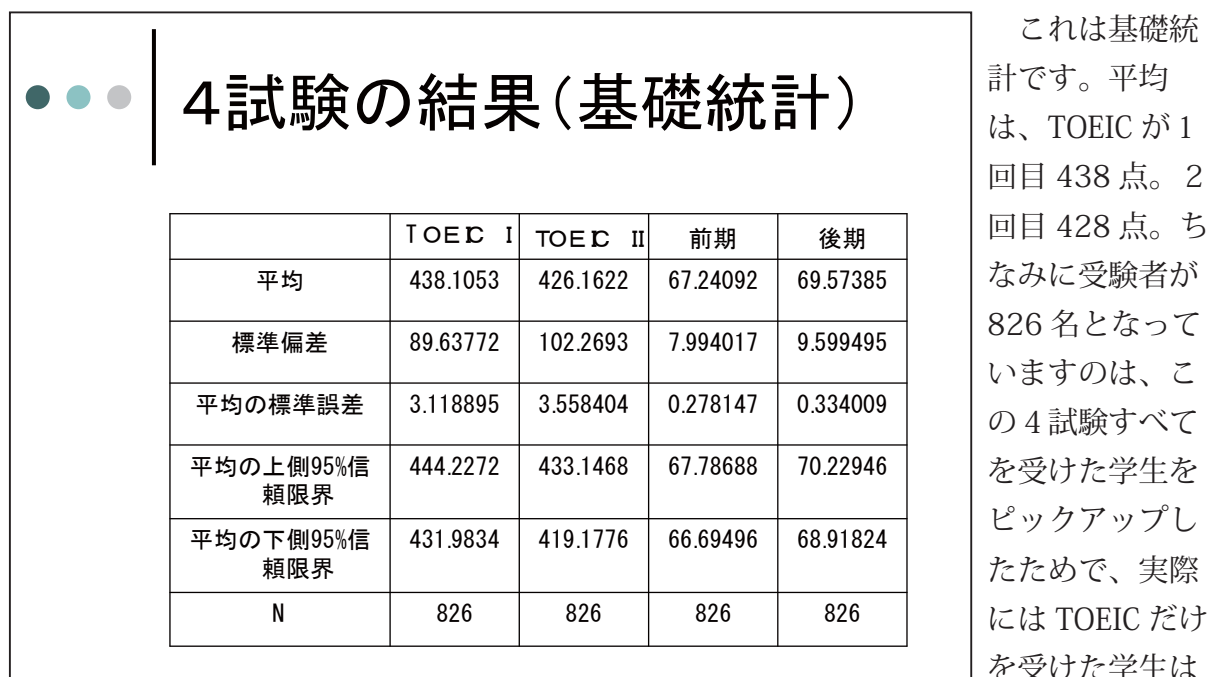
このデータを簡単に分類しますと、質的データの中に「名義尺度」、「順序尺度」というものがあり、量的データの中に「間隔尺度」、「比例尺度」というものがあります。「名義尺度」というのは、たとえば男女、大人と子供、犬と猫などといったように、質的な違いに数値

をつけたものであって本来は数値ではないものです。「順序尺度」とは順序に数値をつけたもので、たとえばアンケートで1が大変良い、2が普通、3が大変悪いというように番号がついていますけれども数値自体には意味があるわけではありません。ところが「間隔尺度」から、数値に意味が出てきます。たとえば気温のように、等間隔であるが原点と単位は任意のものがそうです。実際、華氏と摂氏では原点も単位も違いますね。次に「比例尺度」がありますが、これが measurement のときに数として扱うもので、絶対的な原点 0 と等間隔の単位を持つものです。いちばん分かりやすいものは身長や体重です。何キログラムとか何センチメートル、こういったものが比例尺度といわれるものです。

今回は本学の1年生を対象とするテストを量的データとして扱います。三種類あります。ひとつは TOEIC テストです。これは入学時に TOEIC IP の1回目を実施し、2回目は2月に実施するものです。二つめは学期末に実施する統一テストです。これは、教科書である「Humanity and Technology」から出題します。全部で100問、リスニング、リーディング、ボキャブラリーがそれぞれ約30問程度出題されます。もちろん前期と後期はカバーする範囲が違いますので問題の中身は違います。三つ目は、パイロットとして行いました語彙テストで、この分析も加えます。この場合の語彙は言うまでもなく科学技術語彙です。5分野の科学技術論文コーパス(約350万語)を構築し、5分野すべてに共通に出現した語彙から、既習語彙約1000語を除いたものを工学部の学生に対する教育語彙として、頻度リストを作りました。その頻度リストを基に、なるべく等間隔で語彙を取り出し、「発信語彙」と「受信語彙」の4択のテストを作成しました。対象としましたのは283人。上級、中級、基礎、各クラスから色々なレベルの学生を対象にしてテストを行いました。

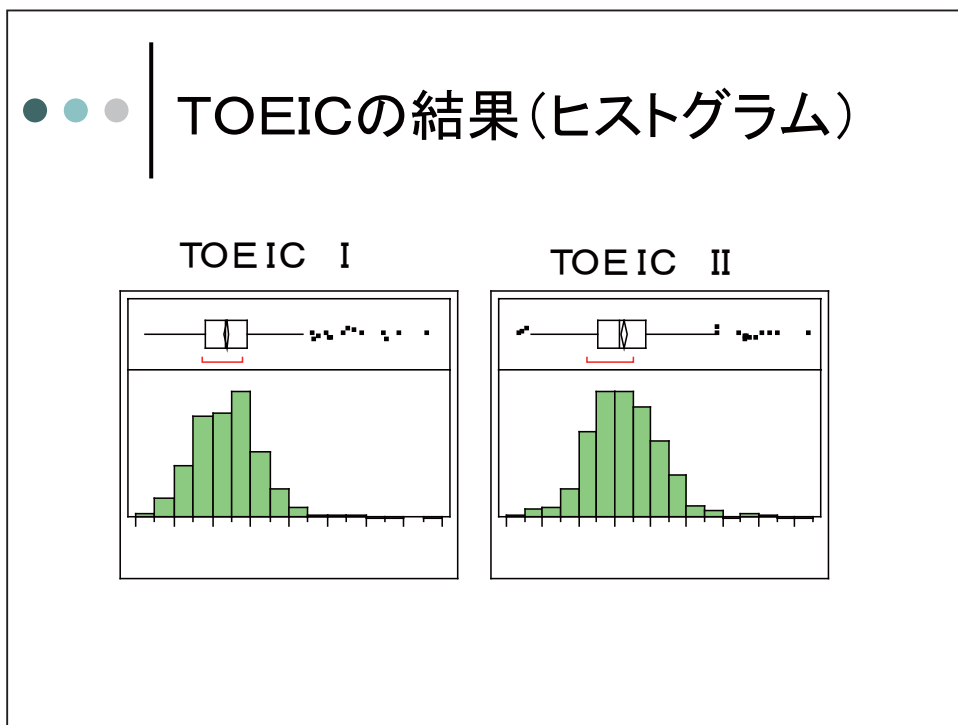
統一テストと語彙テストに関しましては、自作のテストですので、信頼性係数を計りました。クロンバック α が 0.7974。これは統一テストの前期後期全体の信頼性係数です。語彙テストは 0.83 でした。いずれも 0.8 くらいあるのでテストとしての信頼性はある程度高いと考えて良いでしょう。4試験の結果を【図1】に示しておきました。

【図1】

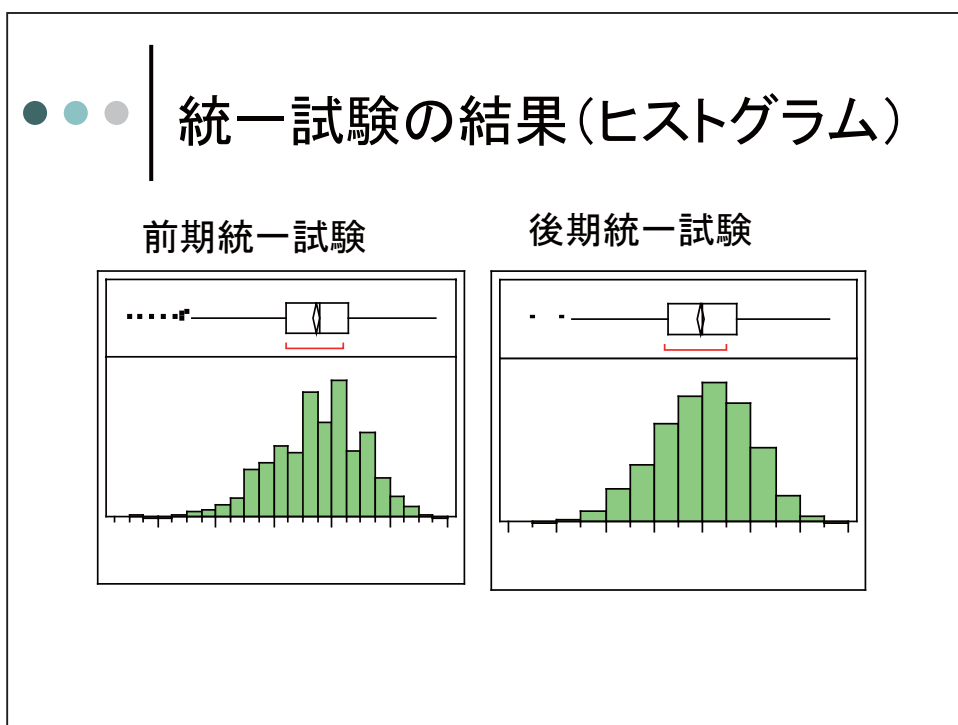


1000 人くらいおります。4 試験全部を受けた学生は 826 人ということです。この二つのテストは中身が違いますから単に平均が書いてあるだけです。標準偏差もこのような結果です。もちろん TOEIC はご存知のように 990 点満点に対し、統一テストの方は 100 点満点ですのでスコアが違います。次に、TOEIC の結果と統一試験の結果をヒストグラムにしてみました。【図 2】と【図 3】です。

【図 2】



【図 3】



【図2】に関して言えば、2回目は山が400点のあたりにちょっと左にずれている感じが見受けられます。統一試験のヒストグラムである【図3】に関して言えば、前期は凹みが出ましたが、後期は大変きれいなベルカーブになっています。そして統一試験の前期をIRT（これは2パラメーターですが）をかけてみました。数は少ないですが、項目困難度がかかなり高いものと低いものがあり、識別力という点から幾つかは問題の差し替えを検討する必要があります。

TOEICにもどりますが、TOEICのレベル別に少し変化を見てみました。実はこの3年間ずっと同じようなデータを私は取ってきたのですが、毎年上位は下がっている、中位はほとんど同じ、下位が上がるというのが毎年の傾向でした。ただ大変残念なことに、今年は上位の下がる率がやや大きく、平均点が例年と比べてさほど上がらなかったという結果が出ています。

これらすべての試験（TOEIC2回、統一試験2回、語彙テスト1回）の相関を見たところ、実は相関はあまり高くありませんでした。TOEICの試験と統一試験の相関もそれほどなく、語彙のテストとTOEICの試験の相関もあまり高くありません。語彙のテストは、科学技術の教育用語彙を用いたものですので、私たちとしては、ここでもう一度、コース（科学技術英語）の目的とそれぞれのテストが、妥当性という観点から見て適切なものであるのかどうか、本気で問い直さなければいけないと感じております。そのためには、コースの目的を明確化することと、その目的に合致した授業を行った上で、学生のパフォーマンスを計るテストを作成することが必要であると思われます。その意味で、今回は、TOEICその他の試験の結果を比較分析することによって、本学の英語教育の内容を省みる良い機会を得ることができました。以上で発表を終わります。

量的評価の観点

学生のアンケート調査から

石川有香

名古屋工業大学大学院助教授



発表概要

本発表では、名古屋工業大学 EGST 教育プログラムとして 18 年度に行った、1 年生対象の英語科目（科学技術英語 I・外国語演習）のうち、ネイティブ教員が担当する上級英語学習者用コース（GP コース）を取り上げ、学生のアンケート調査の分析結果を踏まえて、英語教育活動の定性的評価を試みる。次いで、19 年度プログラムの実施に向けて、当該英語教育コースの改善点を提案する。

1. 質的評価の重要性

一般的に評価と言うと、個々の学生に成績をつけるイメージがありますが、それと同じように、教育カリキュラムやプログラムもまた評価の対象になります。ここでは、本学の EGST 教育プログラムをどう評価すべきか、まずはその観点について考えてみたいと思います。

教育プログラムの評価方法には、量的データを基に評価を行う定量的評価と、質的データから評価を行う定性的評価があります。

定量的評価とは、プログラム受講生のテスト成績などの数値データを用いて評価を行うものです。たとえば、プログラムの受講生が TOEIC(R) (Test of English for International Communication) のような信頼性の高いテストを受験している場合、そのスコアの変化を根拠として、プログラムの有効性を定量的に評価することができます。

定量的評価は、実用性や客観性が高く、プログラムの有効性を一目で判断することができます。しかし、重要なことは、定量的評価だけでは、プログラム参加者の質的な変化を捉えることができないということです。

ここで重要になってくるのが定性的評価です。定性的評価とは、プログラム参加学生の学習行動や学習活動の質的な変化を分析し、その変化の程度にもとづき、プログラムの有効性を見定めようとするものです。

定性的評価では、テストの点数といったはっきりした資料がないのが普通ですから、記述式アンケートやインタビューを行ったり、観察日誌をつけたりして、これらの記述的情報をもとに評価を行います。定性的評価はしばしば膨大な作業量を要し、加えて評価者の主観を統制する難しさもありますが、こうした評価によって、学習者の意欲・態度・関心といった側面の変容の実相を正しくとらえることができるのです。

こうした二つの評価の視点のうち、以下、本発表では、定性的評価の立場から、EGST 教育プログラムの有効性を検討してゆくこととします。

2. EGST 教育プログラムの質的評価

名古屋工業大学 EGST 教育プログラムでは、1 年生を対象とした英語科目として、「科学技術英語 I」と「外国語演習」の 2 科目を設定しています。18 年度は、この 2 科目の中に、上級学習者用特別コースとして GP コースを設置しました。GP コースでは、英語のインプット量を確保することができるよう、全クラスをネイティブの教員が担当して、英語で授業を行うとともに、多読活動を取り入れました。また、教師の目がいきとどき、学習者が「話す」「書く」の活動も十分に行えるよう、クラスサイズは 30 名以下としました。さらに、集中的に英語学習が行える機会を設けようと、「外国語演習」は時間割を変更して、5 日間連続で行っています。

EGST 教育プログラムの質的評価を行うにあたっては、この GP コースを受講した学習者のうち、コースを最後まで受講した 187 名を対象に、自由記述式のアンケートを行うこととしました。アンケートでは、EGST 教育プログラムについて、1) 良かった点、2) 悪かった点・改善が必要な点を、それぞれ、最大 3 点まで、自由に書いてもらいました。

こうして収集した記述データを、その内容に沿って、数種に分類することで、プログラムの評価を行ってゆくこととします。

2.1 肯定的な評価

まず、肯定的な要素を分類してゆきたいと思います。アンケートでは、教員に関する記述が最も多く認められました。たとえば、「先生がネイティブだったので、ネイティブの英語を聞くことができた」「ネイティブの人と話すのに慣れた」「ネイティブの先生と会話できた」など、ネイティブ講師がクラスを担当したことについて、肯定的な記述がなされていました。また、通常のクラスよりも少ない人数でクラスが運営されていたことについても、多くの言及がありました。たとえば、「クラスが少人数制でよかった」「少人数制だったので、先生と接する機会が多く持てた」というものです。それから、「丁寧に指導してもらった」という記述も数点ありました。これも、少人数制度に関係する記述と言えます。

学習活動についても多くの記述がありました。先に述べましたように、このコースでは、多読活動・英作文活動にも多くの時間を割いていますが、特に、英作文活動は、多読活動に付随したブックレポートと独立した活動のエッセイ・ライティングの 2 種を行っています。英作文についての記述は、「英語でレポートを書く機会が得られた」「ブックレポートを書いた」「英作文を添削してもらった」というものが多く、多読に関しては「英語で本を読んだのは良い機会だった」「英語の本なんて読めないと思っていたのに読めた」などというものでした。

さらに、英語で行う授業活動については、「英語で話す自信ができた」「英語で話すことになれた」「ゲームが楽しかった」「意欲がわいた」というような記述がありました。以上の肯定的な記述をひとつの大きなカテゴリーとしてとらえ、ここから、共通因子を抽出してみると、これは、「学習動機の向上」と言えるのではないかと思います。

次に、もうひとつの大きなカテゴリーを見てみましょう。アンケートでは、自己の英語能力に関する記述が多くありましたので、これらをまとめてみます。そうすると、こちら

は、「英語運用能力」に関するカテゴリーになると言えます。これらは、あくまでも自己評価ですが、「リスニング力が向上したと思う」「リスニング力が上がった」「リスニング力がアップ!」という意見がたくさんありました。それから、「話す力がついた」「英会話能力が上がったと思う」というものも多くありました。また、オーラル面だけではなく、「英語の本を読み、リーディング力がついた」「リーディングのスピードが上がった」、さらに「エッセイの力がついた」「作文構成力がついた」という声もありました。このような記述から、GPプログラムの受講生には、オーラル面だけではなく、4技能にわたって、「英語運用能力向上の実感があった」ということができます。

以上、肯定的記述の分類分析から、GPプログラムでは、「学習動機の着実な向上」と「英語技能向上の確かな実感」の相乗効果が認められたということが出来ます。

ここで、もう一度、GPプログラムでの肯定的評価をまとめておきたいと思います。評価されたことは、

1. インタラク션을促進する少人数制
2. ネイティブ講師を中心にした授業
3. オーラル面のトレーニングを中心にした授業
4. 多読・英作の自習プログラム

の4点です。

こうした評価の背景には次のようなことがあると言えます。現在では、高校でのALT配備も進んでいるのですが、まだまだ個人レベルでのインタラクシヨンの機会は限られています。ALT教員は月に一回やって来るだけで、しかもたくさんの人数の学習者に対して、たった一人のALT教員が対応するという状況が多いので、「ネイティブの先生と個人的にインタラクシヨンがとれた」「私の作文を読んでくれた」「添削してくれた」ということが高く評価されたのではないかと思います。ネイティブ講師が担当する場合には、少人数が教育の効果を高める基礎的要件となると言えるようです。また、少人数制によって、教員を身近に感じるため、学習動機が上がり、意欲が湧いた受講生は、実力の向上も実感できたと思われます。すなわち、少人数制とネイティブ講師、4技能にわたる指導の3点が揃ったために、良い評価が得られたと言えましょう。

2.2 否定的な評価

次に、GPプログラムで、悪かった点・改善してほしい点についての記述を分析してゆきます。実は、否定的な意見は、肯定的な意見の1.3倍くらいありました。しかし、こちらも分類して分析してゆきますと、3つの問題点が明らかになりました。具体的に記述を見てみましょう。

まず、「先生と話す時間をもっと多くってほしい」「生徒同士の会話練習ではなくネイティブの先生と話したい」「先生とのコミュニケーションが少ない」「英会話能力は日本人と話しては上がらないと思う」と、一見、授業方法や教員に対する否定的な意見であるかのような記述が多くなされていました。しかし、これらをまとめて見てゆきますと、「ネイティブの先生とのコミュニケーション時間をもっと多く取ってほしい」という要求が満たされないことへの不満であることがわかります。

また、「テキストが簡単」「レベルが低い」「もっと難しくしてほしい」といった意見も

たくさんありました。先ほどの発表でもご紹介しましたように、名工大英語プログラムでは、統一教科書を使っております。GPプログラムは、上級者を対象としていますので、こうした不満も、学生の「もっと勉強したい」という意欲の表れだと言えるでしょう。

さらに、多読プロジェクトに関しては、「多読よりもむしろ、英文を読むテクニックを教えてほしい」「文法を教えてほしい」という意見や、英作文に関しても、「きちんとした文法指導が必要だ」「文法を勉強して正しい英文を書けるようになりたい」などの意見が出され、学習者が、英語学習または英語そのものに、興味を持っており、同時に、「正しい英語」への憧れを抱いていることが明らかになりました。また、「先生がエッセイを読んでくれないことがあった」「レポートをもっときちんと添削してほしい」というような意見もありました。このように、「もっと指導をしてほしい」という要望は、すなわち、コミュニケーションの道具としての英語使用能力の獲得に飽き足らず、学習者が、さらに高いレベルの英語能力の習得を望んでいることを示しています。

また、多読プロジェクトに関しては、「本の貸し出しが1冊しかないので不満だった」「自習リーディングの本の種類をもっと増やしてほしい」という意見や、さらに、「ブックレポートを書くほうが大変で、じっくりと本を読めない」という不満が出されていました。英語作文に関しても、「エッセイ提出までの時間が短くてじっくり書けなかった」というような問題点が指摘されました。また、集中コースの場合には、「毎日授業があるので、家に帰って課題をする十分な時間がなかった」という指摘もありました。こうした記述の分析からは、学習者には、強い学習意欲があり、「もっと自由に学習を進めたい」という希望があるにもかかわらず、プログラムの方が学習者の様々な要求に応えきれていないという問題点が浮き彫りになっています。

さらに、クラス運営・コース運営についての問題も指摘されています。「すべてが英語なので、何をやればいいのか、理解できないことがあった」「試験などの重要なことも英語なので聞き取れないことがあった」という記述が多くあり、意欲があるにもかかわらず、学習活動を進めることができない場合の問題が挙げられていました。学習活動がクラス運営方法によって阻害され、学習者の学習意欲が失われるのであれば、プログラムの有効性は著しく損なわれます。ネイティブ講師が英語クラスを運営する場合には、学習活動が滞りなく行われるよう、工夫が必要であると言えます。

最後になりますが、アンケートでは、「GPクラスに入るかどうか、それは本人が選ぶようにしてほしい」といった意見も散見されました。柔軟なプログラム運営が求められていると言えるでしょう。

以上、否定的な記述をまとめて、GPプログラムにおける問題点が明らかになりました。プログラムに必要な改善点として、以下の3点を指摘しておきたいと思います。

1. 学習意欲と提供プログラムのミスマッチの解消
2. 自律的学習を促す対応
3. 柔軟なクラス運営とプログラム運営

一つ目の問題点は、「学習意欲と提供プログラムのミスマッチ」です。強制的にクラスの割り振りをしてしまったために、「正しい英語」に憧れ、さらに高いレベルの英語能力を欲する学習者に対して、「道具としての英語」を使用するための運用練習に焦点を絞ったネイティブ主体のプログラムを提供してしまっているという点です。次に、「自立的学

習を促すような対応が取れていない」という問題点があげられます。プログラム外の学習活動を想定していないため、ある一定のレベルまで達した学生が、次のステップに進みたい時に、学習者の要求に応えるシステムができていません。三番目としては「柔軟なクラス運営とプログラム運営ができていない」ということであると思います。やる気があるのに、英語の指示がわからないで、ついていけない学生がいるという事実があります。クラス運営を合理化することが肝要かと思われまます。ネイティブ講師によるコース運営には、長所と短所がありますので、それを十分に把握したうえで、学習者のレベルに応じて上手く提供していく必要があるでしょう。

3. 19年度プログラムにむけて —改善の指針—

アンケート記述の分類分析より、上で述べたようなことが明らかになりました。ここでは、上の分析結果を踏まえ、19年度プログラムの改善提案を行いたいと思います。以下に、4つの提案を行います。

1. クラス運営・プログラム運営の合理化
2. 学習者の知識意欲に応えうる各種教材の設置
3. 自律的学習者育成を可能にする学習環境
4. 多面的に学生のニーズを受け止めるシステム作り

具体的には、クラス運営・コース運営については、プログラム評価にかかわる教員が、担当のネイティブ教員との連絡を密にとり、学習者の学習活動が阻害されることのないように、柔軟な運営を求めてゆくことが必要であると言えるでしょう。また、多読プログラムとして設置している教材の多様化や、多読のための解説書や方略書をはじめ、英文読解についての参考書や問題集の設置、さらに、「正しい英語」を求める学生に対しては、英語語法辞典・文法書などの設置が考えられるでしょう。こうした教材の整備に加えて、自律的学習者の育成を見据えた、面として学生のニーズを受け止めるシステム作りが必要なのではないかと思われまます。たとえば、週に1回、英語授業のGPコースのみをネイティブ教員が担当するというような、「点」で、学生のニーズに応じてゆこうとするのではなく、様々な方向から英語教育を行っていく、「面」としてのシステム作りというのが必要ではないでしょうか。

こうした改善策を考える上で、他大学の英語教育への取組を参考にするのも良い方法だと思います。たとえば、東京大学では、十分なアカデミック・ライティングの指導が行えるように、ライティング・センターを設置しています。また、神戸大学では、常駐のネイティブ教員によるアカデミック・プレゼンテーション指導を行っています。大阪大学でも、CALLシステムと留学生のTAを活用して、アカデミック・プレゼンテーション能力を育成するESP教育を行っています。近畿大学では、学生や教職員がいつでも利用できる「英語村」を設置し、ネイティブと楽しみながら英語を学ぶ機会を提供しています。こうした取組を少しずつ参考にしながら、多面的に学生のニーズを受け止められる、より良い英語教育プログラムを開発することが必要なのではないかと思われまます。

量的評価の観点

共通尺度としてのプレイスメントテスト

清水 裕子

立命館大学教授



発表要旨

「大学全入時代元年」も、実際には数年後に先延ばしになる状況のようであるが、少子化や新設大学・学部
のラッシュの中で、今までの入学者層とは明らかにことなる学生を受け入れ始めていると感じている大学が多
いのではないだろうか。英語教育を考えた場合、入学者に対して、何らかの形で英語科目の履修を課している
大学がほとんどだと思われるが、大学生の英語力の低下だけではなく、英語力の二極分化を実感している英語
教員も多い。特に私立大学においては、入学者の選抜方法が多様であり、入学試験を通じて共通の尺度で英語
力の測定を行うことができず（英語テストを課さない入試形態をとっているケースもある）、当該テストとそ
の受験者に依存した結果をもとにして入学者が決定されている。そこで、多様な英語力をもつ入学者に対応す
る一方法として、入学後に統一的な英語テストを実施し、学習者の力を把握すると共に、さらにその結果をも
とにクラス配置を行っている機関も多くでてきているようである。

ここでは、まず私立大学の選抜の多様性の例を紹介し、立命館大学・経済学部で共通の尺度とし実施してい
る英語配置テストおよび外部テストの結果等を紹介しながら、測定誤差の存在やカリキュラムの設計上の問題
について触れ、学習者の英語能力の測定と評価の難しさを考えていく。また、今年度から展開している経済学
部の英語教育の2つの側面、つまりリメディアル教育の側面と経済学系で学ぶ学生を意識した ESP アプローチ
の側面についても紹介する。これは必ずしも英語カリキュラムの二極化と捉えて設計してきたわけではないが、
前者から後者への橋渡しについては困難を感じているのが現実である。

今回の社系の英語教育の現状の報告をもとに、工学系の英語教育との共通点や相違点、そしてお互いが学び
合えることを考えていくための話題提供ができればと思う。

1 はじめに

立命館大学の清水裕子です。私は経済学部にも所属しており、工学英語教育のセミナーに
経済学部の者がやってきてどうするのだと思われるかもしれませんが、ひとつの事例とし
てお聞きいただければと思います。

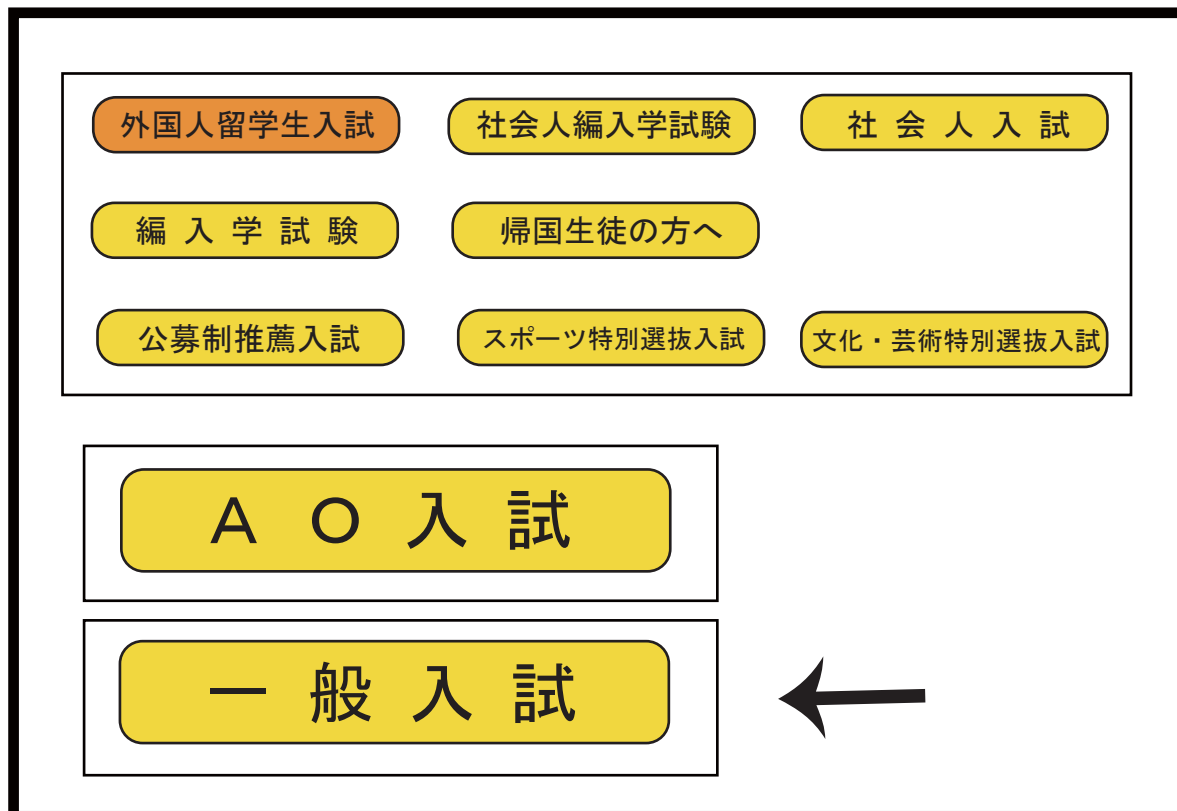
まず立命館大学について簡単にお話しておきたいと思います。京都の金閣寺の近くに
「衣笠キャンパス」がありますが、それに加えて1994年滋賀県草津市に「びわこ・くさ
つキャンパス」ができ、その後2000年には大分県に立命館アジア太平洋大学ができた
。また昨年には「朱雀キャンパス」が京都の二条城に近い所にできましたが、ここは工
科大学院と経営管理研究科があり、また本部機能が衣笠キャンパスより移りました。私
が所属しております「びわこ・くさつキャンパス」には、現在、4学部がございます。私
は経済学部で英語を教えておりますが、新入生の募集人数は785名。ただし、今年度は

1002 名の新入生を迎えました。通常、900 名前後が入学しますので、名工大の 1 学年とほぼ同じサイズだと思っていただければいいかと思います。男女構成については、経済学部の場合は 4 分の 1 くらいが女子学生です。教員数ですけれども、理系は 1 教員に対して学生数 20 名くらいですが、社系は 1 教員に対して 60 名くらいになります。経済学部は 68 名の教員がおりますが、英語教員に関しましては、専任が 5 名、任期制の教員が 1 名、計 6 名体制ですが、それに加えて、嘱託講師と申しまして、1 年契約で 5 年まで継続して教えていただける先生がおります。さらに、非常勤講師として、今年度は 4 学部合わせて 67 名に授業を担当していただいております。非常勤の先生方への依存率は 50%を超えております。こういった状態の中で授業を展開しているわけですが、私立大学である立命館大学は、いろんな形の入試で学生を受け入れております。そこでまず、入学者の選抜方法の種類を紹介し、その後、入学した学生の英語力、特に 2006 年度にカリキュラムの改革がありましたので、そのプログラムで学習を開始した一回生を中心にお話していきたいと思います。

2 入学者の多様な選抜方法

最初に入学者の選抜方法の種類についてお話しいたします。おそらく高校の進路指導をされている先生の方がよくご存知かもしれませんが、まず入試方式を分かりやすく三つに分類しました。これが【図 1】です。

【図 1】 様々な選抜方法



いちばん上の集合は特別入試に対応しています。社会人の方とか編入学の方というように、いろんな背景の方たちを受け入れるための入試形態として用意されています。それから二つ目は AO 入試で、三番目は一般入試となります。もちろん、一般入試で入学してく

る学生の比率が一番大きいです。【図1】では<一般入試>として一括りになっているので非常にすっきりしているように見えますが、実はこれが曲者です。一般入試の中身をさらに詳しくみたものが【図2】です。

【図2】一般入試の種類

一般入試

スタンダード3教科型 英語重視型入試

- ・リスニングを課す英語重視型
- ・英国2教科型
- ・英語総合型

特定科目重視型入試

- ・選択科目重視3教科型
- ・数学重視3教科型
- ・英国数3教科型
- ・理科重視3教科型
- ・3教科受験2教科判定型

センター試験併用方式 センター試験方式

- ・7科目、5教科、3科目型
- ・後期型

3月実施入試

- ・後期分割方式
- ・「〇〇学部で学ぶ感性 +
センター試験」方式

一般入試とひとことで言っても、実はこれだけありました。それでも今までよりは多少なりとも少なくなったようです。そして、これらのほとんどすべてに、何らかの形で英語のテストが関わっています。

3 入学者の英語力の測定

このように、いろんな形で入ってきた学生というのは、いわゆる英語力に関して言えば、minimum entrance requirement のようなものがなくて、出発点がばらばらになっています。そこで、どうしても共通の物差しによって英語力を測定する必要が出てきます。このことについては、6、7年前に全国の4年制大学を対象に、プレイスメント・テストあるいは統一テストの実施について調査したところ、当時、回答してくださった私立大学の半数以上が何らかの形で入学者に対して統一的な試験をやっている、しかもそのうちの半数以上がそれをクラス分けプレイスメントとして利用しているとおっしゃっていました。今ではこの数はもっと増えているかと思います。

とにかく私たちのところでも、プレイスメント・テストをやって、クラス編成を行い、そして英語プログラムを展開しております。そして今まではこれで終わっていたのですが、

今年からは到達度の検証、そして出口テストとして TOEIC® を活用することになりました。プレースメント・テストに関しましては、当初、事務サイドのほうから、60分+20～30分程度で実施できるものでないと困るといったような条件があったので、何年か前にリスニング、文法、リーディングの3つの下位テストからなるテストを作成しました。平均点が180点満点中の97点くらいです。学生の英語力がじりじりと落ちてきていると感じられる先生はいらっしゃるかと思いますが、私たちのところでは全体の平均点はそう大きくは変わってきていません。ただし4、5年前の学生と比べると文法セクションの点数が落ちてきているように感じられます。反対にリスニングは割と伸びてきているように思います。それから学力の「二極化」とよく言われますが、二極化も全体としてみると正規分布を崩すようなものにはなっていないと思います。ただし大きな幅があるわけですから、この学習者たちをランダムにクラス分けして、同じ教材を使って授業をするということは到底無理な話ですので、現在、Upper Intermediate、Intermediate、Pre-Intermediate の三つの水準に分けてクラスを編成しています。2、3年前までは、この上に Advanced と Super Advanced というレベルがありました。しかしミスプレースメントになっていたり、あるいは5つのレベルでそれぞれ異なる教材となると、例えばリーディングのクラスを担当しても、担当者によっては教科書が3種類、4種類というように、授業準備の点からも非常に効率が悪いということもあり、3つの水準の設定に変更しました。

この3水準に関しては、プレースメント・テストの総合点が高く、しかもリスニングの力がある人たちを Upper Intermediate、そして文法が弱く、しかも合計点数が低い人たちを Pre-Intermediate としています。今年からは、第1セメスターの Pre-Intermediate のレベルを補習教育するようにしています。なお、今年度新設の国際経済学科は、ある程度英語ができる学生が入学するという前提でしたので Pre-Intermediate の水準は設定していませんでした。しかしそのクラスに入った方がいいのではと思える学生も入学していますので、その人たちは大変な状態になっています。それから本来でしたらプレースメント・テストの得点をもとに、「Pre-Intermediate クラスは65%以下にする」というようなカットオフ・ポイントを決めたいわけです。ところが、入学者が決定してテストを実施し、クラス配置ができるのは4月上旬なのに、その前の3月の段階でクラス数が決定して、担当される先生には教材を配布しているわけで、結果を見てからクラス数を変更することはできません。今年度は入学者数が多かったこともあって、本来は Intermediate クラスに配置されてもいい学生が Pre-Intermediate クラスに入ってしまう、またクラス数は35人以下のはずが38人クラスというような形になっています。このあたりはどうしようもない状態です。

4 経済学部のカリキュラム

経済学部のほうからの要望としては、英語教育と専門教育の融合、到達目標を明確にしたプログラム、そして海外体験の重視があります。そのため、私たち英語教員としては、まず英語力別のクラス編成をして、共通シラバス、共通教材を用い、ミニマムレベルの設定を行なっています。国際経済クラスに関しましては1クラスが27～28名です。しかも話すことと書くことを中心として、英語ネイティブの先生が担当されるクラスである Communication & Writing では、さらに1クラスを二つに分けて13～14名のクラス編

成にしています。

次に英語のカリキュラムをご紹介します。【図3】をご覧ください。一回生の前期（第1セメスター）では全員が Reading、Communication & Writing 1、Listening 1、CALL 1 の4種類の授業を履修します。その後、専門科目へとつなげていくわけですが、どちらかといえば、EAP：English for Academic Purposes あるいは English for General Academic Purposes という形でスタートしています。

【図3】 経済学部の英語カリキュラム（1）

英語カリキュラム

| セメスター | 1回生 | | 2回生 | |
|-------------------------|--------------------|-----------|------------|---|
| | 1 | 2 | 3 | 4 |
| Reading | 基礎 科目 Basics | コース 選択 | 専門科目へのブリッジ | |
| Communication & Writing | | | | |
| Listening | | | | |
| CALL | | | | |

一回生で言えば、現在、第2セメスターが終わった段階で、これからがいわゆる専門へのブリッジになるところですが、リーディングとライティングに力を入れてほしいというのが学部からの要望でした。それで今のところ4～5週間をひとつのモジュールにして、少しずつ経済関係のコンテンツを入れ始めています。経済学の基本となるような教材があるわけですが、私たちは経済の専門の人間ではないので、Study

Skills 的なものを入れていきます。それと同時に大学英語教育学会の中で、ESP: English for Specific Purposes の研究グループがあって、その先生方の力を借りて教材開発を行っております。なお国際経済学科に関しては、これに加えて英語経済学入門I・IIという科目が2回生になったら入ってきます。これは経済学の専門の先生が、主に英語でレクチャーをされます。たぶん50人くらいのクラスになると思われます。これらを具体的な図にしたのが下の【図4】です。

【図4】 経済学部の英語カリキュラム（2）

| 1回生 | | 2回生 | | |
|-----------------------------|-----|--------|------|------|
| 1st semester | コース | 2nd | 3rd | 4th |
| English R1 English CW1 | 英語 | R2 | R3 | R4 |
| | | CW 2 | CW 3 | CW 4 |
| | | L2 | | |
| English L1 English CALL1 | 二言語 | CALL 2 | | |
| | | R2 | R3 | |
| | | CW 2 | CW 3 | --- |
| | 初修 | --- | --- | --- |

plus 国際経済学科のみ英語経済学入門 I・II

5 評価について——教室レベルとプログラム・レベル

次に評価基準に関してですが、まずアセスメントに関しては以下の【図5】に、また評価基準表を【図6】に示しました。

【図5】 教室レベルの Assessment

| 教室レベル | | |
|------------------------------------|-----------------------------------|--|
| Assessment | | |
| 60% | 30% | 10% |
| Continual Assessment (Quizzes etc) | Examinations Reports Presentation | Extra work (attitude & active participation) |

【図6】 教室レベルでの成績評価基準

| 教室レベル | | | | | |
|-------|----------------|-------|--------|-------|----------|
| 評価基準表 | | | | | |
| | A ⁺ | A | B | C | F |
| UI・IM | 90 以上 | 89-80 | 79-70 | 69-60 | 59 以下 |
| PI | --- | --- | 100-70 | 69-60 | 59 以下 |

これらの目安がどれくらい有効に働いているかについてですが、結果として、Upper Intermediate は A⁺、A、B あたりが非常に多かったですし、Intermediate は B が一番多かったということで、予想以上にうまく分布していたようです。しかしこれを科目別や、あるいは教員別に見ると、多少、ばらつきがあることがわかりました。やはり非常に多くの先生方にいろいろなことをお願いしているので、評価方法に関しては、特にきちんとしたものをこちらから提示することが必要だと常に感じております。

これらが教室レベルの評価の話です。次にプログラム・レベルの成果の測定についてですが、私たちは TOEIC® を活用しています。英語のプログラムが終了するまでには、TOEIC® で経済学科の学生は最低 400 点相当を取得する必要があります。これは最低線ですから、もっと上を目指していただければいいわけです。国際経済学科については TOEIC® 550 点相当が最低の基準点です。しかも、経済学部では、このミニマム・クリアを卒業要件にすることになりました。経済の専門の先生方は最初 670 点や 650 点などの高い数値目標を提示されましたが、私たち英語教員は、470 点前後の数値をだしていました。

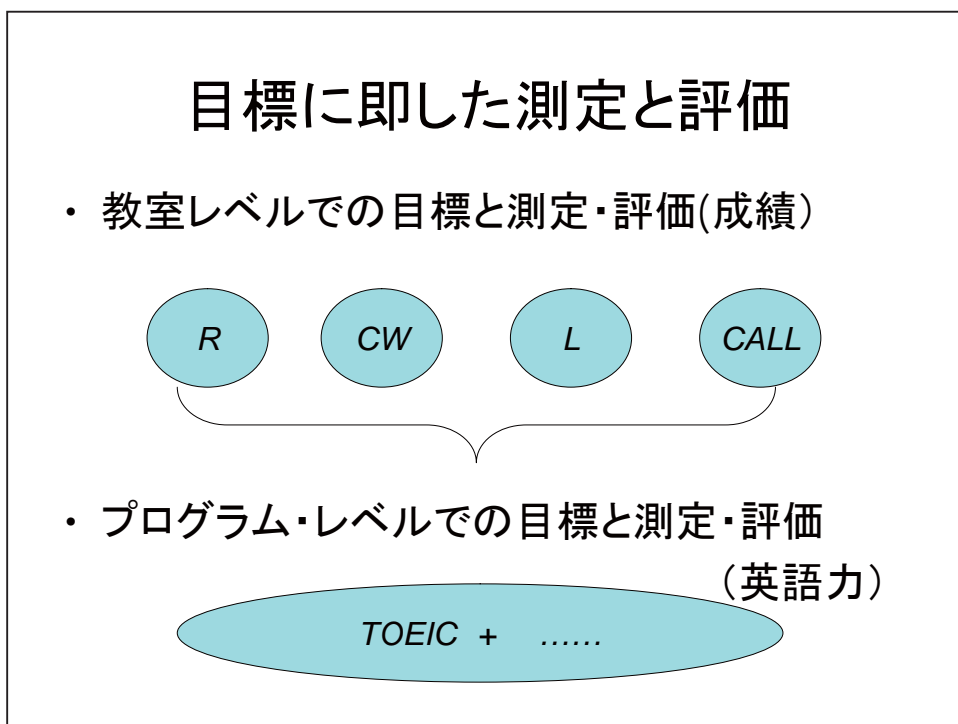
結果、550 点に落ち着いたわけですが、これが今後大変なことになるような気がしております。経済学科ですと、今の段階では、すでに 6 割以上がミニマム値の 400 点をクリアしています。ところが、国際経済学科になると、550 点に達しているのは、6 月の段階では 2 割弱、今の段階でようやく 3 割程度でした。

このように、入試による選抜、プレイスメント・テストによる配置に加えて出口テストという形で TOEIC® が入ってきたわけで、具体的な数字目標を出すということに関しましては客観性があるので大変いいわけですが、下手をすると数字に踊らされてしまう可能性もあります。それから TOEIC® でどれくらい学習効果があったかを見るという点に関して言えば、1 回生に対しては、私自身かなり抵抗がありました。そのため学部の理解を得て、TOEIC® 以外のテストを活用させてもらうことにしました。それは、教育測定研究所が開発運用している「英語コミュニケーション能力判定テスト (CASEC)」と言われているもので、今年度は研究協力校という形で利用しました。このテストで、1 回生の 4 月と 6 月末の比較をしてみました。CASEC は、250 満点の 4 つの下位テストがあり、合計 1000 点満点のテストです。ある私立大学が CASEC で 570 点を目標にしていると聞いたことがあります。Upper Intermediate あたりはそれをクリアしているようです。また、4 月と 6 月末を比べると、すべて統計的に有意な伸びが観察されていたので、プログラムを設計している者としては、少しほっとしました。

6 まとめ

最後にまとめておきたいと思います。「目標に即した測定と評価」ということで、まず教室レベルでの「目標と測定評価」に関して言えば、私たちの場合には 4 種類の科目があり、その科目を通じて学生の成績評価をしているわけです。そしてそれらの科目を学習した学生が、どれだけ英語力がついたかということ判定するために TOEIC® を使っているわけですが、小山先生が先のご発表で、「名工大では統一テストと TOEIC® の相関はなかった」とおっしゃいました。それは大いにありうることだと思います。英語それ自体はできないけれども、与えられたタスクを一生懸命やっていい成績を得る。しかしそれが英語力にはつながっていないという学生もいますし、英語力はあるけれど、さぼっている。だから成績は悪いという学生もいます。だから私はあえて「測定と評価」と書かせていただきました。簡単に言えば、正しく測定するということと、教室レベルでの問題を明確にすること。そしてそれを (測定誤差が少ない状況で) プログラム・レベルで正しく測定してやるということ。さらに言えば、TOEIC® テストはリスニングとリーディングの力を測るもので、ライティング力と言ったプロダクティブ的な力は、直接的には測られておりません。ですからできるだけ多面的に測定してあげる必要があるのではないかと感じています。

【図7】カリキュラムにおける測定と評価



以上で私の発表を終わらせていただきます。