

# 監視の要らない

## 中国語教育を目指して

山崎直樹

### 一 はじめに

COVID-19の感染拡大により対面式の授業が不可能になったとき、我々は、我々のおこなっている外国語の授業が「受講生を同一かつ唯一の時間帯に同一の場所に拘束して、教師が監視できる」ことを条件に成り立っていること、つまり、「拘束と監視」を前提になりたっていることを改めて自覚した。そして、授業のオンライン化で「拘束」が難しくなったとき、教員の中のある者は拘束をしなくとも「監視」ができるよう工夫を凝らし、ある者は監視の要らない授業を実践することを試みた。

近い将来、コロナ禍が収束したのち、つまり、「監視」



ができる授業が可能になったとき、我々はまたそのような授業に戻るべきだろうか、それとも、監視を必要としない授業のよりよいありかたを目指して努力を続けるべきであろうか。当然、後者を選ぶべきである。それは、外国語教育の授業として、後者がより理想に近いからである。この文章は、監視の要らない言語教育のほうが望ましいこと、そして、それを実現するためには何が必要かを考える。

### 二 監視の技術の発達

オンラインでの授業実施を強いられた期間、従来のような授業をおこなうことを望んだ大学教員は、オンライン授業での監視の技術……授業中に実施するテストにおいて、

受講生が不正行為をおこなわないように監視する技術を高める必要に迫られた。

ビデオ会議システムを使ったリアルタイムで進行するオンライン授業であっても、受講生が何をしているかを、教員は細かくチェックすることはできない。不正行為の監視は教室での対面授業では容易であるが、オンライン授業ではさまざまな特別な手段が必要になる。

具体的には、部屋の中に誰か他の人間はいないか、机の上に参照不可なものや通信機器がないかなどをチェックする、あるいは、受講生のカメラを、常時、受講生の手元を映させるのに使う……などの手段である。また、不正行為を防ぐための出題方式や時間配分（例えば、不正行為をおこなう時間がないように、個別に細かい制限時間を設けた出題をする）などにも工夫が凝らされた。

しかし、これらの監視を手落ちなくおこなうスキルは、教育機関における教育の本質とは無関係であると思われる。この時間と労力は他に向けられるべきであろう。

### 三 不正行為について

監視の主要な目的は「テストにおける不正行為」を防ぐことである。では、どのような行為がそう考えられているのか。次の(ア)~(エ)は、代表的な不正行為の例である。

#### (ア) 他人の答案を書き写す

教室でのテストであれば、これは書き写される側の意志に関わりなくおこなわれる。オンラインであれば、誰かが意図的に自分の答案を他の受講生に見せなければ、この行為は成立しない。そして、オンラインでは、受講生の誰かが自身の答案の画像や入力したテキストデータを、何らかの通信手段で他の受講生に送信するのは容易である。

なお、これは容易に防げる。要するに、誰がやっても同じ一種類の正解になることを求められる出題であるから、書き写せばOKとなるわけで、ひとりひとりが異なった解答を要求されるテストであれば、この行為は成立しにくくなる（「ひとりひとり異なった内容」については後述）。

#### (イ) 他人と相談する

これも教室では難しいが、オンラインだと容易である。クラス全員が何らかの通信手段を共有していれば、全員で相談ができる<sup>(1)</sup>。

ここで次の疑問を提出しておきたい。授業では、学習者同士が協働してタスク（何かの文章の読解とか、目的のある文章の作成とか）を遂行する活動はよくおこなわれ、しかも、学習者の相互作用により学習効果が上がることはよく知られている。テストにおいて、これをわざわざ不正行為にしてしまうのはなぜだろうか。

(ウ) 参照してはいけなものを参照する

辞書を調べる、教科書を見る、ノートを見るなどを禁止しているテストにおいて、それらを参照する行為である。

これもまたオンライン授業では容易におこなえる。カメラで手元を映すなどの監視を逃れる方法はいくらでもある。

後述するが、「見たいものは何でも見てよい」という条件にすれば、これは不正行為ではなくなる(筆者の周囲ではこの方法を実施している教員も多い)。

筆者は、「学習者のほんとうの力がわからない」「みんなが良い点を取ってしまおう」という二点を、この方法の欠陥として指摘されたことがある。「ほんとうの力」についての議論は後でおこないたい。「みんなが良い点を取る」可能性については、もしそうだったとしたら、出題自体に問題がある。よほど何も考えさせないテスト、何かを丸写しにすれば済むテストで、なおかつ丸写しにする場所がわかりやすいテストでなければ、そうはならない(最後の条件は重要で、入門期の学習者は、教科書などを参照させても、参照すべき個所がわからないことも多い)。

(エ) 使ってはいけない道具を使う

「辞書を使う」がこれに当たる。ウェブ上の無料の辞書は多いが、現在なら、単語単位で調べるのであっても、機械翻訳を使うほうが効率が良い。機械翻訳の精度は日夜、向上しており、主要な言語間の翻訳であれば、使いかたに

よっては実用に耐えうる(なお、現在は、機械翻訳を使わせるかどうかよりも、どのような使わせかたをしたらより効果的かのほうが、外国語教員の間での興味を中心である)。

教室で授業をしているも、受講生が情報機器を使ってわからないことばを調べる姿は、もはやありふれたものになった(オンライン授業であれば、さらに頻繁におこなわれているだろう)。いっぽう、テストにおいては、機械翻訳の使用を認めないのがふつうである。しかし、機械翻訳の使用を許可して伝統的な筆記テストを解答させたら、学習者はどのような使いかたをするのか、たいへん興味深い。例えば、次のような問題があったとする。

日本語の意味になるよう、「」内の語句を並べ替えよ。  
わたしは昨日三時間中国語を勉強した。

〔一小時、昨天、漢語、了、我、三個、學〕

(以下は、筆者の見聞に基づく) ある学習者は、この問題をどう解いたかというところ、機械翻訳に出題文の日本語をそのまま入力として与え、出てきた翻訳をそのまま解答とした。その翻訳には出題中の中国語とは異なる語彙が使われていた。しかし、そのようなことは意に介さないようであった。この課題解決の手順と「」内の語句を「という指示を無視する態度はとても興味深い」。

なお、教師の側は、機械翻訳を使って解答をすることよりも、機械翻訳の結果が不適切であっても学習者が意に介

さない（＝わからないのではなく、検証をしようとしな  
い）ことのほうを嘆く。しかし、この行動には無理からぬ  
ところがある（詳細は山崎 [2018] 参照）。

先に、機械翻訳は使いかたによっては実用に耐えうると  
書いた。一例を挙げる。以下は台湾留学中に救急車で搬送  
された学生の経験談である。その学生に応急処置をした救  
急隊員は、相手が日本人とわかると、スマートフォンを取  
り出し、それに向かって華語で何かを話し、その画面を見  
せてくれた。そこには、Google 翻訳が出力した日本語「こ  
れから点滴をします」が表示されていた……このように、  
機械翻訳のような「エイド」を使ったコミュニケーション  
は必要不可欠にして日常的なものになっている。テストの  
ときだけその使用を排除するということは、そのようなテ  
ストがいかに日常と乖離しているかを示している。

#### 四 監視の必要なテスト

##### (一) どのようなテストに監視が必要か

もちろん、世間には「監視が必要なテスト」が存在す  
る。公的教育機関の入学試験、TOEFLなどの外国語能  
力試験はその代表である。この種のテストが厳重な監視を  
必要とするのには次の二つの理由がある。

##### (1) ハイステイクス・テスト hi-stakes test である。

(2) 誰がやっても同じになる唯一の正答が存在する。  
hi-stakes test とは、テストや評価の結果が、生徒の将来  
の進路や学校の評価となり、社会全体の注目を浴びるよう  
になるテストのことである [ギップス 2001: 252]。

TOEFL のようなテストの聴解・読解問題は、多肢選  
択式出題なので正解は一つに限定される。当然、個人的背  
景によって正答が変わるといったこともない。

なお、前述の(1)は不正行為への強い動機づけとなり、(2)  
は不正行為をしやすい条件を提供する。よって、この種の  
試験では厳密な監視が必要になる。

##### (二) まねをする必要があるか

大学の授業でおこなわれる「授業の達成度を測るテスト」  
は、前述の監視が必要な hi-stakes test と同じであろうか。

たしかに、成績の良し悪しが受講生の進路に影響を与え  
ることはある。しかし、あるコースの最終成績の重要性  
と、そのコースがおこなうテストの重要性は同じではな  
い。そのコースでは期末のたった一回のテストだけで、他  
は何も考慮せずに最終成績を出すのであれば同じになる  
が、そのような授業は珍しいであろう。

通常の授業でおこなうテストは、複数ある評価手段の一  
つである（そうあるべきである）。よって、世上の hi-  
stakes test とは異なるのであるから、hi-stakes test のような

監視をする必要があるかどうかを考えてみるべきである。

不正行為をせよ／見逃せというのではない。第三節で見たように、不正行為とされている行為の中には、教室での活動や日常のコミュニケーションにおいて、ごくふつうにおこなっている行為も含まれている。それらが不正行為とされるのは、テストという限られた文脈においてである（そこには、ある種の言語能力観が働いているが、それについては後で触れる）。日常とは乖離したそのような評価方法をまねる必要があるのかということを考え直してみたい。

### (三) 監視を必要なくするために

監視をやめるためにはどうすればよいか。単純に考えると、まず次の方法を思いつく。

(1) テストはこれまでどおりで、監視をやめる。「不正行為」を咎めない。

(2) テストはこれまでどおりの出題で、どのような補助手段（参照物、道具）を使ってもよいことにする。

(3) 従来型のテストをやめて、他の評価手段を採用する。

もちろん(1)はよくない。理由はいろいろあるが、主要な理由は、GPA (Grade Point Average) 制度を導入している場合、成績評価が集団準拠型であるからである。明示的な到達基準を定め、そこに到達したかどうかで成績を評価する標準準拠型の評価であれば、他人の成績は関係ない。

しかし、集団準拠型の評価はその学習者が集団の中でどの位置にいるかで評価する。GPA制度では、成績の各等級について、それを与える割合に制限を設ける（でない）、GPAの数値が意味をもたない）。つまり、自分への評価が良いだけでは良い成績が得られない。他の学生への評価が低いことが必要である。

例えば、「クラスの誰と相談してもよい」という条件でテストを実施するとする。「相談」への参加者各々が提供する情報が同量で、議論に同等の貢献を為し、それらにより受け取った恩恵が同量であるなら、問題はない。しかし、学習者のレベルに差があれば、与えるものと受け取るものを正答の個数で測れば差が生じるであろう。ましてや、ただ人の答案を書き写すだけの「タダ乗り」をする者がいれば、相対評価のもとでは、当然、不公平感が生じる。(2)を実施するなら、「これまでどおりの出題」をやめなければならぬ。従来型のテストは、初級レベルの場合には、条件に合った単純な言語形式を生成させるものが多い（空欄補充、語順整序、日文中訳など。「何を見てもよい、何を使ってもよい、誰と相談してもよい」という条件にすると、それを禁じていたときとは、測定対象の能力を構成している要素が異なる。それなのに出題の内容が同じというわけにはいかない。従来は、学習者に何も考えさせない、創造させないテストであっても、記憶力を測るこ

とはできた。「何でも参照可」ではそれすらも測れない。

(3)については、筆記テストなどの従来型の評価方法に代わる新しい評価方法……パフォーマンス評価、ポートフォリオ評価などを考えることができる。これらの評価方法に監視が必要かどうかは、後で検討する。

## 五 不正行為がやりやすいテスト

ここでは「不正行為」を実行しやすいテストの特徴を整理しておきたい。繰り返しになるが、以下の条件を備えているテストは「やりやすい」。

(1) 誰がやっても等しく同じ唯一の正答にたどりつくことが要求される。

(2) 解答が選択式であるか、単純な言語形式のアウトプットである。

例えば、次のような出題がそれにあたる。

(a) 日本語の意味になるよう( )に適切な語を入れよ。

あなたのアイスは冷蔵庫にあります。

你的冰淇淋( )冰箱裡。〔有、在、是〕

(b) 日本語の意味になるよう「」内の語句を適切な順に並べ替えよ。

わたしは毎日、自転車で学校に來ます。

〔騎、學校、每天、自行車、我、來〕

(a)も(b)も、記憶の中にある言語形式と、出題された言語形式のマッチングを要求する再認テストである。「」の選択肢をなくせば、記憶の中の言語形式を自分で表出する再生テストになる。一般に、再認テストよりも再生テストのほうが難度が上がるといわれているが、このような単純な言語形式の場合は大差ないであろう。いずれにせよ、「回答が選択式であるか、単純な言語形式のアウトプットである」出題である。

また、このような出題に対する解答は学習者の個人的な背景などによって変わるものではない。「誰が解答しようと同じ正答にたどりつくことが要求される」出題である。なお、出題内容によっては、この形式でも正答候補が複数考えられるが、それにせよ少数の有 limit の候補に過ぎない。

大学構内の地図を用意し、「構内でコンビニの場所を尋ねられた。コンビニは二号館の地下一階にあるが、二号館自体が見つけにくい建物なので、その場所も含めて説明せよ」というようなタスクを与える出題も考えうる。しかし、これにせよ、大学の第二外国語の対象となるレベルの学習者の場合、解答に使用する言語形式は決まった形式の短文(の羅列)である(××在××的地下一層/××在××和××的中间……)。というよりむしろ、それで済むような課題にすべきである(目印を教えるとか、道案内をするとかの解答まで要求するとしたら、授業での学習範囲を

容易に超えてしまう)。

要するに、入門(初級レベル)の範囲で筆記テストをおこなう場合、単純な文法的/語彙的知識を問う出題であれ、あるいは、コミュニケーションタイプなタスクを遂行する出題であれ、要求される解答の言語形式自体は、単純な言語形式のアウトプットであるということである。

## 六 我々はどんな能力を測るべきか

(一) 言語に対する知識、言語を運用するための知識を測る

学校の授業では、受講生の能力を測る何らかの活動をして成績を判定する。筆記テストもその活動の一部である。では、一つのコースの成績を出すには、どのような能力を測ればよいのか……それはそのコース終了時に到達すべき目標によって決まる。決められた目標をどの程度達成したかを測ればよい。であるから、筆記テストだけで評価するコースは、その筆記テストで測れる能力だけが目標に達したか否かの指標になる。

入門レベルの外国語の授業の筆記テストは、語彙の意味や機能に関する知識を測ったり(空欄補充問題に代表される形式で測る)、文法的に適切な文を組み立てる能力を測ったり(語順整序問題に代表される形式で測る)するの

がふつうである。少しひねったところでは、「こんな場合にはどう言いますか」のような質問により、特定の状況下の言語使用の知識を問うたりもする。

ただ、「こんな場合には……」系の出題には二種類あり、焦点となる言語形式(特定の文型や慣用表現)をターゲットにしてその産出能力を見るものと、そのようなターゲットはなく、純粋に課題解決の能力を問うものである。後者は、設定された状況や解答者に要求する内容によって、むしろ、次の(二)で扱う出題に近い。

(二) コミュニカティブ・タスクを遂行する能力

近年は、外国語学習の目的として、コミュニケーション能力の育成がよく掲げられる。伝統的に掲げられてきた「言語の構造に関する知識の体系的構築」は、「コミュニケーション能力」を下支えするものとして語られることが多い(各校種の『学習指導要領』『外国語』編を参照)。

さて、コミュニケーション能力(厳密に言えば「言語を使ったコミュニケーションの能力」)はどのように測定したらよいか。現実の世界におけるコミュニケーションは、始まりや終わりが明確でなく、また、目的やゴールが明確でないこともあり、テストのためにその一部を切り出すことが難しい。

そこで、言語コミュニケーション能力を測るために、目

的やゴールが明確で、なおかつ解決すべき課題が明確で、それらの遂行のためにことばをやりとりする必要がある。「コミュニケーションなタスク」を設定し、そのタスクを遂行できたかどうかをコミュニケーションをおこなう能力とみなすことが多い。つまり、「現実世界における特定の状況における課題解決をシミュレーションした」タスクの遂行能力と、「そのタスクに類似した現実世界の状況」におけるコミュニケーション能力とを、同一視するわけである。

ここで重要なのは「現実世界をシミュレーションした」という点である。現実の世界で何かしらの課題を解決しようとしたとき、人は、使える道具は何でも使い、協力を求められる相手がいれば協力を求める。このような道具を使いこなす能力・協働作業をおこなう能力も、あるいはそれをおこなうことを肯定する態度も、それらを現実世界の課題解決の能力の一部とみなすことはごく当然である。

であるのになぜ、外国語学習における課題解決（とくに評価のために何かのタスクを遂行させる場合）において、道具の使用や周囲との協働作業が不正行為とされるのだろうか。このおかしさをもう一度考える必要がある。これについては第八節で触れる。

### (三) 言語を学ぶための知識やスキル

近年は、外国語学習に成功する学習者はどのような特徴

(性格、認知特性、行動習慣、動機づけ……などの特徴)をもっているかという研究、つまり、学習者個人の内部にある要因を考える研究も盛んである。その成果として、外国語学習に成功する学習者は、自分自身の特性に合ったいろいろな学習方略を組み合わせて使用している、メタ認知方略を有効に使っているなどの特徴をもつ事例が多いことが指摘されている。さらに、習得目標言語の音声、表記、語彙、文法、語用論的習慣が、既習得の言語(例えば母語)とどこが同じでどこが異なるかを考察することができ(メタ言語的認知の発達した)学習者の強みも指摘されている。

つまり、簡単にいうと、我々が効率的に外国語を習得するためには、「何を学習するか」の「何」、つまり、習得目標言語の構造と運用の知識などを学ぶことも大切であるが、それと既習得の言語とを対比することができるメタ言語的認知の能力(この能力の評価をどのように学習に組み込むかについては、第十一節で詳述する)、およびそれらをどのように学習するかという方略、そして、それらを全部含めて、自分の学習をどのようにコントロールしていくかという方略も重要であるということである。

であれば、授業でもそれらに対する知識・スキル・態度を涵養し、それらを身につけることを到達目標の一部にしてもよいはずである。そして、それを到達目標の一部にす



るのであれば、当然、成績評価の対象とすべきである。

## 七 「不正行為」をなくすために

以下では、第三節で概説した、テストにおいて不正行為とみなされる行為を不正行為でなくなるようにする方策と、そのような行為をすることを無意味にする方策（不正行為をチェックする方策ではない）を紹介したい。

それは、簡単にまとめると次の三種類である。

- (1) エイドのあるコミュニケーションは、より現実世界に即したコミュニケーションと考え、エイドを使う能力もコミュニケーションなタスクを遂行する能力の一部と考える。

- (2) 「丸写し」が無意味になるよう、個人化した課題を与える。

- (3) 「丸写し」が無意味になるよう、目標に到達するまでの過程を把握する。

以下の節においてこれらを順に紹介していく。

## 八 「エイド」を使って成立する

### 「コミュニケーション」

第三節で示した「他人と相談する」「参照してはいけな

いものを参照する」「使っではない道具を使う」で例示した行為を不正行為でなくする方法をここで扱う。以下では、これらの行為における「相談する他人、参照するもの、使う道具」をまとめて「エイド aid」と呼ぶ。ここで主張は次のとおりである。

・ エイドを使って成立するコミュニケーションはコミュニケーションの一形態であり、現実の世界ではふつうに存在する。

・ 「どんなエイドをどう使うか」という知識やスキル、エイドを積極的に使ってコミュニケーションを成功させようという態度もコミュニケーション能力の一部である。

・ エイドを相互作用的に使って課題を解決していく過程自体が言語の学習にとって有益である。

### (一) なぜエイドを嫌うのか

なぜ、テストでは、「参照してはいけなものを参照する」「使っではない道具を使う」ことが嫌われるのか。

筆記テストの不正行為を懸念する教師に対して、「見たものは何でも見させ、使いたい道具は何でも使わせればよい」と主張すると、たいていは否定的な反応が返ってくる。最も多い反応は、「エイドを使った解答は実力で出したものではない、そのような答案ではへほんとうの実力」

は測れない」である。

このような文脈で使われる「ほんとうの実力」という表現の意味は、比喩的にいえば「裸眼視力」のようなものである。……裸眼視力が重要になる場面が日常の生活の中でどれだけあるだろうか。

車の運転を許される視力の条件は裸眼視力ではない。矯正視力（補助器具を使った視力）である。なお、現在では、飛行機の操縦でも、裸眼視力であれ矯正視力であれ、視力が一定以上あればよいし、JAXAの宇宙飛行士の募集要項にも同様の趣旨が書かれている。大学の第二外国語で外国語を学ぶ者が、その言語を、宇宙飛行士以上にクリティカルな場面で使うことを要求されるとは思えない。

## (二) エイドのあるコミュニケーション

われわれは日常ではさまざまなエイドを使ってコミュニケーションをしている。わからないことがあれば辞書を引く（今日では機械翻訳が辞書に取って代わりつつあるが）。とくに、メールを読む／書くなどの非同期的コミュニケーションにおいて、この手段は多用される。聞き取れないことがあれば書いてもらうなどの方略も使う。物の名前がわからないとき、スマートフォンなどで当該の物の画像を示すこともある。救急車の隊員がGoogle翻訳を使った前述の例からもわかるとおり、エイドを使ったコ

ミュニケーションは我々の現実のコミュニケーションでは日常的であり、エイドをうまく使うこともコミュニケーション・スキルの一部である。

## (三) コミュニケーションの相手もエイドのうち

エイドは「物」だけではない。コミュニケーションの相手もエイドたりうる。我々はわからないことがあれば相手に聞き返し、ことばの意味を確認する。自分が伝えたいことが出てこないときは、言い換える、身振り手振りを使う、相手に助けてもらうなどの方略を駆使する。

## (四) コミュニケーション方略

(二)と(三)で述べた方略は、いわゆる「コミュニケーション方略」である。

コミュニケーション方略とは次のように定義される。

- ・十分に習得していない第二言語で意思疎通をするときに起こるさまざまなトラブルに対処する技術（古典的な、ややネガティブな定義）

・どのようにしたらより効果的なコミュニケーションを達成できるかを話し手と聞き手が追求するときに使う

### 技術（前向きな定義）

次のような種類があるといわれている。

- (1) 回避（話題を変える、メッセージを出すことを放棄

するなど)

- (2) 言い換え(類似表現、バラフレーズ、造語、直訳……)
- (3) 他言語による穴埋め(意図的なコード転換)
- (4) 非言語表現(身振り手振り)
- (5) 他からの支援に頼る(ネイティブ話者、辞書……などによる支援)

このうちの(5)の方略を使うのが、本稿でいうところの「エイドを使ったコミュニケーション」である。

言語学習におけるコミュニケーション方略については、「そんな方略に習熟しても、それは言語そのものの習得には繋がらない」という見解もある一方、「言語を習得するということを、言語を使ったコミュニケーションにより何かの目的を達成する能力を身につけることだと考えるなら、コミュニケーション方略の能力は言語コミュニケーション能力の一部である」という考えかたもある。

#### (五) コミュニケーション方略と学習方略

前述したように、コミュニケーション方略の能力を重要視しない立場もある。しかし、そのような立場の者も学習方略の重要性は疑わないであろう。ところが、学習者の立場からすれば、コミュニケーション方略と学習方略とはさほど違わない。

筆者が、ある授業において、「コミュニケーション方略

と学習方略が相互に関連していることを、具体例を用いて説明せよ」という課題を出したところ、複数の受講生が、「ことばのやりとりにおいて、理解できないことばや、言語表現がわからなくて伝えられないメッセージがあるときは、コミュニケーションの相手に支援が欲しいというサインを出す」という方略を使うという経験を挙げ、そして「そこで得られた支援が、自分の言語表現を豊かにするなど、新たな知識の獲得につながる」ことを指摘してくれた。

#### (六) エイドを使ったコミュニケーションでの成果を評価する

ここまで述べてきたことを考えると、「他人と相談する」「参照してはいけないものを参照する」「使っていない道具を使う」などの、これまで不正行為とされてきた行為を不正でなくするのは簡単である。「学習者同士で相談する、教科者やノートなどを参照する、辞書を使う」などのエイド使用を認めればよい。

ただ、第四節(三)で述べたとおり、それらの使用を認めるだけで、出題の内容が旧来の筆記テストのままであれば、それは「エイドを使ったコミュニケーションをおこなう能力も含めた言語コミュニケーション能力」を測定するテストにはならない。なぜなら、旧来の入門く初級レベルのテストによくあるように、単純な言語形式をアウトプットさ

せるだけのテストであれば、エイドを使って答えを出す過程には「気づく、考える、創造する」のような過程がないからである。そのような過程がなければ、コミュニケーション能力という、知性と感性を総動員する必要がある高度な能力を測っているとはいえない。

要するに、エイドを使ってコミュニケーションをする能力をコミュニケーション能力の一部と見なすゆえにエイドを使うことを認めるのであるから、テスト自体がコミュニケーション能力を測るものでなければ、エイドを使って「誰がやっても答えが同じになる単純な言語形式」をアウトプットするだけのテストで終わってしまう。それはコミュニケーション能力を測るものではない。

## 九 他人の答案を丸写しにする

他の学習者と相談をして、自分の頭でも考える……のであれば、その過程自体が「学習」である。しかし、何も考えずに他人の解答を丸写しにするだけでは学習にならない。そして、伝統的な筆記テストでは、これができてしまうことが問題である。

なお、前節では「さまざまなエイド」の例としては機械翻訳を挙げなかったが、機械翻訳というエイドの使用は、それだけでは「他人の答案を丸写しにする」に似ている。

もし機械翻訳の出力をそのまま無検証で使用するのであれば、そこには当該の学習者本人に何かを考えさせる過程がないからである。

余談だが、学習者が機械翻訳を使うことに対して教師が抱く不満の一つに、「学習者は機械翻訳の結果が適切かどうか検証しようとしめない」がある。しかし、これは無理もない。筆記テストにおいて学習者が他者の答案を写そうとするとき、いちいち適切な解答かどうかを検証するだろうか。そうではなく、「自分が答えられない出題に答えている他人の答案」が「たまたま見える」から写すのである。

さて、機械翻訳の利用はいったん置いて、「他人の答案を丸写しにする」を無意味にするためにはどうすればよいか。それは、学習者ひとりひとりが、自らの気づき、考察、創造によって、異なった解答を出すことを要求される課題により回避できる。これについて次節で説明をする。

## 十 アウトプットを個人化するための方策

前節では、エイドを使ったコミュニケーションの能力を評価の対象とすれば、「不正行為」を成立させていた要件がなくなることに、しかし、伝統的な筆記試験の内容のままエイドを認めるだけでは不十分なこと……他人の答案を丸写しにする、機械翻訳のアウトプットをそのまま利用す

ることのできるテスト、つまり、学習者に何も気づかせない、考えさせない、創造させないテストは、「言語コミュニケーション能力」を測る評価とは程遠いことを述べた。

これらの問題を解決するためにまずすべきことは、評価のためのアウトプットの内容を個人化することである。以下でそれを説明する。

### (一) 個人化されたアウトプット(1)

「筆記で答えるテスト形式」に限定しなければ、アウトプットを個人化すること自体は難しくない。例えば、発音のスキルを見るのであれば、課題を与え、それを音声でアウトプットさせて録音したものを提出させればよい。その課題をプレゼンテーションにして、動画で提出させてもよい。学習者が創造的に内容を構成できるプレゼンであれば、おのずから内容が個人化する。

ゼロからの中国語学習歴が二〜三カ月程度であっても、教科書の最初の数課の学習項目をうまく組み合わせれば、「自分のことを語る」くらいは可能である。また、箇条書き形式のテキストを読み上げるのであれば、談話構成を考へることなく同じパターンの文型を繰り返すだけなので(例:「○時にくをします」という形式で一日の日程を紹介する)、さほど複雑な言語構造の知識は必要ない。また、入門レベルの学習者に創造的なプレゼンテーションを求め

た場合、学習者が採るアプローチは「プレゼンの内容を日本語で考え、それを機械翻訳で中国語に翻訳する」である。学習者がこのようなアプローチを採る以上、機械翻訳を使うことを禁ずるのではなく、それを使うことが学習に役立つような課題にすべきであろう。

昨今では、機械翻訳を使ったライティング支援の研究が盛んにおこなわれている。一連の研究の成果によると、(i) 伝えたい内容を母語で考え、(ii) それを機械翻訳で目標言語に翻訳し、(iii) その翻訳がほんとうに自分の伝えたいことを伝えられるかチェックをする……という手順でおこなわれるライティングは、初級の学習者であっても多くのことを気づかせる効果があるようである [Garcia and Pena 2011]。とくに、初級の学習者であっても学習者相互でその出力の適切さについて議論する過程は、メタ言語意識を高める効果があるという報告もある [Lee 2021]。重要なのは、機械翻訳というエイドを使ってもなお、学習者に「気づかせる、考えさせる、創造させる」ことを促進する課題を考案することであろう。

なお、確かに本人が発音していることや、メモを見ずに暗唱できているということを確認するためだけに、指定したテキストを読み上げた記録を動画で提出させる方法もあるが、そうなるとこれはもはや「監視」である。

## (二) 個人化されたアウトプット(2)

発音のスキルを確認するために音声ファイルを提出させる、プレゼンテーションを記録した動画を提出させるなどのアウトプットではなく、筆記テストの形式の中で、学習者に「個人化された」アウトプットをさせたいという要望もあると思われる。筆記テストを評価するのは、プレゼンテーションを評価するのに比べ、ずっと易しいし、時間も労力もかからないからである。

次の(1)の出題に比べ、(2)は個人化された内容の解答を要求しているように見える。

(1) 次の日本語の内容を中国語で表現せよ。「わたしの家は四大家族です。父、姉、わたし、妹です」

(2) あなたの家は何人家族か、家族構成はどうなっているか、中国語で表現せよ。

しかし、(2)はさほど効果的ではない。まず、学習者は、自分の中国語の能力が(2)に答えるのに不十分だと考えた場合(例えば、「姉」に相当する中国語を忘れた場合)、どのような行動を採るか。現実のコミュニケーションであれば、いろいろな方略を駆使して、それを伝えようとするだろう。しかし、何も参照できないテストであれば、最も簡単な解決方法は「姉がいないことにする」である(コミュニケーション方略の分類でいえば「回避」)。

もちろん、現実のコミュニケーションでも「回避」はよく見られる。しかし、「言えないから、言う内容を変える」という方略は、不誠実な態度であるため、相手に好印象を与えない(Crizeの会話の公準にも違反している)。現実のコミュニケーションでなら(IIテストでなければ)他の方略を駆使したかもしれないのに、テストという状況が、あまり推奨できない「回避」という方略の使用を強いるという可能性はおおいにある。

もうひとつ問題がある。「姉」に相当する中国語は「媽媽」だと、学習者が誤って記憶していたらどうなるか。前述の(2)の解答として、「我家有四個人、爸爸、媽媽、妹妹和我」という中国語が出された場合、採点する教師はそのままスルーするであろう。この出題形式では、学習者が「姉」を「媽媽」だと記憶していることは気づきやうがない。「あらかじめ、家族構成のデータを提出させておいて、それと照合する」という方法もあるが、それは「監視」と五十歩百歩である。

## (三) 形式と意味の結合

言語構造に関する知識を身につけるといえるのは、つまるところ、形式と意味の結びつけを適切におこなうことである。文法的な知識を問う出題であってもそこは外せない。産出目標となる文の伝える内容を指定せずに「        」内

の語句を意味がとおるように並べ替えよ」という指示しかない出題、肯定文を提示して「否定文にせよ」という指示しかなかった出題をときおり見かけるが、これらはすべて、使われている言語形式の意味を学習者が正確に把握していなくても答えられる出題である。

学習者だけがもっている情報を、学習目標言語で表現させるのはけっこうだが、表現したい内容を適切にアウトプットできたか検証する仕組みがない出題は、「形式と意味の結びつき」ができていくかどうか検証できない。

検証する仕組みは、入門レベルであれば、例えば以下のような出題で実現できる。

・あなたの部屋の様子をイラストで示せ。そこで描かれた「物」に（漫画のセリフのような）吹き出しをつけ、中国語でその名称を書け。

このような出題でも、中国語の名称を知らない「物」は描かないでいくという〈回避〉は発生する。しかし、これは、前節で述べた中国語で表現できない家族はいないことにするという〈回避〉と同じではない。

家族構成の場合は、表現すべき項目はすでに決まっています、そこから、学習目標言語で表現できない項目があれば減点していくという出題であろう（＝減点方式）。部屋のものを描写するのは、自分が表現できる物だけを選んで積み上げていくという「加点方式」の出題である。伝えた

い内容を主体的に選択するという点で、コミュニケーションに対する態度が異なる。

なお、この「部屋の中の物」のような課題をオンラインでおこなう場合、学習者がどのようなエイドを用いて課題を遂行しようとするかについても、当然、想定しておく必要がある。そして、エイドを使ってこの課題を遂行しようとした場合は、学習者が何かを考え、何かに気づく余地があるかどうかで、この課題の価値が決まってくる。

## 十一 「気づき」を評価する

### (一) 「気づき」を評価することの意義

アウトプットを個人化するための方策としてさらに考えられるのは、習得目標言語の何らかの言語形式をアウトプットさせるのではなく、学習の過程で何に気づいたかを報告させる課題である。

・学習項目となっている習得目標言語の言語形式はどのような規則で構築されていると「思ったか」

・学習項目となっている習得目標言語の言語形式は、既習得の言語の相当する言語形式と、どこが共通しており、どこが異なると「思ったか」

この「言語形式」は音声でもよいし、文法構造でもよい。このような課題をおこなうのは、学習目標言語を使った

コミュニケーションのスキルを上達させるという趣旨の授業とは異質であるように見える。それは、かつては一般的で、現在では否定的に扱われる「文法を宣言的知識として蓄積する」タイプの授業ではないか……という疑問も出てこよう。しかし、前述の課題は「正しい知識を一方的に伝授する」課題ではなく、自分で考え、気づくことを促す課題である。現在の言語教育では、メタ言語的意識を高め、メタ言語的な気づきを促進させる作業を重要視する。成功する第二言語の学習者は、しばしばこのような能力に秀でていることがあるのを、多くの研究が報告している。

## (二) メタ言語的気づきを促進する課題

メタ言語的意識を高め、メタ言語的気づきを促進する課題には、いろいろなアプローチが考えられるが、基本的には、ある言語形式Aとそれと何らかの関係がある言語形式Bを比較するのがよい。AとBは同一の言語語であつてもよいし、別の言語語であつてもよい。

以下では機械翻訳を使った課題の例を紹介する。比較する言語形式は教師があらかじめ選定して与えてもよいのだが、検討すべきデータを学習者に選択させ、機械翻訳を使って動的に構築するのがよい。教師が与えたデータであると、クラス内で同じ言語形式を検討することになり、それこそ「他人の答案を写す」という行動を誘発しかねない。

### 【課題の例】

- (1) 「わたしは××が好きではない」という意味の中国語(C1)を書く(例…我不喜歡××)。
  - (2) C1の文末に「了」という助詞を加え(C2、例…我不喜歡××了)、それを自分で日本語にする(J2)。
  - (3) 三種類の機械翻訳にC2を入力として与え、三種類のJ2を得る。
  - (4) 機械翻訳による三種類のJ2を相互比較する。また、自身のJ2とも比較する。そして、どこがちがうかを見つける。
  - (5) 四種類のJ2のうちどれが最も適切かを選ばせ、そう思う根拠を述べさせる(この助詞はどんな機能をもっているか)を直接尋ねていないことに注意。その種の質問は教科書の説明の丸写しを誘発する)。
- 文末の助詞「了」は、動作動詞を述語にした文を例として導入されることが多く、この場合は日本語のテンス／アスペクト助詞「た」とのちがいがわかりにくいいため、往々にして、学習者はこの助詞の機能を「た」のテンス的側面と同一視し、その後の中国語学習に負の干渉をもたらす一般化をする。その一般化が不適切であることを気づかせるためには、状態性の高い述語をもった文に「了」をつけることのような意味になるかを検討させるのがよい。
- この課題の評価の観点は「どれが最も適切かを選ばせ、



そう思う根拠を述べさせる」の部分であり、適切なものを選べたかどうかではなく、根拠が正しいかどうかでもない（そのような評価基準は「他人の答案の丸写し」を誘発する）。つまり、機械翻訳の出力のちがいをどこまで考察できたか、正しい出力を選ぶ根拠をどこまで具体的に述べられたか……などにする。

機械翻訳は、入力と同じであっても、昨日と今日とでは翻訳結果が異なるし、同じ構文（例：我不喜歡××）であっても、××の部分を入れ替えれば、翻訳後の文型が変わってくる可能性もある（現在、広汎に使われている用例ベースの機械翻訳とはそうしたものである）。もちろん、使用する翻訳サービスが異なれば、翻訳の結果は異なる可能性が高い。

逆に、この状況をうまく利用すれば、学習者同士が「こちらの結果はこうなったが、こちらはどうか？」とお互いの出力を見比べ、そこから「どれが適切か」という議論を始めることもできるかもしれない。それであれば「丸写し」とは異なる協働作業である。

なお、学習者が目標言語のインプットを処理する際には、ワーキングメモリの限界から、意味的に不透明な機能語には注意が及びにくい可能性が指摘されている（VanPatten [1996] など）。日本語話者にとって、中国語の「了」はまさにこれである。

機械翻訳を使えば、特定の機能語の有無で結果がどのように変わるかを動的に実験することができ、通常なら意識しにくいこれらの語彙の機能に着目させることも可能である。なお、この課題を遂行する過程で、教科書の「了」の説明も参照させると、読んだだけではわからなかった説明の意味するところが腑に落ちるという効果も見られる。

あるいは、次のような課題も考えられる。「あなたは昨日何を食べたか？」と「あなたはいつ結婚したのか？」という二種類の質問を提示する。日本語の場合、この二つの文のテンス／アスペクトの形式は同じである（た）。この二つの文を英語に翻訳しても、テンス／アスペクトの形式は同じく過去形である。しかし、機械翻訳でこれを中国語に訳してみると、（ふつうは）異なるアスペクト形式で表現される（前者は助詞「了」を使い、後者は「是……的」の構文を使う）。どのような条件でこの二種類の形式が分岐するのかを、学習者自身が、母語をあれこれ操作することにより試行錯誤しつつ、帰納的に一般化することができるかもしれない。

一般に、「是……的」の構文のように日本語に対応する形式がない形は、文型自体が「意味的に不透明」であり、しかも、不連続要素をセットで用いることを強いられたため、「是」的「それぞれの機能分担も不透明である。このような文型は習得が難しい。」

このような構文に対し、機械翻訳を使いつつ、動的にあればこれと訳文を生成させ、それを比較し分析させるという活動は、メタ言語的な気づきを促進するのに効果があると思われる。前述の「了」を付け加えて全体の意味の変化を追わせるのは、活動の焦点となる言語形式をあらかじめ明示しその機能を探っていく課題であるのに対し、後者の課題は、ふだん意識しないちがいが言語によっては異なる形式で現れてくることを見つけさせ、具体的にはどこが異なるのか、異なっている部分は何れどのような機能をもっているかを考えさせる課題である。繰り返しすが、これは教科書に書かれた文法的な知識を復唱させる課題ではなく、自分で手を動かして見つけたものを、自分の頭で考え、自分のことばで表現させる課題である。<sup>4)</sup>

### (三) 音声に関する気づき

メタ言語的気づきを促進する課題は、文法構造に関する課題だけでなく、音声領域でも実現することができる。以下に一例を紹介する。

対面授業をおこなっている場合、教師は学習者の発音に対して即座にフィードバックができ、学習者がそのフィードバックの意味を理解できたか、そして、指摘された部分を修正できたかを、その場で把握できる。しかし、オンラインでは対面授業のときほど円滑にこれをおこなうことが

できない。ましてや、オンデマンド形式の場合はこのようなフィードバックは不可能に近い。学習者に自らの発音を録音させて提出させ、それに対してフィードバックをすることは考えられるが、リアルタイムではなくなるし、頻度も少なくなる。何よりも、学習者が理解できたか/修正できたかを確認するのに、新たな労力が発生する。このような不利な環境下で、ゼロから中国語の学習を始めた学習者の発音スキルを習熟度の高低だけで評価するのは、ほんとうに学習者のためになる評価であろうか(これまで教室授業でおこなわれてきた評価は、習熟度の高低だけで評価してきたわけだが)。

この問題に対する筆者の回答は、発音スキルだけを評価するのではなく、中国語の音声と既習得の言語(母語を含む)の音声を比べ、どこが異なるかを認知し、どの相違が重要で、その相違を実現すべく努力したことがどこまで成功しているかも併せて評価する方法へ路線転換せよというものである。

表1に示すのは、入門レベルの学習者の発音スキルの評価に使ったループリリックである。<sup>5)</sup>

従来の評価の観点を「旧」とし、ここで示した新しい評価の観点を「新」として、そのちがいを以下にまとめた。

「旧」意味の区別に重要な役割を果たしている $\alpha$ と $\beta$ という二種の音の区別を実現できているか。

表1 音声のスキルを評価するルーブリック

	目標以上を達成	目標を達成	目標まであと一步
(1) 日本語にな い音につい て	日本語にない音があることを理解し、全体を通じて、それを再現しようと意識して発音し、だいたい成功している。	日本語にない音があることを理解し、全体を通じて、それを再現しようと意識しているのがよくわかる。 【例】……(省略)……	日本語にない音を再現しようという意識が見られない箇所が、いくつかある。 【例】……(省略)……
(2) 紛らわしい 音について	紛らわしい音のペアがあることを理解し、全体を通じてそれを区別しようと意識して発音し、だいたい成功している。	紛らわしい音のペアがあることを理解し、全体を通じてそれを区別しようと意識しているのがよくわかる。 【例】……(省略)……	紛らわしい音のペアを区別しようという意識が見られない箇所がいくつかある。 【例】……(省略)……

(実際に使用したルーブリックは、【例】の部分で、「qu を ku のように発音している……」などのように詳細かつ具体的に記述しており、初学者が理解できるチェックリストになっている)

「新」意味の区別に重要な役割を果たしている $\alpha$ と $\beta$ という二種の音の区別が何によってできているかにどれだけ気づいているか、そして、その区別を実現しようとどれだけ注意をはらっているか、そして、それがどれだけ成功しているか。

要するに、ここで示した新しい評価の観点は、言語構造(音声)への気づき、その重要性の認識、そしてそれをどれだけ実現しようとしているかという態度を評価している。「成功しているか」はこれまでの観点(発音の習熟度)と大差ないが、そこに至るまでの「気づき」と「態度」を重視した評価法になっているところが新しい点である。

#### (四) なぜ「気づき」を重視するのか

なぜ「メタ言語的気づき」を重視するのかについて、もう一度整理しておきたい。

このような授業の組み立ては、伝統的な「構造に関する知識を積み上げる」授業のように見えるかもしれないがそうではない。確立した知識を一方的に伝授することも意図しないし、知識として蓄積させることも意図していない。自分が習得しようとしている言語について、学習者に自分で何かに気づかせること、それを学習のその時点での認識として、自分のことばで確認させることを意図している。その認識は日々変わりうるし、ときには「誤った仮説、

過剰な一般化をした仮説」であるかもしれない。しかし、それは問題ではない。ちょうど、「中間言語」がその時点その時点での学習者内部の言語システムの反映であると仮定できるように、日々変わりうるメタ言語的認識もその時点での学習者内部の言語システムの反映である。

ただ、このように言語の構造を分析することを好まない学習者もいる。そのように深く考えることもなく、感覚的・直観的に第二言語を習得していく学習者もいる。しかし、大学のような教育機関での第二言語の習得 (Instructed Second Language Acquisition と呼ばれる習得) は、使える時間もリソースも限定されている。知識として理解できることは知識として理解したほうが効率がよい。

さらに、このようなメタ言語的思考に習熟することは学習者の〈自立〉に役立つ。ある言語を習得しようとするとき、いつでもどこにでも、構造について説明をしてくれる教師がいるわけでもないし、そのような教材があるわけでもない。その場合、学習者は自力でそのような学習をする必要がある。逆にいえば、そのような学習ができる学習者は「自立できる」学習者であるといえる。

## 十二 過程を把握する

### (一) 学習の過程を把握できる評価方法

例えば、卒業論文を書いている学生が、毎週必ず指導教員のところへやってきて、「ここはどうすればよいか」「ここまで書いたが、これでよいか」と相談をしにくるとしよう。そうすると、教員はその論文の作成過程を逐一把握できる。このようにして作成された論文では、部分的な剽窃はともかく、丸ごと全部の盗作は成立しにくい。

このように学習成果物の作成過程を把握することは、その成果物の真正性を担保する重要な方策となる。現在、伝統的な筆記テストによる評価方法以外にも、パフォーマンス評価、ポートフォリオ評価などさまざまな評価方法が試みられているが、この中でも、ポートフォリオ評価は、教師が指定する条件に合致すると学習者が判断した成果物を蓄積した集合体 (ポートフォリオ) で評価するので、過程を把握するにはちょうどよい。

ただし、ポートフォリオに何を納めるかが問題で、よくおこなわれているように、日々の授業の小テスト、ちよつとした作文の課題の成果などを納めていくだけのポートフォリオの場合、それぞれが「誰がやっても同じ唯一の正解」を求める性質の課題であると、その個々の成果も、他

人の答案や機械翻訳の出力の丸写しだけで成立してしまう可能性がある。これでは解決にならない。やはり、個々の課題の個人化と個々の学習者に何かを考えさせ、気づかせ、創造させる課題であることが必要である。

なお、ポートフォリオ評価の目的はいろいろあるが、本来は、「学習の過程を見る」のほかにも「学習者自身による自己の学習への省察の促進」という目的もある。いうまでもなく、「丸写し」した成果を納めたポートフォリオでは、学習に対する省察を促進する効果は期待できない。

## (二) リスク回避のための評価方法として

期末に一回だけ大がかりな筆記テストを実施し、それだけで成績を評価する……という評価方法を採用する授業は、言語を学ぶ授業では（昨今では）珍しいと思われる。しかし、出席を数えたり、小テストを何回かおこなったりして、最後に期末テストを……という評価方法であれば、現在も広くおこなわれている。この「期末テスト」は「受講生を同一かつ唯一の時間帯に同一の場所に拘束して、教師が監視できる」ことを条件に成立してきた。

しかし、昨今では、対面授業に戻っても、この「受講生を同一かつ唯一の時間帯に同一の場所に拘束して、教師が監視できる」は実現が難しい。例えば、COVID-19による感染症のような伝染病が蔓延すると、学期の授業の最終週

に全員が揃うことが難しくなった。それではといふので、追試の日程を設定しても、その日までに症状が回復している保証もない。また、ある大学では、最終週が台風の襲来で休講になり、その補講日にまた台風が襲来して……という事態に陥ったこともあった。

このように考えてくると、受講生の拘束と監視を前提にした評価方法は、リスク管理の点から見ても危うい。ポートフォリオ評価は一例に過ぎないが、学習者の学習過程を把握し、たとえ、最終週の授業が成立しなくても妥当で公平な成績を算出できる評価方法の確立が、不測の事態に対応できる望ましい対策であるといえよう。

## 十三 おわりに

この文章では、監視をしなくても成立する中国語教育の授業を実現させるにどうしたらよいかを、具体例を挙げつつ考えた。このような授業の組み立てや評価の方法が、コミュニケーション能力の育成という外国語教育の目的により合致していると考えられるのであれば、コロナ禍が過去のものとなり、対面授業に何の支障がなくなっても、監視を前提とした外国語教育に戻る必要はないと考える。賛同者が増えることを期待したい。

- 〈1〉余談だが、「クラス全員で相談してもよい」という筆記テストを実施し、全員が正解だと思われる解答の合意に至るプロセスを見てみると、筆者は考えている。そのプロセスは、その授業で与えた教材やインスタレーションを、受講生がどう理解し、どのような中間言語を形成しつつあるかを物語る格好の資料となるはずである。このような協働作業を不正行為にしてしまうのはもったいない。
- 〈2〉筆者は、筆記テストの実施後、同じ受講生に「教科書・ノート参照可」で同じ問題を解かせる実験をしたことが何度かある。結果を見ると、成績下位の学習者は一度目と二度目では成績はあまり変わらない。理由はおそらく、どこを参照してよいかわからないからと、自分の産出した形式とモデルになる形式のどこが異なるかを分析できないからであろうと思われる。
- 〈3〉なお、この点を考えるにあたって、成瀬尚志編〔2016〕の内容を大いに参考にした。
- 〈4〉ただし、学習者の中には分析的に考えることが苦手である者もいる。ここで例示した二つの活動は、比較的複雑な、時間のかかる活動であるので、実際には、学習者の特性を考慮に入れつつ、もっと短時間に成果の得られる簡素な活動にプレイクダウンする必要があるかもしれない。
- 〈5〉この評価方法は山崎〔2021〕ですでに紹介している。

## 参考文献

- キャロライン・V・ギップス 2001 『新しい評価を求めて——テスト教育の終焉』 論創社 [原著] Gipps, C. V. 1994 *Beyond Testing: Towards a Theory of Educational Assessment*, The Falmer Press.
- 成瀬尚志編 2016 『学生を思考にいわなうレポート課題』 ひつじ書房
- 山崎直樹 2018 「権威」の要らない言語学習の可能性——ICTと学習者オートノミー』 『漢字文献情報処理研究』 第一八号、四一—七頁
- 山崎直樹 2021 「初級中国語のオンラインクラスにおける活動と評価」 『関西大学高等教育研究』 第一二号、一五七—一六二頁
- Garcia, I. and M. I. Pena 2011 “Machine Translation-assisted Language Learning: Writing for Beginners,” *Computer Assisted Language Learning*, 24(5), pp. 471–487.
- Lee, Y. J. 2021 “Still taboo? Using Machine Translation for Low-level EFL Writers,” *ELT Journal*, 75(4), pp. 432–441.
- VanPatten, B. 1996 *Input Processing and Grammar Instruction in Second Language Acquisition*. Ablex.