

外国語科目における習熟度別クラス編成方法の課題と改善

— 技能別習熟度を考慮したクラス編成の提案 —

西出 崇（京都外国語大学）

1. はじめに

外国語科目などでは、学習効果の向上、授業運営や効果的な指導などを目的として習熟度別のクラス編成が行われることが多い。特に 2000 年前後に進んだ入試制度の多様化に伴い学生間の入学時点での英語力の差が拡大したことで、外国語科目において習熟度別クラスの検討と導入が急速に広がっている。また管見の限りでも、大谷他(2014)、小笠原(2011)、高橋(1999)、田原他(2001)、勅使河原他(2013)など、習熟度別クラス編成の必要性の検討や、その効果の検証が行われている。

一般に、習熟度別のクラス編成を行うためには、何らかのプレースメントテストを実施し、そのスコアを習熟度の指標としてクラスが割り当てられる。京都外国語大学の英米語学科でも入学時にプレースメントテストとして TOEIC を受験させ、そのスコアによって英語の必修科目のクラス編成を行っている。これに対して、実際に授業を担当する現場の教員からクラス編成がうまく機能していないなどの指摘や意見がしばしば寄せられていた。このような中で、プレースメントテストの実施時期やクラス編成作業の日程など、主に事務運営上の課題からプレースメントテストのあり方を見直す必要が生じ、これを機に現場からの声なども踏まえ、利用するテストやクラス編成上の問題も含めて包括的に検討されることになった。本稿では、このプロセスにおいて明らかになったクラス編成上の課題やそれに対する改善策の提案、および IR が果たした役割などについて報告する。

2. 現在のクラス編成方法とプレースメントテスト変更の検討

ここで検討の対象とするのは、京都外国語大学英米語学科の1年次生の必修科目のクラス編成である。英米語学科の1年次には6つの必修科目が配置され、習熟度別クラスで開講されている。クラス編成は入学時に実施する TOEIC の Total スコアによって、High、Higher-Intermediate、Lower-Intermediate、Low の4レベルに分けられ、各レベル内でさらにランダムに20名程度のクラスに編成される。具体的には、英米語学科の約430名の学生を、High レベルに4クラス、Higher-Intermediate レベルに6クラス、Lower-Intermediate レベルに6クラス、Low レベルに6クラスの合計22クラスに編成している。必修科目のクラス編成は6つの科目で共通となっており、学生は同一のクラスメンバーで必修科目の授業を受講する。なお、科目によっては同一レベルの2クラスを合併したクラスで授業が行われるものもある。

このようなクラス編成は一般的な方法であるように思われるが、先にも述べたとおり授業を担当する教員の一部から、学生に割り当てられるレベルやクラス内の学生間の習熟度の不均衡、それによって生じる授業運営上の問題などがしばしば指摘されてきた。その要因の1つとして、TOEIC による習熟度の測定に問題があるのではないかという推測がな

されていたが、具体的に検証は行われていなかった。そこで、プレイスメントテストのあり方を再検討する際に、利用するテストとともにクラス編成の問題も含めて包括的に検討することとなり、IRがその支援を行うこととなった。

プレイスメントテスト見直しの焦点は、主に事務運営上の課題から TOEIC とは異なるテストに切り替えることにあり、コンピュータを用いて受験するタイプの2つの製品（テストA、テストBと呼ぶ）が検討され、導入に向けて試行的に両テストが実施された。試行的受験は2019年度の新入生を対象に、授業開始から間もない時期に、TOEICに基づいて編成された4つのレベルにおいて各2クラスを選び、それぞれ「テストA」と「テストB」を受験させ、それぞれ約70名のデータが得られた。

この2つのテストによるクラス編成を検討するために、プレイスメントテストの TOEIC スコアと両テストのスコアとの相関、および現在のクラス編成と同様の方法で両テストのスコアに基づいて仮想的にクラスの割り当てを行い、現状の TOEIC に基づいたクラス編成との比較を行った。TOEIC と両テストとの相関係数は、それぞれ「テストA」とは0.757、「テストB」とは0.837となり、いずれも比較的強い相関を示した。また、クラス編成状況の比較でも、現状の TOEIC によるクラス編成と大きな違いは見られなかった。

この結果が示唆するのは、TOEIC による習熟度別クラス編成は代替の両テストを利用した場合と大きな違いはなく、したがってクラス編成の問題は TOEIC に起因するものではないということである。プレイスメントテストの切り替えという点では、いずれのテストを採用しても現状のクラス編成とあまり変化はなく、いずれを導入しても大きな問題は生じない可能性が高いことは示されたが、他方でクラス編成自体の課題は依然として残った。

3. 習熟度別クラス編成方法の問題：技能別習熟度の見落とし

現場の教員が感じるクラス編成の「違和感」の要因が、プレイスメントテストとして TOEIC を用いていたことではないことが示唆されたことから、課題は測定方法ではなく得られたスコアによるクラス分けの方法であることが見えてくる。習熟度別クラス編成の目的は、より効果的な指導や各クラスの授業運営のためにクラスの構成員の習熟度の水準を一定にすることである。現状のクラス編成では、プレイスメントテストで得られた TOEIC スコアをそのまま習熟度の指標として用いているが、そこに見落とされた問題が潜んでいた。

TOEIC では Listening と Reading の2技能を測定し、その合計が TOEIC スコアとして示されるが、総合得点のみでクラス分けを行った場合、Listening と Reading のバランスは考慮されないため、同じ水準の総合得点でも技能別にみれば同一クラス内で比較的大きな差が生じることになる。習熟度別クラス編成の目的はクラスの構成員の習熟度を「等質化」することにあるが、個別技能のスコアが考慮されないために、かえって異質な学生を同一クラスに割り当てることになっていることが明らかになった。具体的には、総合得点と同じ500点だとしても、Listening と Reading のスコアがそれぞれ100点と400点、350点と150点の学生では、技能ごとの習熟度の差は大きい。この技能別習熟度の差が、現場の教員が感じるクラス編成の問題の一端だと考えられる。

プレイスメントテストの TOEIC における Listening と Reading の関係を示す散布図を

図1に示す¹。ListeningとReadingの相関係数を全体でみると0.628で正の相関を示しているが、強い相関とまではいえない水準である。ListeningとReadingの関係をクラス別に切り出し、それぞれに回帰直線を加えたものを図2に、各クラスのListeningとReadingの相関係数を表1に示す²。上位クラスの両技能の相関は正の値を示すが、中間クラスでは負の相関となり、下位クラスでは無相関に近い値となっている。また、クラス内で2技能の差や得意な技能の偏りが大きいことがうかがえる。このようなことが生じるのは、総合得点と同水準の学生は、図1に示す散布図におけるX-Y平面の原点から等距離となる同心円上に位置しており、全体のListeningとReadingの相関が比較的低い集団において総合得点でクラス分けを行うと、クラス内のListeningとReadingが負の相関になりやすいためである。もし両技能の相関が高い集団であれば、総合得点を用いても比較的等質性の高いクラス編成となるだろう。

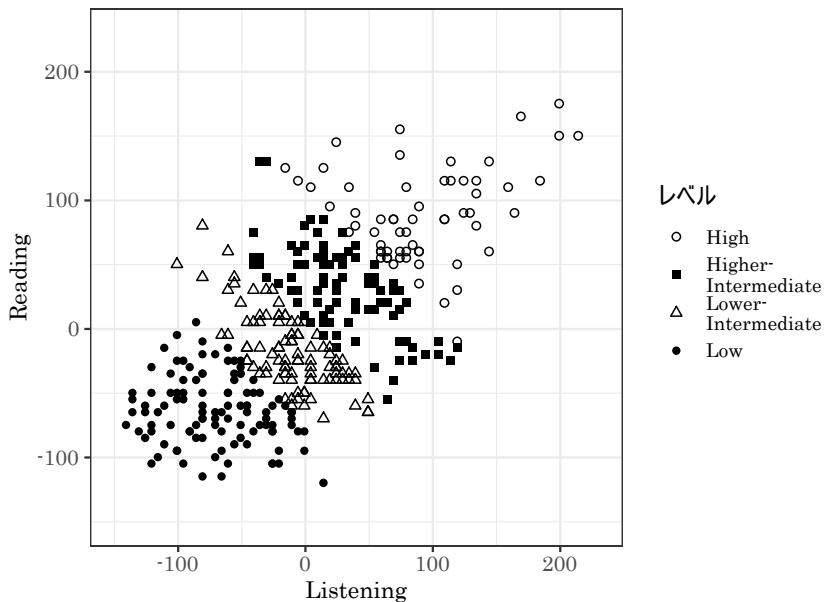


図1 ListeningとReadingの関係

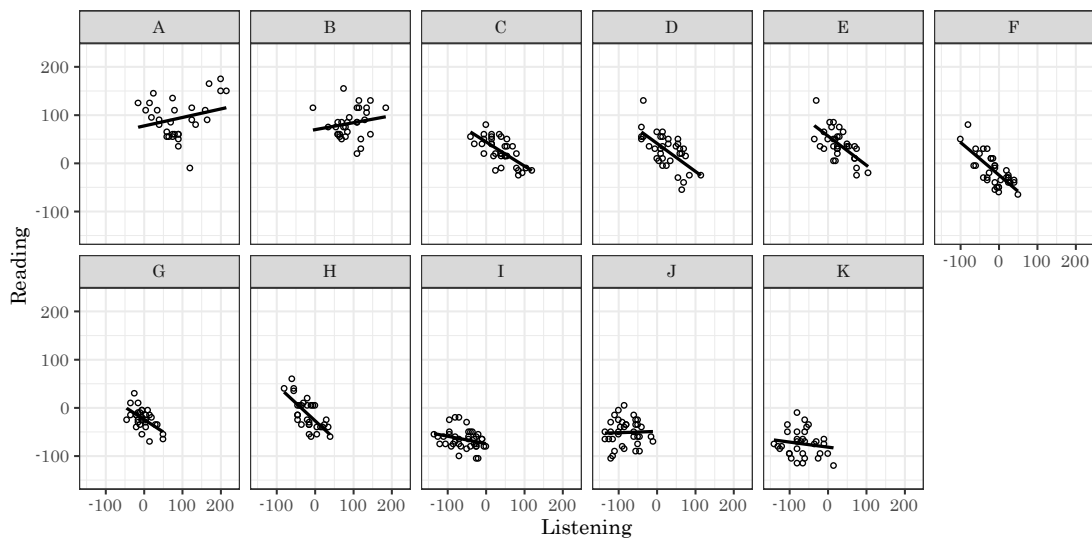


図2 クラス別のListeningとReadingの関係

表1 各クラスの Listening と Reading の相関とクラス内平方距離
現在のクラス

現在のクラス			k-means 法によるクラス提案		
クラス	相関係数	クラス内平方距離	クラス	相関係数	クラス内平方距離
全体	0.628	838814.8	全体	0.628	272414.7
A	0.254	159135.5	A'	0.554	32794.4
B	0.178	72416.1	B'	0.179	20957.1
C	-0.682	74155.7	C'	0.045	27957.9
D	-0.634	97827.1	D'	-0.023	14467.5
E	-0.631	66877.3	E'	-0.290	27791.7
F	-0.731	80958.6	F'	-0.122	20066.0
G	-0.556	29372.6	G'	-0.179	18346.5
H	-0.694	57691.9	H'	0.076	22722.3
I	-0.278	64802.6	I'	-0.139	21522.9
J	0.032	64924.3	J'	-0.027	27661.6
K	-0.150	70653.0	K'	-0.291	38126.7

クラス内の Listening と Reading が負の相関を示すとき、授業運営は相対的に難しくなると予想される。Listening が得意な学生と、Reading が得意な学生という質的に異なる学生が同一クラスに割り当てられることになり、技能別熟度の差が大きな状態で授業を行わなければならないことになる。クラス内の学生の Listening と Reading の相関が正であれば授業運営において習熟度の違いに留意すればよいが、負の相関の場合には得意分野の偏りなど学生の質的な違いにも注意を払う必要が生じる。クラス編成においては、このような技能別の習熟度が考慮されておらず、担当教員もこの点を認識していなかったことが、クラス編成における「違和感」につながっていたと考えられる。

4. 技能別習熟度を考慮したクラス編成方法の提案

問題の要因が明らかになれば、改善案の提案は比較的容易である。ここでは Listening と Reading の両スコアを考慮してより等質なクラス編成を行う方法として、機械学習におけるいわゆる教師なし学習であるクラスターリングを用いる。クラスターリングにも様々なアルゴリズムが提案されているが、ここでは単純なアルゴリズムである k-means 法を利用する。このアルゴリズムは、データの重心に基づいて指定された分類数にデータを分割するもので、本稿に当てはめれば、図1の散布図上の距離が近い点を同一クラスに割り当てるように分割することに相当する。つまり、ここでの Listening と Reading のいずれもが似かよった学生同士を同じクラスにするという目的と一致する。

図3、図4に k-means 法によって現在のクラス数と同じ11グループに分割した結果を示す³。図3はデータ全体、図4はクラスごとに散布図を描いたものである。今回はすべてのデータを単純に指定のクラス数に分割したが、授業運営上の都合でいくつかのレベルに分ける必要があれば、TOEICの総合スコアでレベル分けしたうえで、レベル内のクラス編成をこの方法で分割するなど考えられる⁴。図2、図4から現在のクラス分けと機械学習によるクラス分けを比較すると、明らかに後者の方が各クラス内の学生の Listening と Reading の習熟度が均質であることがわかる。表1右側に示すクラス内の Listening と Reading の相関係数は、現在のクラス分けよりも絶対値がやや小さくなっている。またクラス内平方距離が小さくなり、各クラス内のデータが重心により近く密集していることが確認できる⁵。このようなクラス編成を行うことで、これまでに教員が感じていたクラス編成の「違和感」を軽減し、より効果的な指導に結びつく可能性がある。

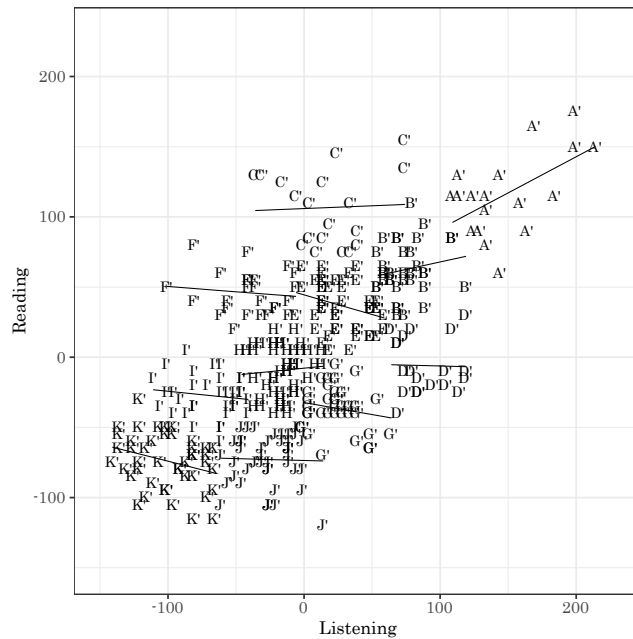


図 3 k-means 法によるクラスタリングの結果

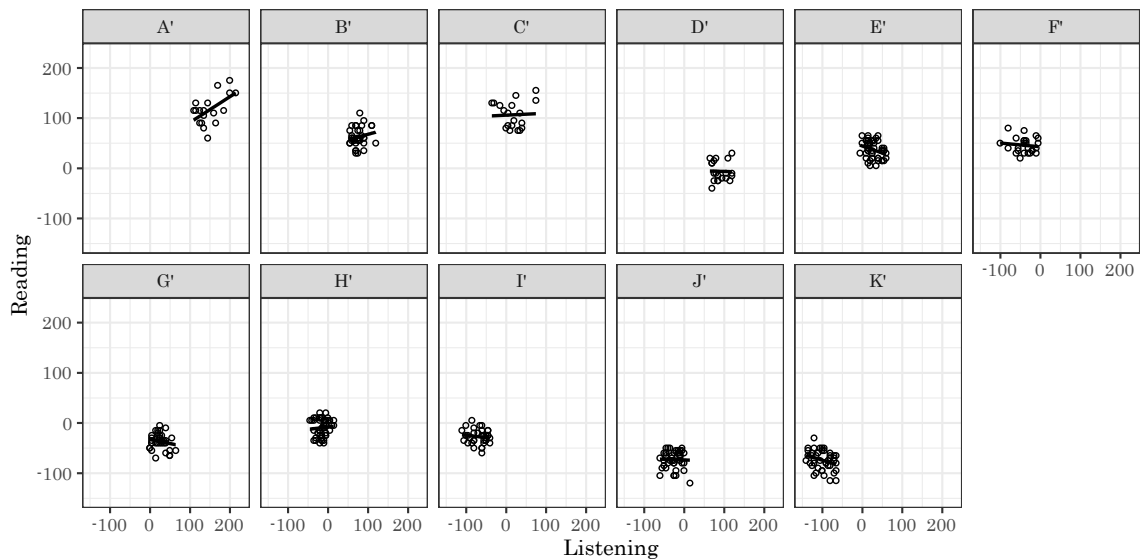


図 4 k-means 法によるクラス別の Listening と Reading の関係

4. むすびにかえて

本稿では、プレイスメントテストのあり方に関する検討のプロセスにおいてクラス分けの問題点を検討し、Listening と Reading のスコアを考慮せずに総合得点のみで習熟度別クラス編成を行っていたため、同一クラスに質的に異なる学生が割り当てられ、現場の教員の「違和感」に結びついていた可能性を示した。またそれに対する改善案として、機械学習によるクラスタリングを用いて Listening と Reading の両者を考慮した等質性の高いクラス編成案を示した。

今回の検証では、IR 担当部署として現場の教員が感じるクラス編成上の「違和感」の原

因を指摘し、それに対する改善案を英米語学科に提示した。これを踏まえて学科では来年度のクラス分けの方法について検討を行っている。ここで提案したクラス編成方法が実際に導入されるのか、それによってこれまで現場の教員が感じていた問題が改善されるのかは、現時点ではわからない。これらの点については、必要に応じて追加的な分析や情報提供などの支援を行いながら、学科での議論を進めている状態である。IRからの提案が受け入れられ、クラス編成方法が変更されれば、改めて効果の検証が新たなIRの課題となる。

既に示したように、ここで指摘したクラス編成の問題点は気づいてみれば比較的単純な構造であるが、データだけを検討しているのみではなかなか発見できないものであった。IRが対処しなければならない問題の多くは、このように探索的に分析の試行錯誤が必要なものである。今回の事例でも、IRにおいてデータを様々な角度から検討したり可視化すると同時に、問題の内容やその要因などについて現場の教員の意見を聞き取ったり、一緒にデータを検討することを通して課題が明らかになった。また、IRからの提案が受け入れられるためにも、こうしたプロセスが重要である。効果的なIRを進めていくためには、やはりIRと現場とが互いに問題意識や専門知識を共有し、双方が密にコミュニケーションを行うことが重要であるといえるだろう。

【参考文献】

- [1]大谷麻美・横山仁視・キム ブラッドフォード ワッツ（2014）「プレースメントテストによる習熟度別クラス編成に関する報告書 ―全学共通言語コミュニケーション科目の英語における事例―」『京都女子大学人文論叢』,62号,pp.27-50
- [2]小笠原真司（2011）「英語習熟度別クラスの効果とG-TELPによる成績分析 ―工学部総合英語IIIのデータを中心に―」『長崎大学 大学教育機能開発センター紀要』,2,pp.9-19
- [3]高橋寿夫（1999）「外国語科目『英語』の改善に向けて ―習熟度別クラス編成について―」『關西大学文学論集』,第49巻第4号,pp.143-162
- [4]田原良子・堀江美智代・竹内光悦（2001）「習熟度別クラス編成に関する考察(1)」『鹿児島純心女子短期大学研究紀要』,第31号,pp.215-244
- [5]勅使河原三保子・上田倫史・林明人（2013）「英語習熟度別クラス編成開始に向けての学内英語レベルの調査(1)」『駒澤大学総合教育研究部紀要』,7,pp.549-562

¹ 各スコアの数値については、データを直接示すことを避けるため、それぞれの平均値が0ポイントになるようにオフセットしている。ただし標準化は行っていない。

² 22クラスをすべて示すと煩雑でクラスごとのデータ数も少なくなるため、ここでは同レベルの2クラス合併で開講されている科目のクラス分けを用いて分析を行う。

³ データの処理はR言語のkmeans関数を用い、クラスタ数以外にnstart=9999を指定した。クラス名は便宜的にクラス内の平均TOEICスコアが高いものからアルファベット順に命名した。

⁴ ここでは新たなクラス編成方法を示すことが主眼であるため検討していないが、k-means法では各グループに割り当てられる人数が一定ではないなど、実際のクラス分けに用いるためには検討すべき課題がある。

⁵ クラスタ内平方距離とは、各グループの重心からグループに所属するデータまでの距離の二乗和で、グループ内のデータの密集度合いの指標である。