

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

L'ÉVALUATION DE LA PRODUCTION ORALE  
CHEZ LES ADULTES EN FRANCISATION : ANALYSE DES CRITÈRES DE CINQ  
ENSEIGNANTES

MÉMOIRE  
PRÉSENTÉ  
COMME EXIGENCE PARTIELLE  
DE LA MAÎTRISE EN LINGUISTIQUE

PAR  
SARA TROTTIER

NOVEMBRE 2007

UNIVERSITÉ DU QUÉBEC À MONTRÉAL  
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

## REMERCIEMENTS

Mes remerciements vont d'abord à Martine Peters, ma directrice, qui, par sa rigueur intellectuelle, sa précision, sa clarté et son immense dynamisme, m'a guidée et motivée dans toutes les étapes de la rédaction de ce mémoire.

Ensuite, mes remerciements s'adressent à Michel Désy, pour ses suggestions, ses conseils, son support et sa compréhension.

Finalement, mes remerciements vont aux étudiants et aux enseignants qui ont généreusement contribué à mon projet : de près, en s'impliquant personnellement dans ma recherche, et de loin, en m'inspirant le sujet. Ce mémoire leur est destiné.

## TABLE DES MATIÈRES

LISTE DES FIGURES.....	viii
LISTE DES TABLEAUX .....	ix
LISTE DES ABRÉVIATIONS .....	xi
RÉSUMÉ .....	xii
INTRODUCTION.....	1
CHAPITRE I	
PROBLÉMATIQUE .....	4
1.1 LA FRANCISATION DES ADULTES AU QUÉBEC .....	4
1.2 L'ÉVALUATION DE L'ORAL .....	5
1.3 LA FLUCTUATION DANS L'ÉVALUATION .....	6
CHAPITRE II	
CADRE THÉORIQUE .....	8
2.1 LA NOTION DE COMPÉTENCE .....	8
2.1.1 VISION CONSTRUCTIVISTE DE LA NOTION DE COMPÉTENCE.....	9
2.2 LA COMPÉTENCE COMMUNICATIVE.....	11
2.2.1 GENÈSE DU CONCEPT DE COMPÉTENCE COMMUNICATIVE.....	12
2.2.2 LE MODÈLE DE CANALE ET SWAIN (1980).....	15
2.2.3 LE MODÈLE DE LYLE BACHMAN (1990).....	17
2.2.3 LE MODÈLE DE GLENN FULCHER (2003).....	21
2.3 LES QUATRE HABILITÉS.....	27
2.4 L'ÉVALUATION .....	28
2.4.1 LA VALIDITÉ .....	29

2.4.2 LA FIDÉLITÉ .....	30
2.4.3 L'APPROCHE PSYCHOMÉTRIQUE ET L'APPROCHE ÉDUMÉTRIQUE .....	32
<b>2.5 L'ÉVALUATION DE LA PRODUCTION ORALE .....</b>	<b>33</b>
2.5.1 LA VALIDITÉ DE L'ÉVALUATION DE LA PRODUCTION ORALE .....	33
2.5.2 LA FIDÉLITÉ DANS L'ÉVALUATION DE LA PRODUCTION ORALE .....	37
2.5.3 LES GRILLES D'ÉVALUATION .....	38
<b>2.6 L'ÉVALUATION DE LA PRODUCTION ORALE À LA CSDM .....</b>	<b>43</b>
<b>2.7 SYNTHÈSE DU CHAPITRE .....</b>	<b>45</b>
<b>CHAPITRE III</b>	
<b>MÉTHODOLOGIE .....</b>	<b>47</b>
<b>3.1 RÉSUMÉ DE L'EXPÉRIMENTATION .....</b>	<b>47</b>
<b>3.2 PRÉSENTATION DES PARTICIPANTES .....</b>	<b>48</b>
3.2.1 JULIE .....	48
3.2.2 ÉLAINE .....	49
3.2.3 PHUONG .....	49
3.2.4 CLÉMENCE .....	49
3.2.5 MARIE-ÈVE .....	50
<b>3.3 INSTRUMENTS UTILISÉS .....</b>	<b>51</b>
3.3.1 QUESTIONNAIRE SOCIODÉMOGRAPHIQUE .....	51
3.3.2 VIDÉOS .....	51
3.3.3 FEUILLES DE NOTES .....	54
3.3.4 CONSIGNES D'EXAMEN .....	54
3.3.5 GRILLE ET PONDÉRATION .....	55
3.3.6 QUESTIONNAIRE D'ENTREVUE SEMI DIRIGÉE .....	55
<b>3.4 DÉROULEMENT .....</b>	<b>57</b>

<b>3.5 TRAITEMENT DES DONNÉES.....</b>	<b>58</b>
<b>CHAPITRE IV</b>	
<b>PRÉSENTATION DES RÉSULTATS.....</b>	<b>61</b>
<b>4.1 RÉSULTATS DE LA PREMIÈRE PARTIE DE L'ENTREVUE.....</b>	<b>61</b>
4.1.1 ÉVALUATION DE L'ÉTUDIANT 1 .....	63
4.1.1.1 <i>Comment Julie a évalué l'étudiant 1</i> .....	63
4.1.1.2 <i>Comment Éline a évalué l'étudiant 1</i> .....	64
4.1.1.2 <i>Comment Phuong a évalué l'étudiant 1</i> .....	65
4.1.1.4 <i>Comment Clémence a évalué l'étudiant 1</i> .....	67
4.1.1.5 <i>Comment Marie-Ève a évalué l'étudiant 1</i> .....	68
4.1.1.6 <i>Bilan des commentaires sur l'étudiant 1</i> .....	69
4.1.2 ÉVALUATION DE L'ÉTUDIANTE 2 .....	70
4.1.2.1 <i>Comment Julie a évalué l'étudiante 2</i> .....	70
4.1.2.2 <i>Comment Éline a évalué l'étudiante 2</i> .....	72
4.1.2.3 <i>Comment Phuong a évalué l'étudiante 2</i> .....	73
4.1.2.4 <i>Comment Clémence a évalué l'étudiante 2</i> .....	74
4.1.2.5 <i>Comment Marie-Ève a évalué l'étudiante 2</i> .....	75
4.1.2.6 <i>Bilan de l'évaluation de l'étudiante 2</i> .....	76
4.1.3 ÉVALUATION DE L'ÉTUDIANT 3 .....	77
4.1.3.1 <i>Comment Julie a évalué l'étudiant 3</i> .....	77
4.1.3.2 <i>Comment Éline a évalué l'étudiant 3</i> .....	78
4.1.3.3 <i>Comment Phuong a évalué l'étudiant 3</i> .....	80
4.1.3.4 <i>Comment Clémence a évalué l'étudiant 3</i> .....	81
4.1.3.5 <i>Comment Marie-Ève a évalué l'étudiant 3</i> .....	82
4.1.3.6 <i>Bilan de l'évaluation de l'étudiant 3</i> .....	83
4.1.4 ÉVALUATION DE L'ÉTUDIANTE 4.....	83

4.1.4.1	<i>Comment Julie a évalué l'étudiante 4</i> .....	83
4.1.4.2	<i>Comment Élane a évalué l'étudiante 4</i> .....	85
4.1.4.3	<i>Comment Phuong évalue l'étudiante 4</i> .....	86
4.1.4.4	<i>Comment Clémence a évalué l'étudiante 4</i> .....	88
4.1.4.5	<i>Comment Marie-Ève a évalué l'étudiante 4</i> .....	89
4.1.4.6	<i>Bilan de l'évaluation de l'étudiante 4</i> .....	90
<b>4.2</b>	<b>PRÉSENTATION DES RÉSULTATS DE LA DEUXIÈME PARTIE DES ENTREVUES</b> .....	<b>90</b>
4.2.1	CRITÈRES DE JULIE.....	92
4.2.2	CRITÈRES D'ÉLAINE.....	94
4.2.3	CRITÈRES DE PHUONG.....	95
4.2.4	CRITÈRES DE CLÉMENCE.....	96
4.2.5	CRITÈRES DE MARIE-ÈVE.....	97
<b>CHAPITRE V</b>		
	<b>INTERPRÉTATION DES RÉSULTATS</b> .....	<b>100</b>
<b>5.1</b>	<b>COMPOSANTE LANGAGIÈRE</b> .....	<b>101</b>
5.1.1	LA PHONOLOGIE.....	101
5.1.1.1	<i>La prononciation des phonèmes</i> .....	102
5.1.1.2	<i>Les liaisons</i> .....	104
5.1.1.3	<i>L'intensité vocale</i> .....	104
5.1.1.4	<i>Synthèse de la phonologie</i> .....	105
5.1.2	LA PRÉCISION.....	106
5.2.2.1	<i>Le nombre et la gravité des erreurs</i> .....	107
5.2.2.2	<i>La conjugaison</i> .....	109
5.2.2.3	<i>La syntaxe</i> .....	110
5.2.2.4	<i>Synthèse de la sous-composante de la précision</i> .....	111

5.1.3 L' AISANCE .....	112
5.1.3.1 <i>Le rythme et la fluidité</i> .....	112
5.1.3.2 <i>La quantité d'information</i> .....	113
5.1.3.3 <i>Synthèse de la sous-composante de l'aisance</i> .....	115
<b>5.2 CAPACITÉ STRATÉGIQUE .....</b>	<b>116</b>
5.2.1 LES STRATÉGIES DE RÉUSSITE.....	116
5.2.2 STRATÉGIE DE PLANIFICATION DU DISCOURS .....	117
5.2.3 SYNTHÈSE DE LA COMPOSANTE STRATÉGIQUE .....	119
<b>5.3 SAVOIR TEXTUEL .....</b>	<b>119</b>
<b>5.4 COMPÉTENCE INTERACTIVE .....</b>	<b>121</b>
<b>5.5 CRITÈRES NON LINGUISTIQUES.....</b>	<b>123</b>
<b>5.6 SYNTHÈSE DES COMMENTAIRES DES ENSEIGNANTES.....</b>	<b>124</b>
<b>5.7 RÉUTILISATION DES ACQUIS PROPRES AU NIVEAU.....</b>	<b>127</b>
<b>5.8 LES OBSERVATIONS PENDANT LA SESSION.....</b>	<b>129</b>
<b>5.9 SYNTHÈSE DU CHAPITRE .....</b>	<b>130</b>
<b>CONCLUSION.....</b>	<b>132</b>
<b>APPENDICE A .....</b>	<b>136</b>
<b>APPENDICE B.....</b>	<b>137</b>
<b>APPENDICE C .....</b>	<b>141</b>
<b>APPENDICE D .....</b>	<b>142</b>
<b>APPENDICE E.....</b>	<b>143</b>
<b>APPENDICE F.....</b>	<b>144</b>
<b>APPENDICE G .....</b>	<b>145</b>
<b>RÉFÉRENCES.....</b>	<b>146</b>

## LISTE DES FIGURES

<b>Figure 2.1</b> Les facteurs de la communication, d'après Jakobson.....	13
<b>Figure 2.2</b> Modèle de la compétence langagière de Bachman (1990) .....	19
<b>Figure 2.3</b> Interaction entre composantes (Bachman et Palmer, 1996) .....	20

## LISTE DES TABLEAUX

Tableaux	Page
2.1	Modèle de la compétence communicative de Canale et Swain (1980) .....17
2.2	Modèle de la compétence communicative de Fulcher (2003) .....26
3.1	Renseignements démographiques sur les participantes .....50
3.2	Classification des codes utilisés .....59
4.1	Notes accordées aux quatre étudiants par les cinq évaluatrices.....62
4.2	Commentaires de Julie au sujet de l'étudiant 1.....63
4.3	Commentaires d'Élaine au sujet de l'étudiant 1 .....64
4.4	Commentaires de Phuong au sujet de l'étudiant 1 .....65
4.5	Commentaires de Clémence au sujet de l'étudiant 1 .....67
4.6	Commentaires de Marie-Ève au sujet de l'étudiant 1 .....68
4.7	Commentaires de Julie au sujet de l'étudiante 2 .....70
4.8	Commentaires d'Élaine au sujet de l'étudiante 2 .....72
4.9	Commentaires de Phuong au sujet de l'étudiante 2 .....73
4.10	Commentaires de Clémence au sujet de l'étudiante 2 .....75
4.11	Commentaires de Marie-Ève au sujet de l'étudiante 2 .....75
4.12	Commentaires de Julie au sujet de l'étudiant 3 .....78
4.13	Commentaires d'Élaine au sujet de l'étudiant 3 .....78
4.14	Commentaires de Phuong au sujet de l'étudiant 3 .....80
4.15	Commentaires de Clémence au sujet de l'étudiant 3 .....81

## **LISTE DES ABRÉVIATIONS**

CSDM : Commission scolaire de Montréal

DGFA : Direction générale de la formation des adultes

MELS : Ministère de l'Éducation, du Loisir et du Sport

MICC : Ministère de l'Immigration et des Communautés culturelles

## RÉSUMÉ

Ce mémoire est dédié à la recherche sur l'évaluation de la production orale en francisation. Il tente de définir le concept de compétence communicative et de faire le point sur l'évaluation de cette compétence. Plus particulièrement, ce mémoire s'intéresse à la façon dont des enseignants, qui travaillent avec une clientèle allophone adulte dans une commission scolaire de Montréal, s'y prennent pour évaluer la compétence communicative à l'oral des étudiants.

Dans la pratique, il existe une grande disparité dans l'évaluation de la production orale parce que cette activité revêt une large part de subjectivité. La littérature révèle que des facteurs personnels (entre autres) influencent grandement le jugement des évaluateurs. Ce problème d'inconstance entre les évaluateurs a été le point de départ de notre recherche.

Le but de la recherche était double. D'abord, nous voulions découvrir quels aspects de la compétence communicative retiennent l'attention des enseignants quand ils évaluent des étudiants. Ensuite, nous souhaitions connaître les critères d'évaluation personnels des enseignants et ce, pour mieux comprendre d'où provient le manque de constance entre les évaluateurs et éventuellement, pour proposer des pistes de solutions à ce problème.

Pour atteindre ce double objectif, nous avons conçu une expérimentation qui a consisté à faire évaluer par cinq enseignantes oeuvrant dans une commission scolaire de Montréal la performance de quatre étudiants de niveau avancé, lors de l'examen oral sommatif. Les cinq enseignantes ont pris des notes au fil du visionnage. Nous avons mené subséquemment une entrevue semi dirigée qui visait à découvrir leurs critères d'évaluation personnels, pour les analyser et les comparer.

De prime abord, nous avons constaté que la note accordée à chaque étudiant variait beaucoup d'une enseignante à l'autre, allant de très bonne à très mauvaise (voire échec). L'analyse qualitative des commentaires des enseignantes nous a permis de découvrir qu'elles n'ont pas toutes les mêmes attentes face aux étudiants, qu'elles n'ont pas la même compréhension des composantes de la compétence communicative et qu'elles prennent beaucoup de liberté vis-à-vis des recommandations ministérielles (allant jusqu'à délaissier la grille d'évaluation du ministère), jugeant souvent les critères d'évaluation prescrits par la grille si imprécis qu'elle devient inutilisable. Le manque d'instrument de mesure précis et l'absence de consensus entre les participantes au sujet de la définition des composantes de la compétence communicative semblent être la cause de l'inconstance entre les évaluateurs.

De plus, nous avons constaté qu'outre les composantes langagière (qui englobe la phonologie, la précision linguistique et l'aisance) et textuelle (la cohésion et la cohérence), qui sont généralement invoquées par les enseignantes pour justifier leurs notes, des composantes telles que l'utilisation de stratégies de communication non linguistiques, la compétence interactive et certaines caractéristiques personnelles des étudiants peuvent avoir, selon les enseignantes, une influence majeure sur le résultat de ces derniers.

Mots clés : évaluation, compétence communicative, francisation, production orale.

## INTRODUCTION

Le nombre de personnes qui arrivent annuellement au Québec ne parlant pas français est important et tout porte à croire qu'il le restera pendant les années qui viennent. Entre 2000 et 2004, 191 477 immigrants ont été admis dans la province. Parmi ceux-ci, 96 079 ne parlaient pas français au moment de leur arrivée (ministère de l'Immigration et de Communautés culturelles (MICC), 2005a). En outre, *Le Plan d'immigration du Québec pour l'année 2007* (MICC, 2006) comptait «Assurer une progression des volumes d'immigration pour que ceux-ci atteignent 48 000 admissions en 2007» (p. 2) Les planificateurs du MICC s'attendaient à ce que parmi ces nouveaux arrivants, 43% ne connaîtraient pas le français. Or, la connaissance du français est un facteur d'intégration sociale important, notamment pour accéder au marché du travail. Selon le recensement de 2001 de l'Institut de la statistique du Québec, à Montréal, 310 345 personnes utilisaient seulement le français au travail, ce qui représente 31,1% des travailleurs. En outre, 492 635 utilisaient le français et l'anglais, soit 50,2 % des travailleurs. Donc, pour occuper un emploi à Montréal, dans plus de 80% des cas, il faut parler français.

Facteur d'intégration au travail, le français est aussi un facteur d'intégration culturelle :

Le français représente non seulement un outil de communication essentiel, mais aussi un symbole commun d'appartenance à la société québécoise et de dialogue interculturel. (MICC, 2006, p.12)

Dans le cadre de la publication de son Plan d'action 2004-2007 intitulé : *Des valeurs partagées, des intérêts communs*, la ministre de l'Immigration et des Communautés culturelles de l'époque, madame Michelle Courchesne, révélait dans un communiqué publié

le 20 mai 2003 sur le site du MICC: «La connaissance du français est non seulement une condition d'insertion en emploi dans la langue de la majorité, elle est aussi un facteur important d'enracinement au Québec.» (MICC, 2003, p.1).

Or, qu'est-ce que la ministre Courchesne entend par «connaissance du français»? À partir de quand une personne maîtrise-t-elle suffisamment la langue pour fonctionner adéquatement dans la société ?

Dès leur entrée au Québec, les nouveaux arrivants sont dirigés vers des cours de français, destinés à favoriser leurs connaissances linguistiques et faciliter ainsi leur intégration<sup>1</sup>. Au moment de l'inscription, les futurs étudiants passent un test de classement qui vise à les diriger au bon niveau. Pendant leur scolarité, ils sont évalués de façon informelle par leur enseignant. À la fin de chaque session, ils doivent se plier à un examen sommatif : au terme du cours, les formateurs s'attendent à ce qu'ils puissent fonctionner dans la langue commune. Lorsqu'ils veulent intégrer un ordre professionnel, les immigrants doivent à nouveau se plier à un examen de français, crucial pour leur avenir, qui leur permettra ou non de pratiquer leur profession (Office québécois de la langue française, 2006). S'ils désirent poursuivre des études dans un établissement scolaire francophone, ils doivent souvent réussir un test d'admission en français.

Ces quelques exemples d'étapes qui jalonnent le parcours des immigrants visaient à mettre en lumière l'omniprésence de l'évaluation linguistique dans leur processus d'intégration. Les occasions où leur niveau de maîtrise du français est évalué de façon formelle sont multiples, que ce soit à l'oral ou à l'écrit. Ces évaluations peuvent avoir un impact important dans leur vie.

Le problème auquel nous nous intéresserons dans ce mémoire est celui de l'évaluation de la connaissance du français, plus spécifiquement de la capacité à le parler. Nous voulons

---

<sup>1</sup> Malgré l'offre de cours de français et les différents moyens utilisés pour franciser les allophones, plusieurs ne maîtrisent pas le français après plusieurs années passées au Québec. À Montréal et à Laval, selon le recensement de 2001 mené par l'Institut de la statistique du Québec, 266 315 personnes (allophones et anglo-québécois confondus) ne parlaient pas français.

savoir comment les évaluateurs, et particulièrement les enseignants, s’y prennent pour évaluer la connaissance du français oral de leurs étudiants. Cela constitue la base de la problématique que nous présenterons dans le chapitre suivant. Elle mettra en lumière les difficultés reliées à l’évaluation de la production orale des étudiants de français, langue seconde. Une fois notre problématique clarifiée, nous présenterons notre cadre théorique dans lequel nous définirons les concepts de compétence, de compétence communicative, et plus spécifiquement à l’oral. Nous définirons ensuite l’évaluation, puis, l’évaluation de la compétence communicative à l’oral. Nous parlerons des divers tests et échelles conçus à cette fin. Nous terminerons notre cadre théorique en parlant de la façon dont se fait l’évaluation dans un établissement de la Commission scolaire de Montréal et en posant nos questions de recherches. Viendra ensuite le chapitre consacré à la méthodologie envisagée pour mener notre expérimentation, expérimentation qui vise à découvrir les critères sur lesquels les enseignants s’appuient pour évaluer leurs étudiants. Dans ce chapitre, nous parlerons de nos instruments de collecte de données et de la façon dont nous avons traité les données. Dans le chapitre suivant, nous présenterons les données. Ensuite, dans l’avant-dernier chapitre, nous les analyserons. Finalement, dans notre conclusion, nous parlerons des retombées et des limites de notre recherche.

## CHAPITRE I

### PROBLÉMATIQUE

Le but de ce chapitre est de montrer que l'évaluation de la capacité à parler le français est au cœur du processus d'intégration des allophones. Pour illustrer cette réalité, nous parlerons d'abord du nombre important de personnes qui vivent au Québec sans savoir parler français. Nous verrons que leur méconnaissance de la langue peut nuire à leur intégration. Ensuite, nous montrerons comment le gouvernement québécois tente d'aider les gens à apprendre la langue officielle. Nous verrons finalement que l'évaluation jalonne le parcours des allophones et des anglophones québécois et que la fluctuation dans l'évaluation est un problème important.

#### **1.1 La francisation des adultes au Québec**

Au Québec, la francisation des allophones semble être une préoccupation majeure, puisque leur intégration en dépend (MICC, 2003). Dès leur arrivée, les immigrants sont donc dirigés vers des cours de français. Deux ministères se partagent la responsabilité d'offrir des cours de français, langue seconde, aux personnes désireuses de l'apprendre. Ce sont les ministères de l'Immigration et des Communautés culturelles (MICC) et de l'Éducation, du Loisir et du Sport (MELS). Dans les deux cas, les cours doivent répondre le mieux possible aux besoins linguistiques courants des immigrants. Au MELS, le ministère qui nous intéresse plus spécifiquement, les objectifs pédagogiques du programme de français, langue seconde de 1994, élaborés par la Direction générale de la Formation des Adultes (DGFA) sont ancrés dans la réalité à laquelle les nouveaux citoyens sont confrontés (DGFA, 1994). L'interaction

orale y occupe une place centrale. L'objectif global du programme est : «D'amener l'élève à communiquer dans les situations où l'usage de la langue française lui est nécessaire et utile» (DGFA, 1994, p.13). Cette volonté d'adéquation du programme de francisation aux besoins des immigrants, tout en réservant une place importante à l'écrit (DGFA, 1994), s'oriente surtout vers l'acquisition de la langue orale, perçue comme essentielle à l'intégration. Les enseignants et enseignantes chargés de la francisation des immigrants sont donc appelés à consacrer la partie la plus importante de leurs cours à l'oral.

## **1.2 L'évaluation de l'oral**

Au terme du cours, il revient à l'enseignante et à l'enseignant d'évaluer le degré de maîtrise du français de leurs étudiants. Selon le dernier programme de la DGFA (1994), l'évaluation de la production orale doit ressembler le plus possible à une interaction orale réaliste, qui pourrait se produire de façon naturelle dans la vie de tous les jours, dans un domaine de la vie courante. À la fin de chaque niveau, c'est un jeu de rôles qui sert à juger de la capacité des étudiants à se servir de la langue en contexte de communication. Il s'agit d'une évaluation sommative, c'est-à-dire d'une évaluation qui sert à déterminer si l'étudiant est apte ou non à passer au cours suivant. Les étudiants sont évalués à l'oral individuellement ou en équipe de deux, selon la préférence de l'enseignant, grâce à un jeu de rôle relié aux objectifs de communication prévus par le programme. L'enseignant participe au jeu de rôle s'il a choisi d'évaluer les étudiants individuellement ou, s'il veut évaluer deux étudiants à la fois, il observe la paire d'étudiants jouer des rôles variés, le tout pendant environ cinq minutes. (DGFA, 1994). Et comment l'enseignant note-t-il la performance des étudiants? Au terme du jeu de rôle, il consulte la grille d'évaluation du MELS (voir appendice A), sur laquelle six niveaux «d'habileté de l'expression» (DGFA 1994) sont décrits, ou bien il utilise une grille que lui ou un collègue a conçue, et il attribue le niveau qui convient, selon lui, à chacun des étudiants. L'enseignant, comme évaluateur, a la tâche d'associer le plus justement possible la performance des étudiants pendant le jeu de rôle au descripteur de la grille qui décrit le mieux leur performance. C'est ce qui déterminera leur note finale à l'oral.

### 1.3 La fluctuation dans l'évaluation

Or, même si elle semble simple, cette façon d'évaluer la production orale est plus compliquée qu'elle ne le paraît. Le caractère souvent très subjectif des descripteurs de la grille peut donner lieu à une grande variété d'interprétations (Bachman, 1990; Kormos et Dénes, 2004; Oller, 1983; Orr, 2002). Les recherches qui démontrent la difficulté d'arriver à un consensus entre les évaluateurs abondent (Brown, 1995 ; Mullen, 1978 ; Wigglesworth, 1993 ; Bachman, Lynch et Mason, 1995 ; Chalhoub-Deville, 1995 ; Lumley et McNamara, 1999). Les étudiants ont l'impression que réussir un examen oral est une affaire de chance : tout dépend de l'évaluateur sur lequel ils sont tombés (Wigglesworth, 1993). De plus, à la Commission scolaire de Montréal (CSDM) (comme dans les autres commissions scolaires, d'ailleurs), comme les enseignants ne sont pas obligés d'utiliser la grille du MELS, les critères d'évaluation sur lesquels repose leur jugement ne sont pas forcément les mêmes d'une classe à l'autre<sup>2</sup>.

Les difficultés liées à l'évaluation de l'oral ont été soulevées précisément dans un centre d'éducation des adultes de la CSDM. Une des enseignantes du centre a étudié la question de l'évaluation des compétences à l'oral pendant la période allant de mars à juin 2006. Elle a consulté ses collègues afin de connaître leur opinion sur le sujet. Il s'est dégagé des consultations, entre autres, que certains enseignants souhaitaient étudier la possibilité d'utiliser une autre grille que celle du MELS, une grille plus précise. En outre, plusieurs enseignants ont exprimé leur désir de recevoir une formation en docimologie<sup>3</sup>. Il est clair que la question de l'évaluation de l'oral préoccupait les enseignants de ce centre. Nous pensons que cette préoccupation existe dans d'autres centres.

---

<sup>2</sup> Nous avons recensé dans un des centres d'éducation des adultes de la Commission scolaire de Montréal, sept grilles d'évaluation de l'oral qui variaient par les critères retenus, par la pondération de chaque critère et aussi par le type d'évaluation (exposé, jeu de rôle, saynète).

<sup>3</sup> Comme nous avons enseigné dans ce centre, nous avons pu participer aux discussions lors de réunions pédagogiques. Les commentaires que nous rapportons viennent de la synthèse des comptes rendus du 22 mars, remis à tous les enseignants, dont nous avons gardé une copie.

Le problème de la variation entre les évaluateurs est soulevé de façon très éloquente par Orr (2002). Dans une recherche portant sur la constance inter évaluateurs en anglais, langue seconde, ce dernier a comparé les commentaires de 20 évaluateurs expérimentés et formés selon des standards précis. Ces commentaires ont été recueillis suite à l'évaluation des mêmes quatre candidats. Le niveau attribué au même étudiant différait d'un extrême à l'autre de l'échelle et les commentaires des évaluateurs étaient très souvent contradictoires. L'auteur explique la différence inter évaluateurs par trois biais principaux : 1) les évaluateurs peuvent avoir une perception différente du même aspect ; 2) les évaluateurs ne portent pas attention aux mêmes aspects et 3) les évaluateurs ne tiennent pas toujours compte de la grille pour évaluer. En conclusion, Orr (2002) relevait l'utilité que peuvent avoir les commentaires des enseignants pour comprendre la façon dont ils évaluent et éventuellement, pour construire des grilles mieux adaptées à leurs besoins.

La grande fluctuation dans le jugement des évaluateurs nous amène à nous questionner sur les critères qui servent de fondement au jugement des enseignants lors de l'évaluation de la production orale de leurs étudiants. L'objectif de notre recherche est de connaître les raisons qui président aux décisions des évaluateurs. Pour y parvenir, nous devons d'abord clarifier les concepts de base qui sont en jeu lors de l'évaluation. C'est ce que nous ferons dans le chapitre suivant, le cadre théorique.

## CHAPITRE II

### CADRE THÉORIQUE

Dans le chapitre suivant, nous définirons les concepts-clés de notre recherche. Nous commencerons par définir la notion de compétence, puis, celle de compétence communicative. Nous présenterons alors trois schémas, celui de Canale et Swain (1980), celui de Lyle Bachman (1990) et celui de Glenn Fulcher (2003) qui modélisent la notion de compétence communicative. Celui de Glenn Fulcher s'intéresse particulièrement à la production orale. Nous parlerons ensuite du découpage de la compétence communicative en quatre habiletés.

Une fois le concept de compétence communicative défini, nous parlerons de l'évaluation en général et ensuite, de l'évaluation de la production orale. Dans cette section, nous présenterons diverses modalités d'évaluation de l'oral en langue seconde et nous parlerons de diverses échelles. Nous continuerons en présentant les modalités d'évaluation de la production orale en cours à la Commission scolaire de Montréal et nous terminerons ce chapitre en posant nos questions de recherche.

#### **2.1 La notion de compétence**

La popularité du terme «compétence» s'accroît depuis une décennie en sciences de l'éducation et ce, dans plusieurs pays de l'ouest de l'Europe, aux États-Unis, au Canada et en Australie (Romainville, 1996; Stroobants, 2000). Pourtant, cette notion est associée à un flou terminologique, voire à un «labyrinthe tautologique de définitions» (Romainville, 1996, p.133). Stroobants (2000, p.61) argue ironiquement que «son mode d'existence principal est

d'entretenir une intense activité de redéfinition». C'est une notion qui n'est «jamais clairement définie» (Castelloti, 2002, p.9).

Jonnaert (2002) soutient qu'en éducation, il y a deux visions de la notion de compétence : une première, dérivée du béhaviorisme et une autre, issue du constructivisme. La vision béhavioriste met l'accent sur l'acquisition de comportements perçus comme nécessaires. La vision constructiviste prône quant à elle le développement de la capacité de l'élève à s'adapter aux exigences particulières d'une situation. Dans les paragraphes suivants, nous définirons sommairement la notion de compétence issue de la deuxième vision.

### **2.1.1 Vision constructiviste de la notion de compétence**

Actuellement, les chercheurs en sciences de l'éducation (Perrenoud, 1997; Jonnaert, 2002, 2003; Rey et al, 2003) tiennent à se distancier de la vision strictement béhavioriste de la compétence. Ils reprochent aux programmes scolaires d'avoir fractionné les contenus enseignés en une multitude d'objectifs. Ces chercheurs, qui adhèrent à une vision renouvelée de la notion de compétence, insistent sur l'importance de créer en classe des situations qui permettront aux élèves de mobiliser une multitude d'habiletés, plutôt que d'essayer d'atteindre des objectifs imposés de l'extérieur, comme le voulait l'approche béhavioriste de la notion de compétence (Biemans et al, 2004; Jonnaert, 2002; Perrnoud, 1997; Rey et al, 2003; Allal, 1999).

La vision de la compétence que prônent Perrenoud (1997), Jonnaert, (2002, 2003). Rey et al, (2003) et Bissonnette et Richard (2001) et qui est mise de l'avant dans le programme actuel du ministère de l'Éducation, du Loisir et du Sport du Québec (2006) est issue du constructivisme. Elle se base sur les principes d'assimilation et d'accommodation, qui veulent que l'individu fabrique lui-même son savoir à partir d'une confrontation avec la réalité (Bronckart, 1977). La compétence s'acquiert donc grâce à la réflexion et à l'action, dans une situation donnée. Les définitions qui suivent s'inscrivent dans la lignée constructiviste de la notion de compétence.

Le ministère de l'Éducation, du Loisir et du Sport (MELS) du Québec (2006) définit la compétence comme «un savoir-agir complexe fondé sur la mobilisation et l'utilisation efficaces d'un ensemble de ressources» (MELS, 2006, p.13). Perrenoud, (1997, p.7) la définit comme une «capacité d'agir efficacement dans un type défini de situations, capacité qui s'appuie sur des connaissances, mais ne s'y réduit pas.». Pour Allal (2000, p.81), la compétence est «un réseau intégré et fonctionnel constitué de composantes cognitives, affectives, sociales, sensorimotrices, susceptibles d'être mobilisé en actions finalisées face à une famille de situations.» Bissonnette et Richard (2001, p. 74) en donnent une définition qui intègre la plupart des concepts utilisés chez les auteurs consultés:

Une compétence [...] constitue un savoir-agir résultant d'une compréhension adéquate des savoirs, savoir-faire et savoir-être intégrés et accessibles en mémoire, «mobilisables» de façon efficiente parce qu'ils ont été utilisés régulièrement et avec succès dans une grande variété de contextes et de disciplines et ce, autant à l'école que dans la vie quotidienne. La compétence ne représente pas un état, mais un processus, qui est d'ailleurs toujours perfectible parce qu'il résulte d'une démarche en construction.

La liste de définitions de la notion de compétence est longue, mais des constantes demeurent : 1) La compétence est reliée à l'action, au fait d'agir; elle se déploie dans l'acte (Perrenoud, 1997; Jonnaert, 2002; Rey, 2003). 2) Le déploiement d'une compétence est relié à une situation, qu'on appelle parfois un «problème» (Perrenoud, 1997; Jonnaert, 2002). 3) Une même compétence peut servir dans plusieurs situations (ou problèmes) différents (Perrenoud, 1997; Jonnaert, 2002; Bissonnette et Richard, 2001). 4) La compétence exige la mobilisation de nombreuses ressources de différents ordres (Perrenoud, 1997; Jonnaert, 2002; Bissonnette et Richard, 2001). 5) L'idée d'un réseau impliquant différents types de savoir est au cœur de la notion de compétence (Perrenoud, 1997; Jonnaert, 2002; Castelloti, 2002; Allal, 2000). 6) La compétence va de pair avec l'efficacité dans l'exécution de la tâche (Perrenoud, 1997; Jonnaert, 2002; Castelloti, 2002; Allal, 2000). 7) La compétence est le fruit d'un processus d'apprentissage, elle se développe avec l'expérience (Bissonnette et Richard, 2001; Dolz, 2000). 8) Il existe une instance capable de reconnaître cette compétence (Castelloti, 2002).

Définie ainsi, la notion de compétence s'insère dans une vision constructiviste de la pédagogie, puisque l'apprenant doit jouer un rôle actif dans son apprentissage. Il faut qu'il confronte ses connaissances à la situation pour arriver à régler un problème et s'adapter à toute nouvelle situation.

## **2.2 La compétence communicative**

Nous venons de voir qu'une compétence est un savoir-faire qui se déploie en contexte et qui exige la mobilisation de plusieurs ressources, de différents ordres (Perrenoud, 1997; Jonnaert, 2002 ; Castelloti, 2002 ; Allal, 2000). Rien n'est plus vrai en ce qui concerne la compétence communicative. S'il est un «savoir-agir complexe fondé sur la mobilisation et l'utilisation efficaces d'un ensemble de ressources» (MELS, 2006, p.13), «un réseau intégré et fonctionnel constitué de composantes cognitives, affectives, sociales, sensorimotrices, susceptibles d'être mobilisé en actions finalisées face à une famille de situations» (Allal (2000, p.81), c'est bien celui mis en branle lorsqu'une personne communique.

La notion de compétence trouve donc un champ d'application particulièrement fertile et pertinent en sciences du langage. Nous devons son introduction dans ce domaine au célèbre linguiste Noam Chomsky (Alptekin, 2002). En 1957, dans *Syntactic Structure*, Chomsky a jeté les bases de sa révolutionnaire théorie de la grammaire générative et transformationnelle (Bronckart, 1977). Avec Chomsky naissent les concepts de compétence et de performance, appliqués en linguistique, mais aussi en sciences de l'éducation (Alptekin, 2002 ; Leung, 2005 ; Canale et Swain, 1980).

Dans les sections qui suivent, nous retracerons d'abord la genèse du concept de compétence en linguistique. Nous verrons ensuite comment différents chercheurs l'ont découpée en fonctions et en composantes, soit Jakobson (1963), Canale et Swain (1980) et Lyle Bachman (1990) et plus récemment, Glenn Fulcher (2003). Ces découpages, même s'ils se révèlent approximatifs et incomplets (Beacco, 2002 ; Fulcher, 2003) sont particulièrement nécessaires en évaluation, puisqu'ils permettent une certaine emprise sur la langue comme objet d'évaluation. En effet, avant de construire des tests permettant de mesurer l'aptitude à

communiquer, il importe de définir ce que nous voulons mesurer (Canale et Swain, 1980 ; Bachman, 1990 ; Bachman et Palmer, 1996 ; Fulcher, 2003 ; Cumming, 1996).

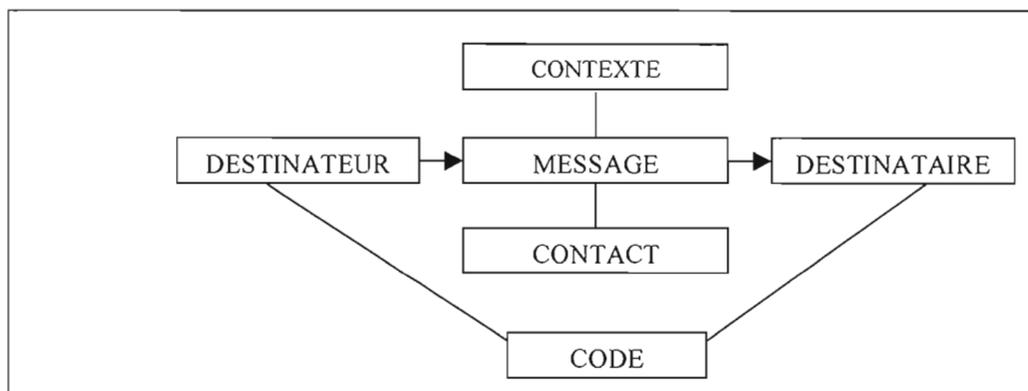
### 2.2.1 Genèse du concept de compétence communicative

Saussure, considéré comme le père de la linguistique moderne, a été le premier à tenter de modéliser la langue de façon scientifique (Bronckart, 1977). Dans son *Cours de linguistique générale*, publié en 1916 grâce à l'initiative de ses étudiants, il a suggéré l'opposition langue et parole, la langue étant définie comme un phénomène social, reposant sur un système de conventions et la parole, quant à elle, comme un acte individuel, pouvant s'écarter de la norme et finissant par la transformer (Bronckart, 1977). Outre cette opposition fondamentale, Saussure a aussi exposé le concept de «langage», s'ajoutant à l'opposition langue/parole, considéré comme une faculté propre à l'espèce humaine, faculté qui n'est pas sans rappeler la compétence linguistique que Chomsky affirmera 41 ans plus tard dans *Syntactic Structure* (1957) (Bronckart, 1977). Dès les débuts de la linguistique moderne, la langue est analysée comme étant un acte individuel balisé par un ensemble de règles partagées par une communauté.

Les travaux de Saussure ont inspiré la plupart des linguistes européens subséquents, dits structuralistes, de 1926 à 1940, regroupés dans le Cercle linguistique de Prague (Bronckart, 1977). Pour eux, la langue est un «système fonctionnel» (Bronckart, 1977, p.139). Plus précisément, ils «affirment d'une part qu'on ne peut comprendre aucun fait de langue sans faire référence au système auquel il appartient et d'autre part, que ce système est avant tout un système de moyens appropriés à un but» (Bronckart, 1977, p. 139-140). Se dessinent les deux fonctions essentielles de la langue : l'une «interne», se définissant par rapport au système qu'est la langue est l'autre «externe», propre au sujet (Bronckart, 1977).

Parmi les linguistes du Cercle de Prague, Jakobson (1963) s'est particulièrement illustré, entre autres par sa «théorie de la communication» (Leclerc, 1989, p.22). Il découpe cette fonction de communication en différents facteurs, illustrés au moyen d'un cadre, connu aujourd'hui sous le nom de «schéma de Jakobson» (Leclerc, 1989, p.22). Ces facteurs sont au

nombre de six : destinataire, message, destinataire, contact, contexte et code, que nous définirons dans les lignes qui suivent.



**Figure 2.1** Les facteurs de la communication, d'après Jakobson

Le schéma de Jakobson est composé d'abord d'un destinataire c'est-à-dire celui qui émet le message et d'un destinataire, qui reçoit le message. Cette relation est tantôt bidirectionnelle, dans la conversation courante par exemple, quand les rôles de destinataire et de destinataire peuvent s'inverser. Elle peut aussi être unidirectionnelle, dans le cas où le destinataire a peu de possibilités d'intervention. Le message varie en durée, contenu, forme et efficacité. Le message passe grâce à un contact, ou canal, qui est direct si le destinataire et le destinataire sont en présence l'un de l'autre ou indirect, si les deux protagonistes ne sont pas ensemble. Selon le canal choisi, la communication peut être immédiate ou différée. Cette communication s'effectue à travers un contexte ou référent, appelé aussi contexte situationnel. Ce référent, commun au destinataire et au destinataire est constitué des connaissances culturelles partagées par les deux parties. Finalement, le facteur de code s'ajoute à l'ensemble du schéma de communication de Jakobson, ce facteur est assujéti à l'ensemble de signes, linguistiques ou non, partagés de façon conventionnelle par les deux acteurs du schéma (Leclerc, 1989). Le rigoureux travail de clarification de Jakobson permet d'aborder la communication méthodiquement.

Parallèlement aux travaux de Jakobson et des autres linguistes du Cercle de Prague, le courant structuraliste s'est développé aux États-Unis, notamment avec la *Linguistic Society of America*, dont les membres se sont attardés à la description structurale de la langue, pour

en isoler les composantes et processus élémentaires (Bronckart, 1977). Leur conception béhavioriste de la langue est à l'origine de la prise de position de Chomsky, qui a mené à la définition de la notion de compétence linguistique. En effet, Chomsky (1965), en réaction aux structuralistes et aux béhavioristes, a proposé une théorie du langage qui insistait sur son aspect créatif. Avec cette théorie, Chomsky s'opposait fermement aux linguistes structuralistes à qui il reprochait de réduire la langue en un «inventaire systématique d'éléments» (in Bronckart, 1977, p. 171) et au courant béhavioriste en psychologie du langage (notamment à Skinner, pour qui la langue se modèle selon les contingences de renforcement du milieu) puisque la théorie de Chomsky est désireuse de «rendre compte de la «créativité» du sujet parlant en formalisant le modèle interne ou «faculté» de langage que possède tout locuteur» (Bronckart, 1977, p. 170). Pour Chomsky, le locuteur possède une grammaire interne qui lui permet de générer des énoncés originaux, respectant des règles syntaxiques précises : c'est la compétence grammaticale. Cette compétence s'actualise dans une performance, véritable reflet des processus cognitifs du locuteur. Chomsky a introduit les concepts de compétence et de performance, mais c'est le sociolinguiste Hymes, en 1972, qui a jeté les bases de ce qui allait être appelé plus tard l'approche communicative en didactique des langues. Il a ouvert la voie aux linguistes subséquents qui, au fil des recherches en linguistique, ont défini la capacité de l'individu à utiliser la langue comme la compétence communicative (Fulcher, 2003 ; Bachman, 1990; Canale et Swain, 1980; Dolz, 2002; Halliday, 1976; Hymes, 1984; Kramsch, 1986; Savignon, 1972; van Dijk, 1977). Hymes (1984) s'est érigé contre la notion de locuteur idéal imaginé par Chomsky pour illustrer sa théorie. Il argue qu'un tel locuteur est une abstraction et que les règles grammaticales ne permettent pas à elles seules de communiquer (Leung, 2005). Pour Hymes (1984), l'acte de communication englobe deux types de savoir : linguistique et socioculturel. Le savoir linguistique (la compétence grammaticale de Chomsky) touche à la connaissance de la langue et le savoir socioculturel, à la capacité d'utiliser la langue en contexte. Le chercheur observe que les locuteurs d'une même langue partagent un certain nombre de conventions sociales, aussi importantes que les règles grammaticales, pour communiquer entre eux. Il est intéressant de noter que Hymes, dans sa définition du savoir socioculturel, reprend grosso modo ce que Jakobson avait dit en 1963, soit que les contextes d'utilisation de la langue répondent à des règles dictées par les impératifs propres à ces contextes. Toutefois, selon

Hymes, il serait impossible (et vain) de présenter tous les contextes d'utilisation de la langue. Les contextes sont indépendants de la grammaire et ne sont pas codifiés comme elle.

Durant les années 70, Paul Grice, suivant le courant pragmatique en philosophie du langage, dans *Logic and conversation*, s'intéresse aux «règles de la conversation» («conversation rules»), plus précisément au sens conventionnel et non conventionnel des mots. Selon lui, les locuteurs arrivent à se comprendre même si les mots qu'ils prononcent prennent différents sens selon les contextes. Selon Grice (1979), la conversation est toujours orientée vers un but plus ou moins précis selon le contexte et les locuteurs, par convention, font généralement ce qu'il faut pour atteindre ce but. Il appelle ce phénomène le principe de collaboration. C'est ce principe qui fait que les individus, à condition qu'ils partagent des référents communs, peuvent se comprendre même s'ils enfreignent, en apparence, les lois de la conversation (quantité, qualité, relation et modalité)<sup>4</sup>. C'est ce qui se produit quand ils font de l'ironie, par exemple. Donc, comme Hymes (1984), Grice (1979) fait partie de ceux qui se sont attardés à montrer l'importance du contexte, des référents et des conventions implicites dans la communication et dont nous devons tenir compte dans toute tentative de modélisation de la compétence communicative. Dans les chapitres IV (présentation des résultats) et V (interprétation des résultats) du présent mémoire, nous verrons que les concepts définis par ces linguistes pragmatiques jouent un rôle de première importance dans l'évaluation.

### **2.2.2 Le modèle de Canale et Swain (1980)**

La notion de compétence de communication introduite par Chomsky a été adoptée rapidement et plusieurs modèles de communication ont été élaborés depuis. Nous en présenterons trois, qui ont été publiés au début des trois dernières décennies. D'abord, celui de Canale Swain, publié en 1980, puis celui de Lyle Bachman, publié en 1990 et finalement, celui de Glenn Fulcher, publié en 2003. Ces trois modèles ont une base commune et chacun a

---

<sup>4</sup> Nous verrons dans le chapitre VI, soit la présentation des résultats, que les évaluateurs ont tendance à pénaliser durement les sujets qui ne respectent pas le principe de collaboration.

le mérite de synthétiser les concepts importants, mais surtout, d'apporter des précisions sur la nature de compétence communicative, à la lumière des travaux de recherche en linguistique.

Le modèle de Canale et Swain, (1980), peaufiné par Canale (1983) a fait école. Dans un article phare, «Theoretical bases of communicative approaches to second language testing and speaking», ces derniers font, entre autres, un bilan des différents modèles proposés par les chercheurs en linguistique moderne pour décrire la langue et l'effet de ces modèles en enseignement des langues secondes. Au terme de leur impressionnante recension des écrits, ils arrivent à la conclusion que la définition de la compétence communicative est encore au «stade embryonnaire» («embryonic stage», p.21). Selon eux, s'il est généralement admis que la compétence communicative englobe une composante grammaticale, qui fait référence au code linguistique (syntaxe, morphologie, sémantique, lexicale, phonologie) et une composante sociolinguistique (encore plus ou moins clairement définie, selon eux) qui fait référence aux différents contextes d'utilisation de la langue, il manque aux modèles existants une composante essentielle, soit la composante stratégique, qui renvoie aux stratégies verbales et non verbales utilisées par le locuteur pour combler des lacunes dans les deux autres composantes. Les stratégies utilisées sont de deux natures selon eux : grammaticale (comme la paraphrase) et sociolinguistique (comme la façon de diriger un jeu de rôle lors d'une évaluation, par exemple). Ils ajoutent donc cette composante stratégique au modèle existant et ils s'efforcent de décrire plus spécifiquement ce qu'ils entendent par «compétence sociolinguistique», en synthétisant l'ensemble des descriptions établies par les chercheurs antérieurs. En gros, selon eux, cette composante fait référence aux règles socioculturelles propres à différents contextes et aux différents facteurs («topic, role of participants, setting, and norm of interaction» (p.30)) qui guident les choix linguistiques des locuteurs.

Déjà, les deux chercheurs entrevoient la nécessité d'ajouter une quatrième composante, reliée à l'organisation du discours, qui engloberait et dépasserait les notions de cohésion (la façon dont les concepts sont organisés dans le texte, à l'aide de connecteurs textuel) et de cohérence (l'adéquation du texte au contexte) définies par Widdowson (1978). C'est d'ailleurs ce que Canale fera trois ans plus tard. Il ajoutera la composante discursive,

qu'il définit comme l'habileté à relier des phrases pour produire un discours étendu qui ait du sens (Canale, 1983). Le modèle canonique de la compétence communicative était donc né.

**Tableau 2.1**  
Modèle de la compétence communicative  
de Canale et Swain (1980) et Canale (1983)

Composante	Définition	Exemples
Grammaticale	Concerne la maîtrise du code.	La concordance des temps de verbes; La prononciation des phonèmes.
Discursive	Concerne : La façon de combiner les formes grammaticales et le sens pour produire un texte; La grammaire du texte.	Le choix des arguments dans un texte qui vise à convaincre; La disposition des paragraphes dans un texte; Le choix des connecteurs logiques.
Sociolinguistique	Concerne la façon dont laquelle les énoncés sont formulés en tenant compte du contexte.	Le choix d'un niveau de langue; Le vouvoiement et le tutoiement; La compréhension d'un proverbe.
Stratégique	Concerne l'utilisation de stratégies verbales et non verbales qui visent à pallier des lacunes dans les autres composantes ou à renforcer la communication.	Faire des gestes; Demander de répéter; Paraphraser; Demander d'expliquer un mot inconnu.

### 2.2.3 Le modèle de Lyle Bachman (1990)

Nous venons de présenter le modèle de Canale et Swain (1980), auquel plusieurs chercheurs en linguistique font référence aujourd'hui. Parmi ces chercheurs, Lyle Bachman (1990) a proposé un autre modèle de «compétence langagière<sup>5</sup>» (language competence), qui a en quelque sorte pris le relais de celui de Canale et Swain (1980) et qui est devenu à son tour le modèle auquel la plupart des chercheurs en linguistique se réfèrent actuellement. Il

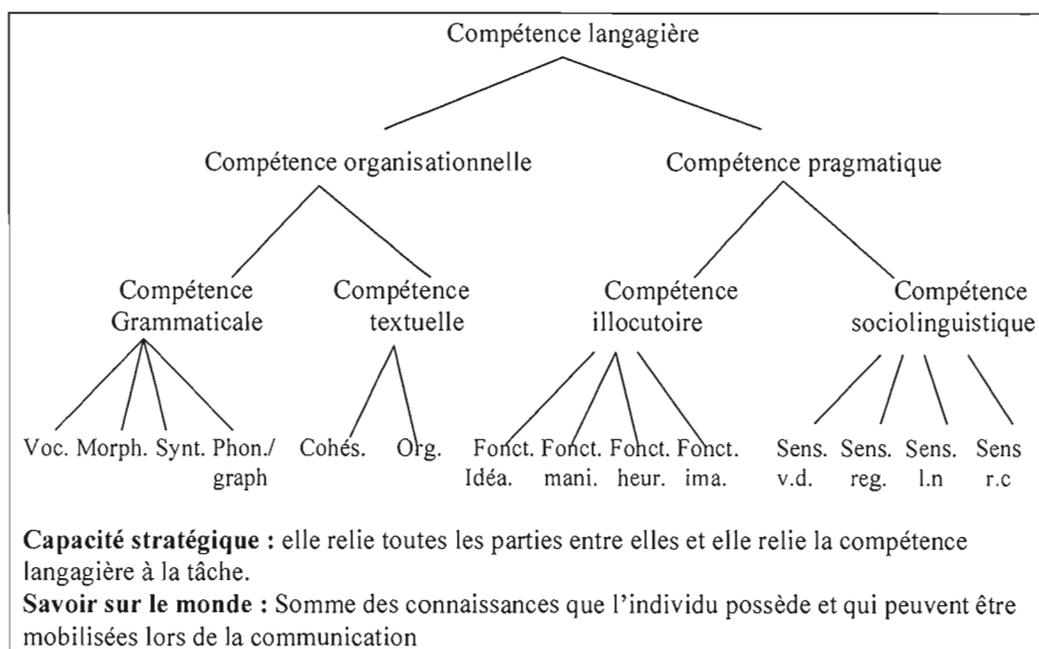
---

<sup>5</sup> Dans son modèle Bachman, fait une distinction entre la compétence langagière, qui réfère au langage et la compétence communicative, qui est plus large et englobe la compétence langagière, les connaissances de l'individu sur le monde, la compétence stratégique, les mécanismes psychophysiologiques et le contexte de la situation.

simplifie en quelque sorte le modèle de Canale et Swain (1980), puisque des quatre composantes générales (linguistique, discursive, sociolinguistique et stratégique), il n'en retient que deux : la composante organisationnelle et la composante pragmatique. Son découpage en deux parties de la compétence langagière vient du constat que lui et d'autres chercheurs (Allen et al, 1983; Bachman et Palmer, 1982) ont fait : il est aisé de séparer les composantes grammaticale et discursive de la composante sociolinguistique, mais il est très complexe, voire impossible, de séparer la composante grammaticale de la composante discursive. Les résultats de ses recherches l'ont amené à définir un nouveau modèle de compétence langagière. Son modèle la divise donc deux grandes composantes, l'une faisant référence au message ou pour reprendre la terminologie de Jakobson, au code utilisé pour transmettre le message et l'autre, au contexte de transmission du message. Sa première composante, la composante organisationnelle, englobe deux sous-composantes : la composante grammaticale, soit : le vocabulaire, la morphologie, la syntaxe et la phonologie/graphie et la composante textuelle, qui englobe la cohésion et l'organisation rhétorique du discours, oral et écrit. La façon d'organiser l'information et la sensibilité aux caractéristiques propres aux différents types de discours (narratif, descriptif, argumentatif par exemple) sont comprises dans cette composante. Y sont également incluses les conventions propres à la conversation (la façon de prendre ou de laisser la parole et de choisir les sujets de conversation, par exemple). Même si Bachman reconnaît que la conversation répond à ses propres règles, il ne fait pas de l'ensemble de ces règles une composante distincte de la compétence textuelle, puisque selon lui, les règles de la conversation sont tout de même régies par des lois organisationnelles de la même nature que celles qui régissent le discours en général.

La deuxième grande composante du modèle de Bachman est la composante pragmatique. Cette composante s'apparente à la composante sociolinguistique de Hymes (1984) et de Canale et Swain (1980). Elle met en évidence l'importance du contexte d'énonciation et des référents culturels dont les locuteurs doivent tenir compte pour livrer un message. Bachman (1990) la divise en deux sous-composantes : illocutoire et sociolinguistique. La sous-composante illocutoire fait référence aux différentes fonctions du langage (idéationnelle (échange d'information concernant les sentiments), manipulateur

(accomplir quelque chose en le disant), heuristique (se servir de la langue à des fins d'apprentissage) et imaginative (jeu de mots : poésie, blague, métaphore)). La sous-composante pragmatique, quant à elle, englobe : sensibilité aux variations dialectales, sensibilité aux registres, sensibilité à repérer le langage naturel et la sensibilité à repérer les références culturelles.



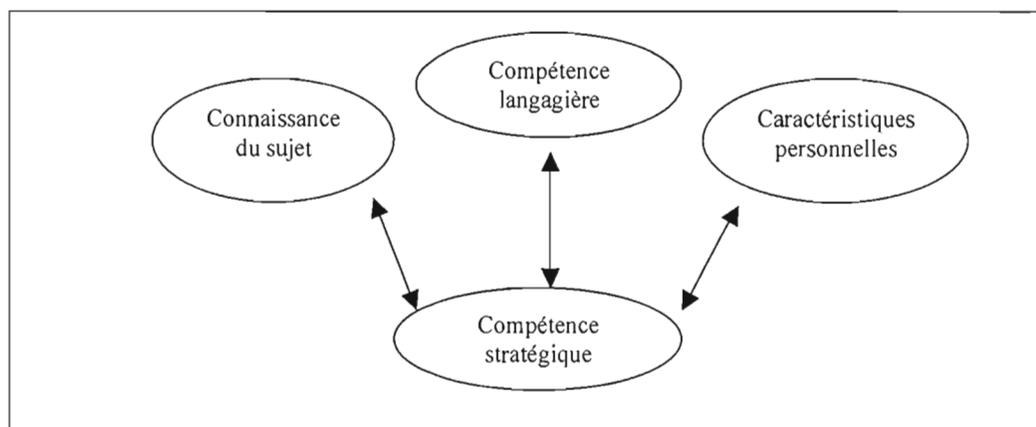
**Figure 2.2** Modèle de la compétence langagière de Bachman (1990)

Contrairement à Canale et Swain (1980), Bachman (1990) présente la composante stratégique non pas comme une façon de pallier les lacunes dans les autres composantes, mais bien comme la disposition du locuteur à puiser dans toutes les composantes pour exprimer un discours qui ait du sens. C'est par cette fonction qu'occupe la compétence stratégique que Bachman (1990) montre que les composantes interagissent. L'originalité de son modèle est donc de faire ressortir l'aspect dynamique de la compétence communicative : toutes les composantes interagissent entre elles grâce à la capacité stratégique du locuteur.

En outre, dans son modèle, Bachman (1990) accorde une place au «savoir sur le monde» («knowledge of the world». p.85). Dans son modèle, ce savoir ne fait pas partie intégrante de la compétence langagière, il est situé juste à côté et c'est la compétence

stratégique qui permet de prendre des morceaux de ce savoir sur le monde et de les exprimer grâce à la compétence langagière. Ce savoir n'est pas à proprement parler linguistique, mais il est plutôt la somme des connaissances de l'individu. De plus, dans son modèle, la composante stratégique acquiert des fonctions qu'elle n'avait pas dans le modèle de Canale et Swain (1980). Il ne s'agit plus seulement de recourir à sa capacité stratégique pour utiliser adéquatement ses ressources personnelles, mais aussi de s'en servir spécifiquement dans différents contextes, en contexte d'examen par exemple. Il ajoute trois aires d'utilisation de la compétence stratégique, soit 1) se fixer un but ou identifier la tâche demandée; 2) évaluer la tâche et évaluer sa propre façon de répondre à la tâche; 3) planifier sa réponse en fonction de ses connaissances. La compétence stratégique est donc selon lui du domaine métacognitif et elle est garante du succès lors d'un test visant à mesurer la compétence communicative.

Plus tard, en 1996, avec Adrian Palmer, Bachman est revenu sur ce modèle et ils l'ont mieux adapté au contexte de l'évaluation. Ils ont ajouté au schéma précédent une nouvelle dimension, celle des caractéristiques personnelles de l'individu (âge, sexe, nationalité, statut, langue maternelle, scolarité, degré et type de préparation au test). Comme le savoir sur le monde, cette nouvelle dimension du schéma ne fait pas partie de la compétence communicative, mais elle doit, selon les chercheurs, être prise en compte lors de la création de tests servant à mesurer la compétence communicative.



**Figure 2.3** Interaction entre différentes composantes (Bachman et Palmer, 1996)

### 2.2.3 Le modèle de Glenn Fulcher (2003)

Nous avons choisi de présenter le modèle d'un dernier chercheur, Glenn Fulcher, publié en 2003, parce qu'il a la particularité d'adapter les modèles présentés précédemment au discours oral, plutôt que de modéliser le langage en général. Sa contribution à la recherche sur l'évaluation de la langue orale est donc particulièrement importante. Fulcher (2003) dit s'inspirer largement du modèle de Bachman et Palmer (1996), mais il s'en distancie en partie, puisque selon lui, comme il est impossible de percevoir l'activité métacognitive d'un individu, il est inutile de baser l'évaluation sur un modèle dont certaines composantes sont imperceptibles. Il ne rejette nullement la pertinence des recherches portant sur les mécanismes sous-jacents du langage, mais les déclare peu pratiques pour l'évaluation. Suivant cette idée, il trouve peu utile l'opposition de compétence et performance, puisque seule la performance est visible. Donc, son travail de modélisation s'oriente sur les parties saillantes du langage. Son objectif est de choisir de façon rationnelle et empirique, ce qu'il faut inclure dans le construit à tester.

Son modèle est divisé en cinq grandes catégories : la compétence langagière («language competence»), la capacité stratégique («strategic capacity»), le savoir textuel («textual knowledge»), le savoir pragmatique («pragmatic knowledge») et le savoir sociolinguistique («linguistic knowledge»). Comme dans le modèle de Bachman (1990), chacune de ces composantes est sous-divisée en sous-composantes. La première composante, la compétence langagière, reprend pour l'essentiel la composante grammaticale de Bachman (1990) et de Canale et Swain (1980) et ajoute les sous-composantes propres à l'oral. Fulcher (2003) sépare cette composante en trois : d'abord, la phonologie, qui englobe la phonétique, les règles phonologiques et la prosodie. Sa deuxième composante est la précision, qui englobe la syntaxe, le vocabulaire et la cohésion. Fulcher (2003) n'explique pas précisément ce qu'il entend par cohésion. Il faut se référer au modèle de Bachman (1990) pour le comprendre. Puisque Fulcher dit s'être inspiré grandement du modèle de ce dernier pour arriver à son nouveau modèle, ce qu'il n'explique pas dans sa monographie se trouve plus

clairement détaillé dans les travaux de Bachman (1990)<sup>6</sup>. À cette sous-composante de la précision, Fulcher ajoute l'aisance («fluency»). L'originalité de l'auteur est d'intégrer ce concept, l'aisance, au sein d'une de ses composantes, plutôt que de le laisser flotter au dessus des autres composantes, comme le faisaient les modèles précédents. En effet, dans les modèles précédents, l'aisance ne fait pas partie explicitement des schémas de la compétence communicative. Peut-être est-ce parce que ces modèles schématisent la compétence communicative dans son ensemble, et pas spécifiquement le discours oral : l'aisance à communiquer, même si elle se manifeste dans l'écriture (Germain et Netten, 2002), n'y est pas aussi perceptible que dans le discours oral. L'aisance à communiquer est saillante dans le discours oral, d'où sa place à côté des autres composantes dans le schéma de Fulcher (2003). D'ailleurs, des études révèlent que l'aisance à communiquer est aisément perçue par les évaluateurs et est généralement un critère d'évaluation consensuel (Kormos et Dénes, 2004 ; Chambers, 1997). Fulcher l'a donc intégrée comme une des composantes de son modèle de la compétence communicative. Cette composante de la compétence communicative fait référence au rythme de la parole du locuteur : au nombre et à la longueur des pauses, aux répétitions de mots ou de syllabes, aux hésitations et aux reformulations. Les adjectifs «lent», «saccadé», «hésitant» ou «inégal», opposés à «fluide», «égal» (Fulcher, 2003, p. 30) sont souvent utilisés pour décrire le discours d'une personne et sont associés à l'aisance à communiquer.

Nous ouvrons ici une parenthèse pour approfondir ce concept d'aisance que d'autres chercheurs, Germain et Netten (2002), se sont attardés à définir, pour en faire ressortir l'utilité dans l'évaluation des langues secondes. Ces derniers ont proposé une définition plus opérationnelle de l'aisance (et aussi de la précision) et ont bâti une grille d'évaluation de la production orale (et écrite) en langue seconde, fondée sur les critères de précision et d'aisance. Pour eux, l'aisance à communiquer dépasse la définition de Fulcher (2003). Elle se traduit non seulement par le rythme du discours du locuteur, mais aussi par la quantité de

---

<sup>6</sup> C'est d'ailleurs une critique que lui adresse Annie Brown (2005), soit de mal définir ses concepts et d'obliger le lecteur à chercher ailleurs la définition de certains concepts clés. En consultant les travaux de Bachman (1990), nous voyons que dans son modèle, la notion de cohésion est reliée à la compétence textuelle et concerne l'organisation générale du discours et la rhétorique et pas la compétence grammaticale. Il devient alors plutôt difficile de déduire ce que Fulcher a voulu inclure dans ce qu'il appelle la cohésion dans la précision.

phrases qu'il réussit à formuler et par la complexité des phrases (c'est-à-dire des phrases comportant des adjectifs, des adverbes et des subordonnées). La notion de complexité des phrases n'apparaît pas dans le modèle de Fulcher (2003).

Germain, Netten et Movassat (2004) remarquent que dans l'évaluation de la production orale d'élèves du primaire, «il s'est avéré impossible, dans le cas de la production orale, de distinguer les critères de précision linguistique du critère d'aisance à communiquer.» Ils ajoutent que «Il se peut que ces deux aspects soient si liés, à l'oral, qu'il est impossible de les distinguer l'un de l'autre. [...] Les caractéristiques de fluidité, de pause et de sens, preuves d'aisance à l'oral, risquent d'être affectées par la précision.» (p.326). Ce constat les a amenés à proposer le concept de «praisance» : la précision dans l'aisance. Selon eux, «la précision n'est pas seulement un savoir, mais aussi un savoir-faire, intimement lié à l'aisance» (p.326). Toutefois, la compétence grammaticale n'est pas toujours prédictive de l'aisance à communiquer.

Revenons au modèle de Fulcher (2003). Nous avons déjà défini ce qu'il appelle la compétence langagière (la composante qui englobe la phonologie, la précision et l'aisance). À cette compétence s'ajoute la capacité stratégique. Si, pour Canale et Swain (1980), pour Bachman (1990) et pour Bachman et Palmer (1996) la capacité stratégique faisait référence aux habiletés métacognitives, donc «invisibles», qui assurent l'interaction entre les différentes composantes de la compétence communicative, dans son modèle, Fulcher (2003), a choisi de limiter cette capacité aux phénomènes observables. Il sépare la capacité stratégique en deux grandes catégories : les stratégies de réussite (achievement strategies) et les stratégies d'évitement (avoidance strategies). La sur-généralisation, l'approximation, la paraphrase, la circonlocution, la restructuration du message, la coopération (quand le locuteur demande l'aide de son interlocuteur pour formuler son message) et l'alternance codique sont des stratégies de réussite. Fulcher (2003) inclut parmi elles des stratégies non linguistiques, soit la gestuelle et l'utilisation de l'environnement. Toutes ces stratégies sont dites de réussite, puisqu'elles sont des moyens utilisés par le locuteur pour atteindre un objectif de communication. À l'opposé, les stratégies d'évitement, ou d'omission, consistent pour le locuteur à ne pas utiliser une forme langagière qu'il ne maîtrise pas: la voix passive par

exemple, ou encore un temps de verbe comme l'imparfait. Il s'agit alors de l'évitement d'une forme. Selon Fulcher, cette stratégie peut être difficile à détecter. De plus, si le locuteur n'utilise pas une forme langagière, cela ne veut pas nécessairement dire qu'il ne la maîtrise pas : l'absence d'une forme ne veut pas toujours dire son inexistence. En plus de l'évitement d'une forme, le locuteur peut avoir recours à l'évitement fonctionnel, plus «sérieux» selon Fulcher (2003). Cet évitement consiste à littéralement abandonner la conversation, ou à décider de ne pas parler d'un sujet. À ces deux manifestations extrémistes, s'en ajoute une troisième, plus courante, l'évitement sémantique, qui consiste à sur utiliser un nom générique, par exemple «chose», au lieu d'un terme plus précis.

À la compétence langagière et à la capacité stratégique s'ajoutent trois types de savoirs dans le modèle de Fulcher (2003) : le savoir textuel, le savoir pragmatique et le savoir sociolinguistique.

Dans le modèle de Bachman (1990), le savoir textuel (appelé «textuel competence») faisait référence à la capacité de structurer son discours, au moyen de connecteurs, pour en assurer l'intelligibilité et la progression. La rhétorique entrait dans cette composante. Pour Canale et Swain (1980), il s'agissait de la composante discursive, regroupant à peu près les mêmes caractéristiques que son vis-à-vis dans le modèle de Bachman. Dans le modèle de Fulcher (2003) tout comme dans les modèles précédents, le savoir textuel fait référence à la structure du discours, mais il se différencie des définitions de Bachman (1990) et de Canale et Swain (1980) parce qu'il s'intéresse uniquement à la structure de la conversation et non aux autres types de discours. Cette composante comprend trois sous-composantes. D'abord, il y a le tour de parole, qui porte sur les conventions plus ou moins explicites auxquelles se plient les interlocuteurs lors d'une conversation, pour déterminer quand ils doivent prendre la parole ou la laisser à l'interlocuteur. Puis, vient la sous-composante, appelée «paires adjacentes». Ces paires complémentaires sont considérées comme l'unité de conversation la plus fondamentale, qui consiste à regrouper les deux parties d'une structure discursive binaire telle que question et réponse, salutations et salutations, invitation et refus ou acceptation, etc. La dernière des trois sous-composantes du savoir textuel est la capacité à ouvrir et à fermer la conversation. La façon de commencer et de finir une conversation varie et les locuteurs doivent se conformer aux contingences spécifiques aux différents types de conversation :

prise de rendez-vous, demande d'information, conversation amicale sur un sujet d'actualité. Les interlocuteurs doivent savoir comment ouvrir et surtout, comment et à quel moment clore la conversation. Nous voyons que le savoir textuel, selon le modèle de Fulcher (2003), fait donc appel non seulement aux qualités d'émetteur d'un locuteur, mais aussi à ses qualités de récepteur, puisque c'est grâce à sa capacité d'écoute et à sa compréhension de son interlocuteur qu'il peut montrer son savoir textuel.

Reste les savoirs pragmatique et socioculturel, qui correspondent respectivement aux fonctions illocutoire et sociolinguistique définies plus haut dans le modèle de Bachman (1990). Fulcher (2003) reprend les définitions de Bachman (1990) et il serait donc redondant de les énoncer ici.

En plus de ces composantes, Fulcher (2003) entrevoit la nécessité d'en ajouter une dernière : la composante interactive («interactional competence»). Il croit que la limite des schémas de Canale et Swain (1980) et de Bachman (1990) est de décrire la compétence communicative de façon unidirectionnelle, en mettant l'accent sur un seul locuteur ou en limitant l'interaction avec l'interlocuteur aux composantes textuelle, pragmatique et stratégique. Cette remarque est cohérente avec la théorie de la compétence interactive («Interactional Competence Theory») définie par Kramsch en 1986 qui démontre que la conversation se construit avec le ou les partenaires. Selon les définitions, la compétence interactive est souvent associée en partie au savoir textuel et à la composante stratégique. C'est pourquoi plusieurs chercheurs ne voient pas la nécessité de l'inclure dans un modèle de la compétence langagière. De plus, elle concernerait plus le comportement d'un individu dans un groupe et sa personnalité que ses capacités linguistiques. Elle est à ce sujet plus proche de la psychologie que de la linguistique. Bien que cette composante ne fasse pas partie des modèles traditionnels, selon Fulcher (2003), elle pourrait s'y tailler une place. En effet, une conversation se construit par l'interaction des participants, par leur collaboration. En plus des composantes pragmatique, stratégique et textuelle, la compétence interactive mobiliserait plus spécifiquement des aspects de la personnalité des locuteurs. C'est pourquoi Fulcher (2003) ne l'inclut pas finalement dans son modèle : elle est encore mal définie et il lui semble

controversé de tenir compte de la personnalité d'un individu pour juger de sa compétence communicative<sup>7</sup>.

**Tableau 2.2**  
Modèle de la compétence communicative de Fulcher (2003)

<b>Composantes</b>	<b>Sous composantes</b>	
Compétence langagière	Phonologie	Prononciation Accent tonique Intonation
	Précision	Syntaxe Vocabulaire Cohésion
	Aisance	Hésitation Répétition Re-sélection de mots inappropriés Restructuration de phrases Cohésion
Capacité stratégique	Stratégies de réussite	Sur-généralisation Paraphrase Circonlocution Restructuration Stratégie de coopération Alternance codique Stratégies non linguistiques
	Stratégies d'évitement	Évitement formel Évitement fonctionnel
Connaissances textuelles	Structure de la conversation	Tours de parole Paires adjacentes Ouverture et fermeture
Savoir pragmatique	Être approprié Implication Faire en disant	
Savoir sociolinguistique	Savoir situationnel Connaissance sur le sujet Savoir culturel	

Nous venons de présenter trois modèles importants, celui de Canale et Swain (1980), de Bachman (1990) et de Fulcher (2003) qui tentent chacun à leur façon de modéliser la compétence communicative. Les deux premiers modèles sont devenus les références habituelles des chercheurs en linguistique. Le dernier, qui s'inspire largement des deux premiers, a la particularité de modéliser la partie saillante de la langue, conférant une certaine

---

<sup>7</sup> Nous verrons cependant dans les chapitres IV et V que les évaluateurs ont souvent tendance à justifier leur note en invoquant la personnalité des individus évalués.

emprise sur elle. Notons au passage que l'analyse des résultats de notre recherche, dans le chapitre V, reposera en grande partie sur la modélisation de Glenn Fulcher (2003). Mais avant d'y arriver et afin de conclure la partie de notre cadre théorique consacrée à la compétence communicative, nous verrons un dernier découpage de la langue, soit le découpage en quatre habiletés.

### **2.3 Les quatre habiletés**

Outre ce découpage de la compétence communicative en différentes composantes, nous retrouvons dans la littérature une autre façon de présenter la langue, soit en quatre habiletés. Au début des années 1960, Lado (1962) et Carroll (1961, 1983) ont modélisé la langue en la découpant en quatre habiletés, soit : 1) la compréhension orale, 2) la compréhension écrite, 3) la production orale et 4) la production écrite. Ces quatre habiletés, même si elles sont fortement corrélées, sont considérées distinctes puisqu'elles peuvent se développer inégalement (Carroll, 1983). La popularité de ce découpage en quatre habiletés ne s'est jamais démentie : l'enseignement des langues s'articule bien souvent, encore aujourd'hui, selon la taxinomie de Lado et Carroll, notamment dans le programme général de formation du MELS (2006). Certes, il peut être discutable de présenter la compétence communicative sous l'angle des quatre habiletés, puisque certains discours, tels l'exposé oral informatif et le texte explicatif, qui appartiennent respectivement à la production orale et à la production écrite, partagent une structure commune et se ressemblent davantage que le débat d'idées et le discours par exemple, classés cependant dans la même habileté, soit la production orale (Bachman et Palmer, 1996 ; Biber, 1998). De plus, lors d'une interaction orale, les habiletés de production orale et de compréhension orale sont sollicitées. Évaluer un exposé oral nécessite des critères d'évaluation différents de ceux d'un jeu de rôle, notamment en ce qui concerne la composante textuelle. Pour ces raisons, aujourd'hui, la pertinence de découper la langue de cette façon est remise en question, en pédagogie et en évaluation.

Bref, nous pouvons modéliser la langue de plusieurs façons, nous pouvons la décomposer en plusieurs composantes, mais il est compliqué de repérer séparément les

composantes de la compétence communicative, parce qu'elles sont difficiles à isoler dans le discours (Bachman, 1990; Carroll, 1983; Germain et Netten, 2002; Orr, 2002; Fulcher, 2003 ; Walter, 2004). Les différentes tentatives de modélisation demeurent imparfaites. Mais, pour ceux dont l'intérêt est d'évaluer la compétence communicative, l'important est, selon Fulcher (2003), d'arriver à définir le plus clairement possible le construit qu'ils veulent évaluer. Si l'évaluation porte uniquement sur la composante grammaticale, il faut définir le mieux possible cette composante. Si c'est la capacité stratégique que nous voulons évaluer, c'est cette capacité que nous devons décrire. Le tout est de définir un construit cohérent avec ce que nous cherchons à évaluer.

Maintenant que nous avons présenté des modèles de la compétence communicative et plus spécifiquement, de la production orale, modèles qui nous aident à définir le construit que nous cherchons à évaluer, nous définirons la notion d'évaluation en général, et plus spécifiquement, l'évaluation de la production orale.

## **2.4 L'évaluation**

L'évaluation, en pédagogie, est aussi appelée docimologie C'est la science qui étudie les différents moyens d'évaluer. Ses formes (entrevue, test à choix de réponse multiples, essai, etc.) et ses fonctions (diagnostique, formative, sommative, etc.) sont multiples. Dans ce mémoire, nous nous intéressons à l'évaluation sommative, c'est-à-dire celle qui sert à assigner un grade à une personne ou à déterminer si un étudiant a atteint les objectifs d'un cours au terme de sa formation (Gronlund, 1985). La définition de l'évaluation que nous développerons dans la section suivante, bien qu'elle ne s'y confine pas exclusivement, sera cohérente avec ce type d'évaluation.

Nous pouvons résumer le processus d'évaluation sommairement en disant qu'il comporte quatre étapes : 1) définir ce que nous voulons évaluer ; 2) colliger l'information ; 3) interpréter l'information et 4) prendre une décision (Genesee et Upshur, 1996). La première partie de notre cadre théorique portait sur la définition de la compétence communicative, plus

spécifiquement la production orale. Elle se rapportait donc à la première étape de l'évaluation. Dans la partie suivante de notre cadre théorique, nous parlerons des autres étapes de l'évaluation. Pour ce faire, nous définirons les concepts de validité et de fidélité. Nous terminerons cette section consacrée à l'évaluation en définissant deux approches utilisées en évaluation, la première, éduométrique et la deuxième, psychométrique.

#### **2.4.1 La validité**

L'évaluation, peu importe la matière évaluée, doit répondre aux deux critères suivants : la validité et la fidélité (ou fiabilité) (Gronlund, 1985; Bachman, 1990). Dans cette section, nous définirons le premier critère.

La validité d'un test dépend de sa capacité à mesurer exactement ce qu'il cherche à évaluer. De façon générale, elle concerne la pertinence de l'information recueillie en regard de ce que nous prétendons mesurer (Bachman, 1990; Gronlund, 1985; Lussier et Turner, 1995; Genesee et Upshur, 1996). La validité se rapporte donc à la façon de colliger l'information (la deuxième étape de l'évaluation) (Gronlund, 1985). Elle se divise en plusieurs types. En contexte de classe, les types de validité les plus connus sont les quatre suivants : de contenu, critérielle, de construit et apparente (Lussier et Turner, 1995). Le premier type, la validité de contenu, indique l'adéquation entre la tâche demandée lors d'un test et ce que nous prétendons mesurer. Le second type, la validité critérielle, concerne la relation entre le test et le critère. Elle se vérifie lorsque nous comparons «les résultats obtenus à un test avec un autre critère comme les notes à un autre test» (Lussier et Turner, 1995, p.181). Le troisième type de validité, soit de construit, est le reflet de «l'adéquation entre la structure d'un instrument de mesure et les modèles théoriques qu'il prétend mesurer» (Lussier et Turner, 1995, p.181). Enfin, le quatrième type de validité, la validité apparente, est l'impression générale produite sur les évaluateurs et les évalués de la pertinence du test (Lussier et Turner, 1995).

Nous voyons que la validité ne repose pas uniquement sur l'échantillonnage ou sur le test lui-même, mais aussi sur les critères d'évaluation retenus pour la correction du test (Bachman, 1990; Gronlund, 1985). Par exemple, si nous demandons aux élèves de rédiger un

texte, sans imposer de limites de nombre de mots, nous devons pondérer la correction afin de ne pas pénaliser un élève qui aurait écrit un texte très long et qui, par le fait même, aurait commis plus d'erreurs qu'un élève qui aurait écrit un texte plus court. Évaluer la compétence orthographique de l'élève en fonction du nombre de fautes total ne serait pas valide.

En contexte scolaire, la validité d'un test repose aussi sur son adéquation au contenu pédagogique (Wall et Alderson, 1996). Un test qui porterait sur autre chose que la matière vue en classe n'informerait pas au sujet de la compétence acquise par les élèves pendant le cours. Les inférences que nous pourrions faire au sujet du degré de maîtrise des élèves des concepts enseignés, à partir des résultats obtenus, ne seraient pas valides.

#### **2.4.2 La fidélité**

La fidélité d'un test quant à elle repose sur sa capacité à rendre compte du degré de maîtrise de l'étudiant au sujet de l'aspect mesuré : jusqu'à quel point la note dépend de l'aptitude qu'on veut mesurer, et non pas d'autres facteurs comme la fatigue ou le stress de la personne évaluée ; la sévérité de l'évaluateur ; la chance ou le hasard lors du choix d'une réponse dans un test à choix multiple, etc? C'est une des questions à poser pour connaître le degré de fidélité d'un test (Bachman, 1990; Gronlund, 1985; Lussier et Turner, 1995). La fidélité s'applique à la personne testée, à l'évaluateur et au test lui-même. La fidélité est donc reliée à la troisième et à la quatrième étape de l'évaluation, soit l'interprétation et la prise de décision (Genesee et Upshur, 1996).

D'abord, puisque la fidélité concerne la constance avec laquelle un test mesure ce qu'il est censé mesurer, pour qu'un test soit fidèle, il faut que celui qui le passerait à deux intervalles différents obtienne le même score. Or, l'état de la personne (la fatigue et le stress, par exemple) et l'environnement dans lequel se déroule le test sont des facteurs qui peuvent avoir une influence sur le résultat et qui ne sont pas reliés à la compétence de l'évalué (Fulcher, 2003 ; Bachman, 1990 ; Bachman et Palmer, 1996). Pour s'assurer que le test rende compte fidèlement de la compétence d'un élève, nous pouvons le tester à plusieurs occasions ou encore nous assurer que l'élève est dans de bonnes dispositions et qu'il

comprend bien ce qu'il doit faire (Genesse et Upshur, 1996). Mais si les évaluateurs ne peuvent faire grand chose en ce qui concerne l'état de la personne, ils peuvent en revanche, s'assurer que les conditions de passation d'un test sont adéquates. Des études (Fulcher, 2003 ; Oller, 1983) révèlent la grande influence que peut avoir l'environnement sur les résultats d'un test.

De plus, la fidélité concerne l'interprétation des résultats. Elle repose entre autres sur la constance intra-évaluateur : un test répond au critère de fidélité si le jugement d'un même évaluateur reste constant d'une évaluation à une autre, c'est-à-dire si ses critères de correction sont appliqués avec la même rigueur. Un exemple de manque de constance intra-évaluateur serait celui d'un enseignant qui corrigerait 20 textes dans la même journée et qui deviendrait moins attentif à mesure que sa fatigue augmenterait et qui laisserait passer des erreurs qu'il n'aurait pas laissées passer au début de sa journée de correction. La baisse d'attention causée par sa fatigue diminuerait sensiblement la fidélité du test.

La constance inter-évaluateurs quant à elle est assurée si deux évaluateurs différents qui interprètent le même test arrivent au même score. Un exemple de manque de constance inter-évaluateurs serait d'obtenir deux notes différentes pour un même texte corrigé par deux évaluateurs différents. L'absence de constance inter-évaluateurs diminuerait donc la fidélité du test. En général, plus l'objet évalué est sujet à l'interprétation, plus la constance intra et inter-évaluateurs est difficile à obtenir (Bachman, 1990; Gronlund, 1985; Lussier et Turner, 1995). Une façon d'assurer une certaine constance intra et inter-évaluateurs est de créer une clé de correction univoque, qui n'est pas sujette à différentes interprétations (Brown, 1983; Gronlund, 1985; Shohamy, 1983). Les conclusions de ces chercheurs au sujet de l'absence de fidélité inter-évaluateur sont directement reliées à notre objectif de recherche.

Une autre façon de pallier l'absence de constance inter-évaluateurs est de concevoir des tests dont les réponses ne sont pas sujettes à interprétation. Les tests à choix multiples, les vrais ou faux et les tests à réponses courtes facilitent la constance intra et inter-évaluateurs, grâce au peu de risques de biais dans la mesure. Par contre, de tels tests ne sont pas absolument fidèles, dans la mesure où la réponse peut dépendre de la chance ou du hasard et non de la compétence de la personne testée. Imaginons deux élèves, de compétence égale, qui

passent le même test composé de 50 questions. Les deux élèves connaîtraient tous deux les réponses aux 30 premières questions, mais ils ignoreraient les réponses des 20 dernières questions et y répondraient de façon aléatoire. Le premier élève, le plus chanceux, choisirait 15 fois sur 20 la bonne réponse. Le deuxième, moins chanceux, ne choisirait que cinq bonnes réponses. Le premier élève obtiendrait donc un résultat de 45/50 et le deuxième, 35/50. Le deuxième élève est-il moins compétent que le premier ? Non, il est simplement moins chanceux. Un tel test, même s'il assurerait la constance entre les évaluateurs, ne serait pas fidèle puisqu'il ne mesurerait pas absolument la compétence des deux élèves.

Nous devons préciser en outre que les concepts de validité et de fidélité sont complémentaires (Bachman, 1990 ; Genesse Upshur, 1996). En fait, la fidélité est une condition sine qua non de validité: si les résultats d'un test sont impossibles à interpréter, ou s'ils sont mal interprétés, le test est nécessairement «non valide», peu importe la rigueur avec laquelle il a été conçu. C'est-à-dire qu'un test valide est nécessairement fidèle, mais qu'un test fidèle n'est pas nécessairement valide.

Nous venons de définir deux concepts de base en évaluation, concepts que nous devons garder en tête pour bâtir des instruments de mesure et d'évaluation. Dans la partie suivante, nous parlerons de deux approches évaluatives : psychométrique et édumétrique, qui concernent la façon d'interpréter les résultats.

### **2.4.3 L'approche psychométrique et l'approche édumétrique**

L'approche psychométrique, aussi appelée interprétation normative (Cziko, 1983), consiste à interpréter la performance d'un individu en fonction de sa position par rapport à un groupe défini. Le but d'un test relevant d'une approche psychométrique est de différencier les individus les uns par rapports aux autres, de les hiérarchiser. Généralement, les items du test sont de difficulté moyenne à complexe, puisque le but est de séparer les «forts» des «faibles». Sans référence à la norme, le résultat ne peut être interprété (Gronlund, 1985). Par exemple, si nous apprenons qu'une étudiante a obtenu 46% dans un test de vocabulaire, cela ne nous éclaire pas beaucoup sur son degré de maîtrise du vocabulaire. Mais si nous savons que 3000

personnes qui ont passé le même test que cette étudiante ont obtenu une moyenne de 25%, nous pouvons déduire que cette étudiante a un vocabulaire plus vaste que la moyenne (ou que les 3000 individus sont plus jeunes, ou encore qu'ils n'ont pas la même langue maternelle que l'étudiante). Donc, dans une approche psychométrique, la note en elle-même n'a de signification que si nous la comparons avec la note d'un groupe et que nous connaissons la composition du groupe de comparaison.

L'approche éduométrique, ou interprétation critériée (Cziko, 1983) quant à elle s'applique plutôt à décrire la compétence d'un individu en regard des critères d'évaluation élaborés. En enseignement, cette approche a tendance à supplanter la précédente, parce que le but des enseignants n'est pas tant de discriminer les «bons» et les «mauvais» étudiants, mais de voir jusqu'à quel point ils ont intégré les connaissances enseignées (Cziko, 1983). Les items de l'examen sont sélectionnés non pas selon leur degré de difficulté, mais selon leur lien avec la compétence à évaluer. L'interprétation du résultat se fait donc en comparant la performance de la personne évaluée avec la définition de la tâche demandée.

Nous avons défini brièvement quatre concepts de base de l'évaluation. Dans la partie suivante, nous appliquerons ces concepts à l'évaluation de la compétence communicative, plus spécifiquement de la production orale.

## **2.5 L'évaluation de la production orale**

L'évaluation des apprentissages évolue au même rythme que les connaissances en pédagogie. L'évaluation de la production orale n'échappe pas à cette règle. Elle s'adapte constamment aux théories nouvelles (Fulcher, 2000 ; Fahardy, 1983). L'avènement de l'approche communicative a eu une influence majeure en évaluation (Fulcher, 2000). Elle a forcé les chercheurs à revoir les critères de validité et de fidélité sous l'éclairage fourni par la définition de la compétence communicative.

### **2.5.1 La validité de l'évaluation de la production orale**

Les tests utilisés pour mesurer la production orale, antérieurs à la diffusion de l'approche communicative en langue seconde, seraient caractérisés par un très grand souci de fidélité, souci qui dans certains cas aurait pour effet d'en diminuer la validité. Ces tests apparaissent le plus souvent sous la forme vrai ou faux, réponse à choix multiples ou sur la traduction. Souvent, il s'agit de tests dits à éléments isolés («discrete point test»), qui évaluent séparément des aspects de la composante communicative (Fahardy, 1983 ; Oller, 1983 ; Palmer, 1983). Les tests qui mesurent des éléments isolés peuvent servir à diagnostiquer des difficultés précises en vue de les corriger : prononciation d'un phonème, désinence des verbes, contraction des articles sont des exemples d'éléments isolés qu'un test peut vérifier. À l'exception des éléments reliés à la phonétique, il n'est pas nécessaire que l'évaluation des éléments isolés se fasse à l'oral. Elle peut se faire sur papier. Il est généralement entendu que la portée du test se limite à ces éléments isolés et qu'elle ne peut s'étendre à l'évaluation de la production orale (Fulcher, 2000 ; Bachman, 1990 ; Canale et Swain, 1980 ; Bachman et Palmer, 1996)<sup>8</sup>. Les résultats obtenus lors de tests à éléments isolés ne rendent pas compte de la compétence communicative à l'oral, mais plutôt de la maîtrise des éléments du discours qui ont été testés. Selon Fahardy (1983), ces tests mesurent la plupart du temps des aspects reliés à la précision. Leur validité pour mesurer la compétence communicative, plus précisément la production orale, est donc faible. En effet, nous avons vu que la compétence communicative englobe plusieurs composantes et que ces composantes interagissent entre elles. Un test fait de questions qui portent sur des éléments isolés ne pourrait pas la mesurer.

Déjà au début du siècle, des pédagogues<sup>9</sup> s'interrogeaient sur la validité des tests «classiques». Dès les années 50<sup>10</sup>, et avec encore plus de vigueur dans les années 80, avec la

---

<sup>8</sup> Toutefois, un chercheur, Hossein Fahardy (1983) démontre qu'il existe des «évidences statistiques» («statistical evidences») selon lesquelles les résultats qu'obtiennent les étudiants à ces deux types de tests, isolés et intégrés, soient fortement corrélés, si bien qu'il discute la pertinence de les distinguer radicalement sur le plan de la validité.

<sup>9</sup> Fulcher (2000), dans un article qui relate l'influence du courant «communicatif» en évaluation des langues mentionne trois pédagogues : Mercier (1933), Wood, 1927 et Hayden (1920) qui remettaient sérieusement en question la validité des tests conventionnels pour mesurer la production orale.

diffusion de l'approche communicative et de ses implications pédagogiques, les chercheurs se sont appliqués à développer de nouveaux outils permettant d'évaluer avec précision la compétence communicative (Fulcher, 2000). Ce besoin de créer de nouvelles évaluations est venu du constat qu'un étudiant pouvait exceller dans des examens portant sur ses connaissances sur la langue, ou dans la traduction sur papier, mais s'avérait incapable de s'exprimer oralement dans la vie quotidienne. Puisque l'expression orale est une compétence, qui se manifeste à travers une performance, par souci de validité, il est dorénavant admis que les tests de production orale doivent placer les candidats dans une situation où cette compétence se manifeste. Parmi les tests intégrés qui servent à évaluer la compétence communicative, nous en retiendrons deux, soit l'entrevue individuelle et le jeu de rôle.

Aujourd'hui, la façon la plus courante d'évaluer la compétence communicative est l'entrevue orale individuelle ou OPI (Oral Proficiency Interview) (Bolton, 1987; Fulcher, 2000; Lussier et Turner, 1995 ; McNamara, 1996; Muller, 2002; Orr, 2002; Springer, 2002 ; Walter, 2004). L'entrevue consiste à interroger la personne évaluée de façon à recueillir un échantillon linguistique suffisant, à partir duquel l'évaluateur jugera de la compétence communicative de la personne en question. Pendant l'entrevue, ce peut être une tierce personne qui joue le rôle de l'évaluateur ou l'intervieweur lui-même. Elle se déroule normalement en quatre temps. D'abord, il y a une prise de contact, ou une introduction. Cette introduction n'est pas censée faire partie de l'évaluation, elle vise seulement à rendre le candidat à l'aise. La deuxième partie est celle de la vérification du niveau plancher (le minimum qu'un candidat peut faire ) et du niveau plafond (le maximum qu'il peut atteindre). Les façons de les déterminer sont dictées par le protocole d'entrevue que l'intervieweur est supposé connaître. Vient ensuite la période de détente. Dans certains organismes, comme à la Fonction publique canadienne et au ministère de l'Immigration et des Communautés culturelles (MICC), cette période peut être précédée d'un jeu de rôle qui fait donc partie de la vérification du niveau plafond (de Fonteney, 1991).

---

<sup>10</sup> Avec la Deuxième guerre mondiale, le besoin de communiquer efficacement en langue étrangère est devenu un enjeu stratégique et l'enseignement et surtout l'évaluation de la compétence communicative a fait l'objet d'intenses réformes en vue d'accroître leur efficacité dans l'armée américaine (Fulcher, 2000).

Si l'entrevue est utile pour récolter un échantillon du langage, sa validité a quand même ses limites. L'échantillon recueilli contiendra une grande variété de réponses, difficiles à prévoir, puisqu'il est dans leur nature même d'être imprévisibles, spontanées, créatives. Comment s'assurer alors que l'échantillon sera significatif et que ce qu'il contiendra reflètera les capacités réelles de la personne évaluée ? C'est ici que le rôle de l'intervieweur entre en jeu. Des chercheurs ont démontré que l'attitude de l'intervieweur a une grande influence sur la performance de la personne interviewée (de Fontenay, 1991 ; Brown, 2003). Il a été démontré que la même personne interrogée par deux intervieweurs différents pouvait obtenir des notes fort différentes, voire des jugements contradictoires de la part des évaluateurs (Brown, 2003). Une façon de remédier à ce problème est la formation des intervieweurs (de Fontenay, 1991). Outre le rôle de l'intervieweur, d'autres aspects peuvent affecter la validité du test. Des chercheurs se sont intéressés à l'influence que peut avoir le sujet de discussion abordé dans l'entrevue sur le résultat (Jennings, Fox et Shohammy, 1999). Ils se sont demandé si les résultats du test pouvaient dépendre du sujet sur lequel les participants ont à s'exprimer. Il semblerait que, bien que le fait de donner le choix du sujet aux participants semble diminuer leur anxiété, le choix du sujet abordé n'a pas d'influence significative sur les résultats obtenus. Les études ne sont pas concluantes parce que beaucoup de variables entrent en ligne de compte.

Outre l'entrevue individuelle, le jeu de rôle constitue une autre forme de test intégré souvent utilisé en évaluation de la production orale (Fulcher, 2003). Le jeu de rôles consiste à créer une situation, souvent calquée sur une scène de la vie courante (par exemple : prendre un rendez-vous avec un professionnel, commander un plat au restaurant ou encore demander des informations sur un produit) et de demander à la personne évaluée de se comporter comme elle le ferait dans cette situation. Plusieurs auteurs (Bolton, 1987; Brumfit, 1980; Holec, 1979) en ont cependant critiqué l'emploi. Ils lui reprochent son manque d'authenticité. En effet, malgré tous les efforts déployés pour créer une situation réaliste, le jeu de rôle reste un «jeu» et selon ses détracteurs, il ferait appel aux qualités de comédien plutôt qu'à celles de communicateur puisque lors d'un jeu de rôle, la personne évaluée doit s'imaginer dans une situation hypothétique et discuter comme si elle vivait réellement cette situation. Certaines

personnes auraient du mal à le faire. En outre, le jeu de rôle favoriserait les personnalités extraverties au détriment des plus timides (Bolton, 1987).

Par ailleurs, selon plusieurs chercheurs (Fulcher, 2003 ; Milanovic, Saville, Pollitt et Cook (1996)) l'évaluation basée sur l'interaction entre pairs lors d'un jeu de rôle constituerait la meilleure façon de recueillir un échantillon significatif de la compétence communicative, et ce pour quatre raisons. D'abord, en contexte scolaire, ce genre d'évaluation serait plus cohérent avec le type de pédagogie généralement adopté en classe et réduirait ainsi l'anxiété causée par le fait d'être évalué, puisque les étudiants seraient habitués à produire ce genre de discours. Ensuite, la position égalitaire des participants leur permettrait de réaliser un échange qui ressemblerait plus à une interaction authentique. De plus, lors de ce type de test, comme ce sont toujours les personnes évaluées qui ont la parole (et non l'évaluateur ou l'intervieweur), l'échantillon recueilli est donc plus vaste. Finalement, le nombre de composantes mobilisées par ce type de test serait plus grand que pour tout autre type de test. Toutefois, de la même façon que l'intervieweur peut malgré lui affecter la validité de l'entrevue, le pair dans le jeu de rôle (son appartenance culturelle, son sexe, sa personnalité, sa compétence communicative par exemple) pourrait avoir une influence sur la performance, mais les études à ce sujet ne réussissent pas à le montrer de façon significative (O'Sullivan, 2002).

À la lumière de ce qui précède, il apparaît que nous devons surmonter plusieurs difficultés pour créer un test de production orale valide, c'est-à-dire qui permette vraiment de tester ce qu'on veut tester. Une fois le problème de la validité surmonté, il en demeure un autre, qui est de taille : celui de la fidélité du test.

### **2.5.2 La fidélité dans l'évaluation de la production orale**

Nous avons vu que la fidélité du test repose sur deux fondements : d'abord sur le fait que les résultats du test ne dépendent pas d'autres facteurs que ceux qu'il veut mesurer et ensuite, sur sa constance à mesurer ces mêmes facteurs. La fidélité peut donc être abordée sous l'angle de ce qui est évalué et sous l'angle de l'évaluateur. Puisque le premier angle est

surtout du ressort de l'évalué et que nous avons somme toute peu d'emprise sur ce dernier, nous examinerons le concept de fidélité sous le deuxième angle, celui de l'évaluateur.

En ce qui concerne l'évaluation de la production orale, en général, il existe un paradoxe qui veut que plus la validité du test augmente, plus sa fidélité inter et intra évaluateur diminue (Bachman, 1990; Fulcher, 2000; Germain et Netten, 2002; Kaplan, 1991; Kormos et Dénes ; Springer, 2002). En effet, nous avons vu plus haut qu'un test valide tend à recréer une situation de communication la plus authentique possible afin de recueillir un échantillon significatif de la compétence communicative du locuteur. Or, la nature même de l'échantillon le rend difficile à évaluer de façon objective et constante. L'acuité des évaluateurs et l'importance qu'ils accordent aux diverses composantes de la compétence communicative étant variables, les résultats des tests de production orale sont souvent fort différents. Comment remédier à ce problème? Une solution, bien qu'elle ne soit pas une panacée, a été la création d'échelles descriptives de niveaux de compétence.

### **2.5.3 Les grilles d'évaluation**

Il existe plusieurs types de grilles d'évaluation, mais nous allons nous restreindre à en présenter deux, soit celles qui sont utilisées le plus couramment pour évaluer la compétence communicative: holistique et analytique (Fulcher, 2003 ; Bachman, 1990). Nous parlerons de ces deux types de grilles d'évaluation dans la partie suivante.

Les grilles holistiques ont été créées pour répondre à un besoin d'uniformité dans la mesure des compétences (de Fonteney, 1991 ; Kaplan, 1991). Elles reflètent la tentative d'augmenter la fidélité de l'évaluation (Fulcher, 2000; Kaplan, 1991; Gronlund, 1985). Elles sont souvent le fruit d'un travail consensuel d'envergure nationale et internationale. Les premières grilles d'évaluation ont été créées par l'American Council for the Teaching of Foreign Languages (ACTFL). Ces grilles servent à décrire de façon claire le niveau linguistique de la personne évaluée. La grille de l'ACTFL est largement utilisée aux Etats-Unis et ailleurs dans le monde et elle a inspiré des chercheurs dans la création de leurs propres échelles (Fulcher, 2000; Germain et Netten, 2002 ; Kaplan, 1991). Récemment, une

grille commune a été adoptée en Europe par le Cadre européen commun de référence (2005). Au Canada, le MICC a adopté en 2004 une grille de mesure de l'interaction orale, de la production écrite et de la compréhension écrite qui est largement inspirée des travaux du Canadian Language Benchmarks. Comment se présentent ces grilles ? Elles proposent des niveaux de compétence, de débutant à avancé. Chaque échelon décrit avec plus ou moins de précision le degré de complexité de la langue et le type de discours que le locuteur est capable de produire. Plus ils sont complexes, plus l'échelon est élevé.

Ces grilles holistiques sont cohérentes avec une approche éduométrique de l'évaluation, c'est-à-dire que les individus y sont évalués selon la réussite d'une tâche fixée par des critères, par opposition à une approche psychométrique où ils sont évalués en comparaison avec d'autres individus (Carroll, 1983; Cziko, 1983; Lussier et Turner, 1995). Les critères des grilles descriptives en langue seconde est fixée par la performance du locuteur natif, critères qui sont par ailleurs remise en question par plusieurs chercheurs à cause de leur caractère abstrait ou éloigné de la réalité d'un locuteur de langue seconde (Alptekin, 2002 ; Hamilton, Lopes, McNamara et Sheridan, 1993 ; Springer, 2002). Dans les grilles holistiques, les diverses composantes de la compétence communicative sont présentées à travers des descriptions de comportements langagiers (North, 1995). Donc, les composantes ou critères choisis ne peuvent être évalués séparément. Ceci n'est pas sans causer des problèmes aux évaluateurs. Ces problèmes sont mis en évidence par Tom Lumley (2002), dans une étude qui a permis d'observer la façon dont cinq évaluateurs expérimentés sont arrivés à évaluer des productions écrites au moyen d'une échelle de type holistique. Dans cette échelle, chaque échelon, de 1 à 5, correspondait à un degré de maîtrise de la compétence de scripteur de l'étudiant. Chaque descripteur d'échelon incluait quatre critères : la maîtrise de la grammaire, la cohésion et l'organisation, les conventions de présentation et l'accomplissement de la tâche. Le premier échelon correspondait à une maîtrise minimale des quatre critères et le cinquième, à une grande maîtrise des mêmes critères. Le principal problème des évaluateurs était de se trouver en face d'un texte qui ne remplissait pas un des critères (par exemple, celui de la maîtrise de la grammaire), mais qui, par ailleurs, remplissait très bien un autre ou d'autres critères (celui de la cohérence, par exemple). La grille ne permettant pas de noter séparément les critères, les évaluateurs devaient déterminer lequel

des critères l'emportait sur les autres au final, ce qui ne se faisait pas sans un intense questionnement. L'originalité de la recherche de Lumley (2002) était de recueillir les commentaires des évaluateurs, à mesure que ceux-ci verbalisaient leur processus évaluatif. Lumley a analysé ces commentaires et il a découvert que la plupart du temps, les évaluateurs s'en remettaient dans la mesure du possible aux descriptions de l'échelle pour justifier leur choix, mais que, souvent, aux prises avec des productions difficiles à faire correspondre à un échelon, ils avaient tendance à se référer aux productions antérieures afin de justifier leur décision. Nous voyons donc que la nature édumétrique des grilles holistiques (comparaison avec des critères plutôt qu'avec d'autres individus) était difficile à respecter. De plus, l'étude démontrerait que cette grille était incapable de refléter la diversité des productions. Malgré ses limites, dans sa conclusion, Lumley (2002) plaide en faveur de la grille holistique parce qu'elle permet d'atteindre, selon les résultats du chercheur, une assez grande constance inter-évaluateurs. Cette constance est somme toute assez surprenante selon lui, parce que l'analyse des commentaires des évaluateurs démontrait une grande diversité dans le degré de sévérité et dans l'importance accordée à chacun des critères.

D'autres chercheurs (Venema, 2002 ; Bachman, 1990) sont partisans d'une grille analytique, plus apte à rendre compte des point forts et des points faibles de l'étudiant. Les grilles analytiques évaluent séparément différents critères (prononciation, syntaxe, débit, cohérence, etc.) à l'aide d'une échelle descriptive pour chacun de ces critères. En effet, puisque, comme le fait remarquer Kaplan (1991), deux participants peuvent se retrouver sur le même échelon sans toutefois se ressembler, la grille holistique n'est pas d'un grand secours pour décrire précisément la compétence communicative des individus. Contrairement aux grilles holistiques, les grilles analytiques distinguent les diverses composantes de la compétence communicative et les évaluent séparément. Ces grilles seraient plus faciles à utiliser par les évaluateurs et plus faciles à interpréter par les principaux intéressés : les personnes évaluées. Ces dernières seraient mieux informées de leurs forces et de leurs lacunes. Toutefois, les grilles analytiques ne sont pas infallibles. L'effet halo (Gronlund, 1985) est souvent observé lors de leur utilisation. Cet effet se caractérise par la propension de l'évaluateur à donner le même pointage à toutes les catégories. De plus, un autre problème

soulevé lors de l'utilisation de ces grilles est l'erreur de logique (Gronlund, 1985) qui consiste pour les évaluateurs à interpréter différemment les critères.

Nous voyons donc que les grilles, qu'elles soient holistiques ou analytiques, ne sont pas garantes de la fidélité inter-évaluateurs. Les études qui le démontrent abondent (Brown, 1995 ; Lumley et McNamara, 1995 ; Mullen, 1978; Wigglesworth, 1993 ; Bachman, Lynch et Mason, 1995 ; Chalhoub-Deville, 1995). Plus grave encore, une étude de Stanfield et Mann Kenyon (1996) démontre que des pédagogues, à qui ils avaient demandé de classer selon leur degré de difficulté (intermédiaire, avancé et supérieur) 38 tâches communicatives répertoriées sur l'échelle de l'ACTFL (par exemple : «Décrire un objet complexe en détails», «Se présenter» ou encore «Persuader quelqu'un de changer d'habitude»), ne s'entendaient pas sur le degré de difficulté de ces tâches. Difficile alors d'avoir une idée précise des capacités d'une personne selon son rang sur l'échelle de l'ACTFL. Certains chercheurs arguent qu'une façon d'assurer une certaine constance entre les évaluateurs est de leur offrir une formation concernant la façon d'interpréter les critères. Ils ont constaté une plus grande uniformité dans les jugements des évaluateurs après une formation (Lumley, 2002 ; Brown, 1995 ; Wigglesworth, 1993). D'autres études révèlent que ces formations ont une certaine efficacité à court terme, mais qu'à long terme, elles sont peu utiles (Lumley et McNamara, 1995). Ces différences entre les évaluateurs dépendent de plusieurs facteurs. Selon Chalhoub-Deville (1995), cela peut dépendre de leur formation académique, du milieu d'où ils viennent ou de leur expérience. Selon Orr (2002), la différence dans la notation des étudiants résiderait dans les différentes perceptions qu'en ont les évaluateurs, dans l'importance qu'ils accordent aux critères d'évaluation et dans leur degré de sévérité. L'expérience personnelle serait un facteur décisif selon Brown (1995). Ce sont autant de raisons qui provoquent une baisse de la fidélité inter-évaluateur dans l'évaluation de la production orale. Une étude de McNamara (1996) révèle que parmi toutes les composantes de la compétence communicative et ce, peu importe l'échelle utilisée, les enseignants sont toujours plus sensibles au critère de la précision qu'à tout autre critère. Ceci est confirmé d'ailleurs par Chaloub-Deville (1995)

Toutefois, dans l'abondance d'études qui mettent l'accent sur les limites des grilles, d'autres, comme celle de Hinofotis (1983) proposent des pistes de recherches intéressantes pour arriver à bâtir des grilles aptes à rendre compte de la compétence des locuteurs et dont

l'interprétation par les évaluateurs serait plus naturelle. Hinofotis utilise l'analyse factorielle pour montrer l'interdépendance de certains aspects lors de l'évaluation de la production orale. Dans son étude, elle a filmé les exposés oraux de dix étudiants d'anglais, langue seconde, et a demandé à six évaluateurs professionnels et à dix anglophones non professionnels de les évaluer selon neuf critères. À l'analyse, elle voit qu'il existe une certaine cohérence dans l'évaluation faite par les deux groupes d'évaluateurs. Certains aspects apparaissent comme étant corrélés : 1) le vocabulaire et 2) la grammaire sont souvent notés de la même façon ; 3) le développement et l'explication, 4) la clarté de l'expression, 5) l'exemplification et 6) l'habileté à relier le contenu de l'exposé aux valeurs des étudiants reçoivent des marques similaires ; 7) la confiance dans les manières et 8) la présence vont de pair et finalement, 9) la prononciation fait cavalier seul. Ces résultats amènent l'auteure à conclure à l'existence de cinq facteurs ou critères d'évaluation : le premier, le facteur «Communication et information», le deuxième, le facteur «Livraison du message», le troisième facteur : «Les aspects non verbaux», le quatrième, la «Maîtrise linguistique» et le cinquième, le facteur «prononciation». Le facteur «Prononciation», selon ses résultats, serait responsable à lui seul de la détermination d'un niveau seuil de compétence communicative à l'oral. De plus, l'étude démontre l'importance d'un facteur non linguistique, le facteur 3 (Aspects non verbaux) dans l'évaluation.

Finalement, il est clair que les questions concernant la validité et la fidélité des tests de production orale préoccupent toujours les chercheurs. La difficulté de créer des tests valides et fidèles est grande, puisque «il est plus facile d'évaluer les connaissances d'un élève que ses compétences, car pour saisir ces dernières, on doit l'observer aux prises avec des tâches complexes, ce qui prend du temps et ouvre la voie à une contestation.» (Perrenoud, 1997, p.21). Nous verrons dans les paragraphes suivants comment, à la Commission scolaire de Montréal (CSDM), qui relève de la Direction générale de la formation des adultes (DGFA) du Ministère de l'Éducation, du Loisir et du Sport (MELS) les enseignants s'y prennent pour évaluer de la façon la plus valide et fidèle possible.

## 2.6 L'évaluation de la production orale à la CSDM

Dans la section suivante, nous verrons comment se déroule l'évaluation de la compétence communicative orale des étudiants à la CSDM. Pour ce faire, nous donnerons l'exemple d'une situation d'évaluation et de la grille utilisée pour sanctionner les acquis en production orale au niveau 5, soit un niveau avancé. C'est le niveau auquel nous nous intéressons dans notre recherche.

Selon les directives de la DGFA (1994), conformément à l'orientation communicative adoptée par le MELS, pour évaluer la compétence communicative à l'oral des étudiants à la fin de session, les enseignants doivent se servir de deux jeux de rôle d'une durée approximative de cinq minutes chacun. Ces jeux de rôle peuvent se passer entre l'enseignant (qui est en même temps l'évaluateur) et l'étudiant ou entre deux étudiants. L'enseignant s'en tient dans ce cas à un rôle d'évaluateur, quoiqu'il puisse intervenir si la situation plafonne (DGFA, 1994). Fulcher (2003) propose un cadre pour décrire la tâche évaluative (son cadre s'inspire de cadres précédents de Bachman et Palmer (1996)). Afin de présenter toutes les facettes du test de la DGFA (1994), nous nous servons du cadre de Fulcher. Ce cadre englobe six aspects : l'orientation de la tâche (ouverte, guidée ou fermée) ; l'interaction (non interactif, interactif à sens unique, bidirectionnel ou multi-directionnel) ; le but (sans but, but convergent ou divergent) ; le statut des interlocuteurs et leur degré de familiarité (sans interlocuteur, statut supérieur, statut inférieur, même statut ; degré de familiarité avec l'interlocuteur) ; le ou les sujets ; la ou les situations.

D'abord, le sujet des jeux de rôle n'est pas totalement ouvert, puisqu'il est guidé, par un scénario auquel les étudiants doivent se plier. Parfois, le scénario prévoit un choix, mais ce choix est restreint. Par exemple, au niveau 5, une des tâches consiste à raconter un match sportif auquel les étudiants auraient assisté. Leur choix se limite à décider quel sport ils sont allés voir. Donc, la tâche n'est pas complètement ouverte. La tâche est interactive et bidirectionnelle : elle implique deux interlocuteurs. Le but des participants est convergent : ils doivent ensemble arriver à mener à bien une conversation cohérente. Les interlocuteurs ont le même statut (ce sont deux étudiants du même groupe) et généralement, puisqu'ils ont étudié dans la même classe depuis au moins huit semaines (la durée d'une session à la

CSDM), leur degré de familiarité est grand. Le sujet, pour cette situation et pour toutes les autres, est généralement familier et il a obligatoirement fait l'objet d'études pendant la session et la situation est calquée sur une scène de la vie quotidienne.

Avant le jeu de rôle, l'enseignant remet aux étudiants des notes qui expliquent le but du jeu de rôle et qui le mettent en scène. Les étudiants les lisent. Ils peuvent poser des questions à l'enseignant s'ils ne comprennent pas bien la tâche. Toutefois, ils ne peuvent consulter de notes personnelles. Les deux étudiants doivent ensuite se mettre d'accord sur les grandes lignes de la mise en situation. Ils ont un maximum de cinq minutes pour se préparer et ce, pour chaque jeu de rôle. Quand ils sont prêts, l'enseignant démarre le chronomètre en leur donnant le signal du départ.

Pendant le jeu de rôle, l'enseignant écoute attentivement et peut intervenir pour relancer ou réorienter la conversation si les participants plafonnent ou si l'échantillon recueilli est jugé insuffisant. La DGFA ne recommande pas à l'enseignant de prendre des notes, car cela pourrait affecter la performance des étudiants. Elle suggère plutôt d'enregistrer le tout sur cassette audio, pour écoute ultérieure si l'enseignant en ressentait le besoin.

Une fois la performance terminée, l'enseignant prend quelques minutes pour consulter la grille d'évaluation (appendice A) et accorder à chaque étudiant le niveau qui lui semble le plus approprié. Donc, il troque son rôle d'enseignant pour un rôle d'évaluateur. Or, selon Figueras (2005), il n'est pas nécessairement préparé pour le faire. Enseigner et évaluer sont deux activités distinctes et le manque de formation des enseignants serait une source d'erreurs dans l'évaluation. Une étude de Glisan et Foltz (1998) montre que les enseignants, même s'ils connaissent leurs étudiants, ne sont pas toujours les mieux placés pour donner un jugement objectif au sujet de leur performance à l'oral notamment. Les auteurs montrent que lorsque l'on compare la note en production orale accordée par les enseignants à celles d'évaluateurs formés et expérimentés, dans 53% des cas, les enseignants avaient surévalué la compétence de leurs étudiants. Pourtant, en contexte scolaire, c'est presque toujours aux enseignants que revient la tâche d'évaluer les étudiants. À ce sujet, Venema (2002) parle de la nécessité d'étudier les réflexions que les enseignants font au moment où ils évaluent leurs étudiants, afin de mieux comprendre la façon dont ils interprètent les critères d'évaluation,

parce qu'au final, la grille n'est pas d'un grand secours pour connaître les raisons qui motivent la décision.

Cette grille d'évaluation utilisée à la DGFA (1994) est holistique. Elle repose sur cinq critères principaux : l'organisation et la clarté du message, le respect de la consigne, la justesse des structures et la richesse du vocabulaire et le degré d'appropriation des éléments linguistiques. Ces critères ne sont pas évalués séparément, mais tout d'un bloc. Notons que l'étudiant, qu'il soit débutant, intermédiaire ou avancé, est toujours évalué avec la même échelle. La grille n'explique pas la différence entre un étudiant débutant, intermédiaire ou avancé. Elle donne des indications très générales sur la réussite de l'épreuve. Chaque échelon correspond à un degré de performance, qui est traduit ensuite en un pointage. Ce pointage est ensuite additionné aux pointages des trois autres examens (production écrite, compréhension orale et compréhension écrite) et le total est consigné sur le bulletin de l'étudiant.

## **2.7 Synthèse du chapitre**

En résumé, la compétence communicative à l'oral est la manifestation d'une habileté langagière complexe. C'est une compétence qui peut être séparée en plusieurs composantes, mais qui est difficile à modéliser. Cette compétence est visible lors d'une performance où elle est mobilisée. Elle est aussi difficile à reproduire artificiellement, en contexte d'évaluation. Et même si le test élaboré pour la mesurer se révèle valide, la production orale reste difficile à évaluer. De nombreuses études font état de la difficulté d'arriver à la constance inter-évaluateurs. Ce problème entourant l'évaluation de la production orale a fait surgir les questions que nous présenterons dans la section suivante.

## **2.8 Questions de recherche**

Comment les enseignants à la CSDM arrivent-ils à évaluer la production orale de leurs étudiants? Cette interrogation est le point de départ de notre recherche exploratoire qui a pour but de découvrir les aspects de la compétence de communication recherchés par les

évaluateurs lors de l'évaluation. Quels sont ces critères et diffèrent-ils d'un évaluateur à un autre? Voilà nos questions, que nous reformulons plus précisément de cette manière : Question 1) Quelles composantes de la compétence communicative sont les plus recherchées par les enseignants lors de l'évaluation de l'oral? Question 2) Y a-t-il un consensus entre les évaluateurs quant aux compétences recherchées? Ces deux questions constituent la base de notre recherche exploratoire. Dans le chapitre suivant, nous exposerons la méthodologie envisagée pour la mener à bien.

## CHAPITRE III

### MÉTHODOLOGIE

Dans la partie suivante, nous présenterons le cadre méthodologique conçu pour répondre à nos questions de recherche. D'abord, nous parlerons de nos participantes, ensuite, nous décrirons les instruments de collecte de données utilisés. Puis, nous parlerons du déroulement de notre expérimentation et enfin, nous expliquerons la façon dont nous avons analysé les données.

#### **3.1 Résumé de l'expérimentation**

Comme nous avons pu le voir, l'évaluation de la production orale soulève des problèmes de définition du construit, de validité et de fidélité de l'évaluation. C'est le dernier problème, la fidélité de l'évaluation, qui nous intéresse. Plus particulièrement, nous voulons savoir ce qui retient l'attention des enseignants au moment où ils évaluent la compétence à l'orale de leurs étudiants. Afin de découvrir les aspects de la compétence communicative qui sont les plus importants pour les enseignants, nous avons tenté de recréer, lors d'une expérimentation, une situation d'évaluation la plus authentique possible, c'est-à-dire qui ressemblait à ce qui se passe normalement en fin de session, lors des examens sommatifs. Pour ce faire, nous avons choisi d'observer les pratiques en cours dans un centre d'éducation

des adultes du réseau des Commissions scolaires. Nous avons choisi de faire notre expérimentation dans ce réseau (plus particulièrement dans un centre faisant partie de la Commission scolaire de Montréal) principalement parce que les cours de francisation y sont en nette progression depuis 2000<sup>11</sup>. Nous pensons qu'il est pertinent de s'intéresser aux pratiques évaluatives dans un milieu où la demande de cours de francisation est croissante. Une autre raison qui a guidé notre choix est que les enseignants de la Commission scolaire à qui nous avons parlé de notre projet se sont montrés très réceptifs. Il a donc été facile de sélectionner des enseignantes intéressées à participer à notre expérimentation. Nous les présenterons dans la partie suivante.

### **3.2 Présentation des participantes**

Nous avons fait appel à cinq enseignantes pour mener à bien notre recherche exploratoire. Le grand critère retenu pour les sélectionner était l'expérience en enseignement et en évaluation à la formation des adultes dans le réseau des commissions scolaires. Nous désirions interroger des enseignantes qui connaissaient le programme, la clientèle et les pratiques évaluatives de la DGFA. Pour les présenter, nous utiliserons un pseudonyme, afin de respecter la confidentialité de leurs commentaires.

#### **3.2.1 Julie**

La première enseignante, Julie, est dans la vingtaine. Sa langue maternelle est le français. Elle parle aussi anglais et espagnol. À la maison, elle utilise le français et l'espagnol puisque son conjoint est hispanophone. Julie a commencé à enseigner en 2003. Elle cumule trois ans d'expérience en enseignement du français comme langue seconde aux adultes. Elle a donné tous les niveaux d'enseignement, c'est-à-dire de débutant (niveau 1) à avancé (niveau

---

<sup>11</sup> En 2000-2001, le nombre d'inscription aux cours de francisation dans l'ensemble du réseau des Commissions scolaires était de 9629. Ce nombre a grimpé à 17 716 en 2004-2005, dont 13 408 dans la région de Montréal (Ministère de l'Éducation du Québec, Direction de la recherche, des statistiques et des indicateurs, 2007).

6) au moins une fois. Il s'agit d'une jeune enseignante de français langue seconde qui est en début de carrière.

### **3.2.2 Éline**

La deuxième participante, Éline, est dans la cinquantaine. Sa langue maternelle est le français et elle parle aussi anglais. À la maison, elle utilise le français. Cette enseignante cumule au total 20 années d'expérience comme professeur de français langue seconde. Elle a donné une quinzaine de fois tous les niveaux, de 1 à 6. Elle a aussi enseigné le français, langue maternelle, pendant trois ans, en Louisiane. Elle a enseigné pendant quinze étés le français langue seconde à l'université et le reste du temps, à la CSDM. Elle a aussi fait de la suppléance au Ministère de l'Immigration et des Communautés culturelles. Il s'agit donc d'une enseignante très expérimentée, qui a travaillé dans plusieurs milieux.

### **3.2.3 Phuong**

La troisième participante, Phuong, est dans la cinquantaine. Sa langue maternelle est le vietnamien. Elle parle aussi anglais et allemand. À la maison, elle utilise le français. Cette enseignante a 19 ans d'expérience en enseignement aux adultes, dont seize en français langue seconde et trois en français, langue maternelle. Elle a donné tous les niveaux un nombre à peu près égal de fois. Elle a accumulé la plus grande partie de son expérience à la CSDM, mais a aussi enseigné dans un organisme communautaire, au MICC, au cégep et dans une école privée, et ce, entre quelques mois à trois ans à chaque fois. Cette enseignante est donc très expérimentée et elle connaît bien tous les niveaux enseignés à la Commission scolaire.

### **3.2.4 Clémence**

La quatrième participante, Clémence, est elle aussi dans la cinquantaine. Sa langue maternelle est le français. Elle parle anglais, mais utilise sa langue maternelle à la maison. Cette enseignante cumule 18 années d'expérience en enseignement aux adultes, dont seize en

français langue seconde et deux en français, langue maternelle. Elle a donné surtout les niveaux débutants, c'est-à-dire le niveau 1 et le niveau 2. Elle a donné des cours au niveau préalable, c'est-à-dire des cours visant à rendre les élèves aptes à se débrouiller au niveau 1. De plus, elle a donné des cours d'alphabétisation à des travailleurs agricoles, au début de sa carrière d'enseignante. Elle n'a donné qu'une seule fois les niveaux 3, 4 et 5 et n'a jamais donné le niveau 6. Elle enseigne à la CSDM, mais a entre zéro et trois années d'expérience au MICC. Cette enseignante a donc beaucoup d'expérience avec les débutants, mais très peu avec les groupes de niveaux intermédiaire et avancé.

### 3.2.5 Marie-Ève

La dernière participante, Marie-Ève est dans la trentaine. Sa langue maternelle est le français. Elle parle anglais et espagnol, puisque, comme Julie, son conjoint est hispanophone. Elle a onze ans d'expérience en enseignement du français, langue seconde. Elle a surtout enseigné aux adultes (huit ans), mais aussi au primaire et au secondaire (trois ans). Elle a donné les cours de niveaux 1 à 5. Elle n'a jamais donné le niveau 6. Marie-Ève démontre une grande polyvalence, puisque son expérience se répartit entre les organismes communautaires (un an), deux commissions scolaires (de Montréal et Marguerite-Bourgeoys), une université (un an) et une école privée (un an). La carrière de cette enseignante est donc bien amorcée : elle connaît bien presque tous les niveaux et a travaillé dans plusieurs types d'institutions.

**Tableau 3.1**  
**Renseignements démographiques sur les participantes**

	Julie	Élaine	Phuong	Clémence	Marie-Ève
Années d'expérience en enseignement	4	20	19	18	11
Années d'expérience en français L2 aux adultes	3	18	14	16	8
Nombre de sessions au niveau 5 à la DGFA	2	14	10	1	3
Langues parlées à la maison	Français, espagnol	Français	Français	Français	Français, espagnol
Autres langues connues	Anglais	Anglais	Anglais, allemand	Anglais	Anglais

### **3.3 Instruments utilisés**

Pour mener à bien notre expérimentation, nous avons utilisé les instruments suivants: 1) un questionnaire sociodémographique ; 2) la grille d'évaluation de la production orale de la DGFA (1994) que les participantes étaient libres d'utiliser ou non; 3) des feuilles de notes; 4) deux vidéos montrant l'examen oral sommatif de quatre étudiants de niveau 5 ; 5) les consignes relatives aux jeux de rôle ; 6) et finalement, un questionnaire d'entrevue semi dirigée. Nous décrirons ces instruments dans la partie suivante.

#### **3.3.1 Questionnaire sociodémographique**

Nous avons élaboré un questionnaire qui nous a permis de sélectionner cinq enseignantes selon leur expérience avec le niveau 5 à la DGFA (voir appendice B). La première page de notre questionnaire visait à recueillir les noms et coordonnées des participantes ; la deuxième, des informations relatives au sexe, à l'âge et aux langues parlées ; la troisième portait sur le nombre d'années d'expérience en enseignement et le type d'enseignement prodigué et la clientèle visée et finalement, la quatrième page portait sur les lieux d'enseignement des enseignantes.

#### **3.3.2 Vidéos**

Nous avons filmé l'examen de production orale de quatre étudiants en cours de formation dans un centre d'éducation des adultes de la Commission scolaire de Montréal, afin de les faire évaluer par les enseignantes. Notre recherche ne portant pas sur la performance des étudiants, mais plutôt sur les commentaires des évaluateurs, les étudiants sont donc considérés comme des instruments plutôt que comme des participants. Le contenu des vidéos constituaient notre corpus.

Nous avons choisi des étudiants de niveau cinq parce qu'à la fin de ce niveau, les étudiants sont jugés fonctionnels en français (DGFA, 1994). Plusieurs d'entre eux quittent les cours de français après la réussite du cinquième cours, souvent pour aller sur le marché du

travail. Nous croyions intéressant de connaître l'avis d'évaluateurs sur la compétence communicative des étudiants qui sont en fin de programme.

Nous avons sélectionné deux hommes et deux femmes, faisant partie de la même classe, pour que leur degré de familiarité soit élevé, comme lors d'une situation d'évaluation authentique. Les quatre étudiants étaient dans la trentaine. Nous avons attribué un numéro à chacun des étudiants afin de simplifier la présentation des données. L'étudiant 1, un homme d'origine mexicaine, était au Québec depuis deux ans. Sa langue maternelle était l'espagnol. Lors de l'évaluation nous l'avons placé en équipe avec l'étudiante 2, une Moldave arrivée au Québec il y a un an et demi. Sa langue maternelle était le roumain. L'étudiant 3 était mexicain, tout comme l'étudiant 1. Il vivait au Québec depuis plus de 10 ans. Sa langue maternelle était l'espagnol. L'étudiante 4, qui a été placée en équipe avec l'étudiant 3, est née au Canada. Sa langue maternelle était l'anglais. Elle vivait au Québec depuis un peu moins d'un an. Les étudiants se connaissaient depuis au moins deux mois et ils avaient tous réussi l'examen oral de niveau 5. Ils avaient commencé le niveau 6 depuis deux semaines. Comme nous ne voulions pas que leur participation à notre expérimentation ait une quelconque influence sur le résultat de leur véritable examen sommatif, nous avons attendu qu'ils aient fini le cours et les examens du niveau 5 avant de les filmer. Ils étaient donc familiers avec le type d'examen du niveau 5, puisqu'ils venaient de le passer et l'avaient tous réussi.

L'examen de production orale du niveau 5 consiste, comme nous l'avons expliqué dans notre cadre théorique (chapitre II), en un jeu de rôles séparé en deux parties où deux étudiants qui interagissent sont évalués. Pour notre expérimentation, nous avons choisi le domaine social pour le premier jeu de rôle. La tâche consistait à : «Conclure une entente de service en vue d'obtenir de l'aide à domicile» (DGFA, 1994). Un étudiant qui jouait le rôle d'un préposé au bénéficiaire répondait par téléphone à une offre d'emploi placée par l'autre étudiant, qui cherchait quelqu'un pour prendre soin, à son domicile, de sa mère malade et âgée. Les étudiants devaient discuter des conditions de travail et conclure l'entretien téléphonique par un engagement, un refus ou encore un rendez-vous. Le deuxième jeu de rôle appartenait au domaine social et consistait à : «Discuter avec une amie ou un ami sur un événement sportif» (DGFA, 1994). Ce jeu de rôle mettait en scène deux amis qui venaient

d'assister à un match du sport de leur choix. Ils devaient demander et donner une opinion sur cet événement.

La première dyade était composée de l'étudiant 1 et de l'étudiante 2. Ensemble, les deux étudiants ont convenu que l'étudiant 1 jouerait le rôle de l'employeur et que l'étudiante 2 jouerait celui de la préposée au bénéficiaire. Ils ont convenu que la mère de l'étudiant 1 requerrait des soins à cause de la faiblesse généralisée causée par son vieil âge. Le jeu de rôle a duré quatre minutes 53 secondes.

Lors de la deuxième situation, les étudiants de la première dyade (l'étudiant 1 et l'étudiante 2) ont choisi de parler d'un match de hockey junior auquel l'étudiant 1 avait assisté. L'étudiante 2 devait lui poser des questions au sujet du match et lui demander son appréciation. Notons que les étudiants ont dérogé à la consigne, puisque selon le scénario proposé, ils avaient assisté ensemble au match. Nous ne sommes pas intervenue pour rectifier le tir, puisque le respect de la consigne fait partie des critères d'évaluation de la grille du MELS (1994). Nous voulions savoir si cette erreur allait être remarquée plus tard par nos participantes.<sup>12</sup> Le jeu de rôle a duré cinq minutes et trois secondes.

La deuxième dyade, composée de l'étudiant 3 et de l'étudiante 4, a convenu que ce serait l'étudiant 3 qui chercherait quelqu'un pour s'occuper de sa mère malade et que, conséquemment, ce serait l'étudiante 4 qui jouerait le rôle de la préposée aux bénéficiaires. Nous constatons que les rôles ont été attribués de façon similaire dans cette dyade : ce sont les hommes qui ont joué le rôle de l'employeur et les femmes celui de la préposée au bénéficiaire. Ce jeu de rôle a duré quatre minutes 56 secondes.

Pour le deuxième jeu de rôle, les étudiants 3 et 4 ont aussi choisi de parler d'un match de hockey, mais, à la différence de la première dyade et conformément aux directives de la DGFA, ils ont décidé qu'ils y avaient assisté ensemble. Le match imaginaire s'était déroulé au Centre Bell et après le match, les étudiants s'étaient attablés au restaurant La Cage aux sports pour en parler. Les deux étudiants ont convenu avant le début du jeu de rôle du

---

<sup>12</sup> Dans les faits, l'erreur n'a été soulevée par aucune des cinq participantes. Nous en reparlerons plus tard, dans la section Interprétation.

pointage final du match et du nom des équipes (Les Canadiens de Montréal et les Maple Leafs de Toronto) qui s'étaient affrontées. Ce jeu de rôle a duré cinq minutes et seize secondes, mais nous avons dû intervenir à la troisième minute parce que les étudiants plafonnaient. Nous avons relancé la discussion en leur demandant de parler de l'ambiance générale qui régnait au Centre Bell pendant le match, de façon à ce que la discussion prenne un tour différent et puisse se poursuivre pendant deux minutes de plus, pour atteindre les cinq minutes réglementaires de l'examen. L'enregistrement s'est poursuivi pendant notre brève intervention. Notons que la DGFA suggère à l'enseignant d'intervenir quand le jeu de rôle stagne et elle indique qu'une brève intervention ne devrait pas avoir de portée négative sur la note des étudiants.

Les étudiants ont été filmés au moyen d'une caméra VHS. La disponibilité des lecteurs VHS et la facilité de stockage des données ont motivé le choix de ce support. En effet, il a été simple de trouver l'équipement pour faire visionner les vidéos. Nous avons choisi le support vidéo plutôt qu'audio parce qu'il permettait aux enseignantes d'évaluer les aspects non verbaux de compétence communicative : gestes, regards, physionomie. Le support vidéo était le plus apte à présenter une situation d'évaluation authentique.

### **3.3.3 Feuilles de notes**

Pendant le visionnage des jeux de rôle, les enseignantes ont consigné au besoin leurs remarques par écrit sur deux feuilles de notes que nous leur avons remises (appendice C). Chacune des feuilles était séparée en colonnes correspondant à chacun des étudiants.

### **3.3.4 Consignes d'examen**

Les étudiants et les enseignantes ont pu prendre connaissance du contenu des jeux de rôle à l'aide du document conçu par la DGFA (1994). Ce document indique quelle est la tâche des étudiants, l'objectif qu'ils doivent atteindre et la durée du jeu de rôle. Les rôles de chaque étudiant y sont décrits en trois phrases et quelques pistes de discussion sont présentées. Notons que la nature confidentielle du document nous empêche d'en reproduire une copie dans ce mémoire.

### 3.3.5 Grille et pondération

Il importe ici de parler de la pondération de la DGFA (1994). Elle réserve 30% de la note finale à l'examen de production orale du niveau 5, le reste du 70% étant réparti entre les trois autres habiletés (production et compréhension écrite, compréhension orale). Le seuil de réussite de l'épreuve est fixé à 60%. Pour évaluer le jeu de rôle (l'expression orale), la DGFA (1994) suggère une grille d'évaluation qui prévoit une note sur quinze points. L'enseignant doit donc évaluer les deux jeux de rôle sur quinze points, puis additionner les résultats pour arriver à un total sur 30 points. Il rapporte ensuite ce total dans le tableau de notation générale, additionne les notes des quatre épreuves différentes (production orale, production écrite, compréhension orale et compréhension écrite) pour arriver à la note finale qui décidera si l'étudiant est promu au niveau suivant.

Nous avons remis une copie de la grille du MELS aux participantes (voir appendice A). Le contenu de cette grille a été décrit dans le chapitre II, le cadre théorique. Cependant, les participantes étaient libres d'utiliser ou non cette grille. Certaines d'entre elles ont préféré utiliser leur propre grille, ou aucune grille. Dans ce dernier cas, leur pondération a été justifiée par les notes prises en cours de visionnage.

Comme nous voulions observer les pratiques personnelles des enseignantes, nous n'avons imposé ni grille ni pondération. Le résultat de ce choix est que nous avons obtenu suite à l'évaluation des participantes des notes sur 30 et des notes sur 100. Dans un souci d'uniformité, nous avons reporté ces notes sur le même quotient, c'est-à-dire sur 100. Nous avons arrondi à l'unité supérieure les notes dont la première décimale était supérieure ou égale à cinq. Par exemple, une note de 67,5 % a été arrondie à 68%. Nous croyons que ce choix facilite la lecture.

### 3.3.6 Questionnaire d'entrevue semi dirigée

Un questionnaire en deux phases a été conçu afin de recueillir les commentaires des enseignants (voir appendice D). La première partie de l'entrevue visait à découvrir ce que les enseignantes avaient pensé de la performance des étudiants. La deuxième partie de l'entrevue, quant à elle, avait pour but de connaître les critères d'évaluation personnels des participantes. Ce questionnaire était composé de questions ouvertes qui tentaient de soutirer le plus de commentaires possible, sans qu'ils soient biaisés. Notre questionnaire était une adaptation de celui utilisé par Pollit et Murray (1996) lors d'une recherche qui avait pour objectif de fournir un fondement empirique (c'est-à-dire plus proche des perceptions des évaluateurs), plutôt que logique (c'est-à-dire souvent calqué sur un modèle théorique de la compétence communicative difficile à transposer dans la réalité) aux descripteurs des échelles d'évaluation utilisées pour évaluer l'oral des étudiants. Pour y parvenir, les deux chercheurs ont eu recours à des procédés d'interview qui ne dirigeaient pas l'attention des évaluateurs sur des aspects précis de la compétence communicative, comme c'est souvent le cas dans des recherches portant sur l'évaluation de l'oral (Orr, 2002 ; Hinofotis, 1983 ; Venema, 2002). Pollit et Murray (1997) ont demandé aux participants de comparer les étudiants évalués entre eux («The method of paired comparisons», de Thurstone (1959), p. 78) et ensuite, d'expliquer en quoi ils sentaient que les étudiants se ressemblaient et différaient : «Immediately following each comparative judgement, judges were told simply to talk about how they felt the two performance were similar and different.» (p. 81).

Ce type d'entrevue offre, selon les auteurs, l'avantage d'observer les processus cognitifs mis en branle par les évaluateurs dans une activité comme l'évaluation et surtout, de mettre à jour les construits personnels («personal constructs», p.77) des participants sur un sujet en particulier. Comme nous nous intéressons à la perception des évaluateurs, nous avons choisi d'utiliser un questionnaire inspiré de cette méthode. En première partie, nous avons demandé aux enseignantes de noter les étudiants, de justifier leurs notes et ensuite, de comparer les étudiants entre eux. En deuxième partie d'entrevue, nous leur avons demandé de nous dire quels étaient, en général, les critères les plus importants pour elles lors de l'évaluation de l'oral, dans leur pratique régulière. Une fois les critères mis à jour, nous leur avons demandé de les noter sur une échelle de 1 à 5, et de nous dire pourquoi elles avaient choisi cette note. Ensuite, nous leur avons demandé de comparer les critères entre eux, de

façon à ce qu'elles définissent les critères en question, sans que nous orientions leurs réponses. Ce type d'entrevue nous est apparu comme étant le plus apte à répondre à nos questions de recherche.

### **3.4 Déroulement**

Notre recherche s'est déroulée en trois étapes : nous avons filmé les jeux de rôles, ensuite, nous avons conduit les entrevues avec les enseignantes et nous avons ensuite transcrit et codifié les données.

Nous avons d'abord filmé les jeux de rôle de quatre étudiants. Cette étape s'est déroulée sur deux périodes de 45 minutes, une pour chacune des dyades. Elle s'est déroulée pendant la dernière semaine de novembre 2006.

Dès la semaine suivante, début décembre 2006, et ce jusqu'à la mi-décembre, nous avons mené les entrevues avec les enseignantes. Chaque séance a duré environ 75 minutes. Nos entrevues comportaient deux grandes étapes. La première étape consistait à évaluer et commenter le jeu de rôle et la deuxième étape consistait à révéler les critères personnels.

Avant le visionnage des jeux de rôle, nous avons remis aux enseignantes les consignes relatives à l'examen de la DGFA, pour qu'elles connaissent le sujet des deux jeux de rôles et les fonctions langagières évaluées. Pendant le visionnage, nous avons encouragé les enseignantes à prendre des notes. Elles pouvaient aussi, si elles le désiraient, utiliser leur propre grille d'évaluation au lieu de notre feuille de notes pour consigner leurs commentaires, ceci pour que leur évaluation soit le plus près possible de leur pratique habituelle. Après le visionnage de chaque vidéo, nous avons laissé aux enseignantes le temps requis pour accorder la note aux étudiants, en se servant de la grille d'évaluation du Ministère ou de leur grille personnelle. Ce temps a varié selon les besoins exprimés par les enseignantes, entre deux et cinq minutes. Une fois la notation terminée, nous avons interrogé les enseignantes. Nous leur avons demandé d'expliquer les raisons ayant guidé leur jugement et de comparer les étudiants entre eux. Les données quantitatives (sous forme de score attribué aux étudiants) et qualitatives (sous forme de commentaires) nous semblaient, à l'instar de celles recueillies

lors d'expérimentations similaires à la notre (Pollit et Murray, 1997 ; Orr, 2002), aptes à apporter des réponses à nos questions. Ensuite, nous sommes passés à la deuxième étape de l'entrevue, qui visait à faire définir par les participantes les critères personnels invoqués lors de l'évaluation. L'analyse qualitative des réponses des enseignantes nous a permis de répertorier les critères d'évaluation les plus importants pour elles. Nous parlerons de la façon dont nous avons analysé ces réponses dans la partie suivante.

### 3.5 Traitement des données

Les entrevues (enregistrées sur cassette audio) ont été transcrites dans Word puis codifiées grâce au logiciel NVIVO. Ce logiciel nous a permis d'organiser et plus particulièrement de hiérarchiser les commentaires, ce qui a facilité l'interprétation des données.

Les réponses des participantes, conformément au plan de notre expérimentation, ont été séparées en deux catégories. La première concernait les notes et les remarques émises par les cinq enseignantes au sujet de la performance des quatre étudiants pendant les jeux de rôle et la deuxième concernait les critères personnels des enseignants.

Nous avons classifié les réponses de la première partie de notre entrevue dans une grille d'analyse de la compétence communicative inspirée principalement par le modèle de Fulcher (2003), qui est conçu pour répondre spécifiquement aux besoins de l'évaluation de l'oral (voir notre cadre théorique, chapitre II). Nous avons donc choisi son modèle, auquel nous avons apporté des modifications à la lumière du modèle de l'évaluation de la précision et de l'aisance de Germain et Netten (2004), qui clarifient et élargissent le concept d'aisance, ce qui le rend plus opérationnel. Dans cette catégorie, à l'instar des deux chercheurs, nous incluons la quantité d'information, la complexité des phrases et les hésitations. Par exemple, une remarque comme : «Ça coulait bien», difficile à classer dans le modèle de Fulcher (2003) entre dans la catégorie de l'aisance selon Germain et Netten (2004). Nous avons ajouté deux catégories au modèle de Fulcher, soit *La compétence interactive* qui concerne la façon dont les étudiants interagissaient et *Les critères non linguistiques* qui concernent le langage

corporel et les remarques sur la personnalité de l'évalué. La raison de cet ajout est que des enseignantes ont émis des commentaires que nous n'arrivions pas à classer dans le tableau de Fulcher, mais qui avaient une incidence sur le score de l'étudiant. En plus de codifier les entrevues selon les composantes de la grille d'analyse de la compétence communicative, nous avons dû tenir compte de la portée positive ou négative de chaque commentaire. Le tableau suivant montre les codes utilisés pour répertorier les commentaires de la première partie de l'entrevue :

**Tableau 3.2**  
Classification des codes utilisés

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie		
	Précision		
	Aisance		
Capacité stratégique	Stratégies de réussite		
	Stratégies d'évitement		
Savoir textuel			
Savoir pragmatique			
Savoir sociolinguistique			
Compétence interactive			
Critères non linguistiques			

La deuxième catégorie de commentaires concernait les réflexions des enseignantes au sujet de leurs critères d'évaluation. Cette deuxième étape visait à découvrir quels sont les critères sur lesquels les enseignantes disent s'appuyer lorsqu'elles évaluent leurs propres étudiants. Nous avons analysé l'ensemble des critères relevés par les enseignants.

Nous avons fait vérifier notre classification et notre codification par un étudiant au doctorat spécialisé en philosophie du langage et habilité à travailler avec un logiciel de codification. Nous lui avons demandé de passer nos codes en revue et de commenter la structure de nos tableaux. Ses remarques pertinentes nous ont permis de simplifier notre codification (nous avons éliminé plusieurs codes redondants) et de clarifier nos tableaux.

Enfin, nous avons relu trois fois nos transcriptions pour nous assurer que nous avons codifié toutes les remarques pertinentes pour notre recherche.

Dans la prochaine section, la présentation des résultats, nous décrirons en détail ce que notre recherche nous a permis de découvrir.

## CHAPITRE IV

### PRÉSENTATION DES RÉSULTATS

Le chapitre suivant sera consacré à la présentation des résultats. Nous présenterons d'abord les résultats de la première partie d'entrevue, c'est-à-dire les notes obtenues par chaque étudiant et les commentaires émis par les enseignantes au sujet de leur performance.

Ensuite, nous présenterons les résultats obtenus en deuxième partie d'entrevue, concernant la façon générale dont chaque enseignante aborde l'évaluation dans sa pratique, notamment au sujet de la grille utilisée, sur les aspects qu'elles disent prendre en considération lors de l'évaluation.

#### **4.1 Résultats de la première partie de l'entrevue**

En conformité avec ce que prévoit la DGFA (1994), Julie et Marie-Ève ont accordé une note sur quinze pour chaque jeu de rôle, notes qu'elles ont ensuite additionnées. Phuong, quant à elle, a accordé une seule note sur quinze points, la même pour les deux jeux de rôles et elle l'a ensuite multipliée par deux. Pour sa part, Éline a accordé deux notes sur 25 (une note différente par jeu de rôle) notes qu'elle a ensuite additionnées puis multipliées par deux

pour arriver à un total sur 100. Finalement, Clémence a préféré donner une note globale sur 100. Nous avons rapporté ces résultats sur 100. Nous les voyons résumés dans le tableau 4.1.

**Tableau 4.1**  
**Notes accordées aux quatre étudiants par les cinq évaluatrices**

	Étudiant 1	Étudiante 2	Étudiant 3	Étudiante 4	Moyenne
Julie	73%	73%	80%	67%	73%
Élaine	80%	92%	80%	70%	81%
Phuong	80%	93%	87%	80%	85%
Clémence	68%	72%	70%	70%	70%
Marie-Ève	87%	83%	93%	53%	79%
Moyenne	72%	83%	82%	68%	78%
Écart entre la plus haute et la basse note	19 points	21 points	23 points	27 points	15 points

En regardant le tableau, le premier constat que nous pouvons faire, c'est la grande disparité dans le jugement des cinq évaluatrices au sujet de la performance globale des étudiants. L'écart le plus marqué se trouve dans l'évaluation de l'étudiante 4 : quatre enseignantes lui accorde au moins la note de passage ou même une bonne note, mais une enseignante, Marie-Ève, la fait échouer. Nous constatons que les deux enseignantes ayant le plus d'expérience avec les niveaux avancés, Élaine et Phuong, ont rendu un jugement presque similaire. En effet, elles ont donné la même note à l'étudiant 1, et presque la même note à l'étudiante 2. En ce qui concerne la deuxième paire d'étudiants, l'écart est plus important, mais dans l'ensemble des notes qu'elles ont accordées, un faible écart de 4,5% les sépare. Clémence, l'enseignante qui a le moins d'expérience avec les niveaux avancés (mais qui a 19 ans d'expérience avec les niveaux débutants) est celle qui a rendu les jugements les plus sévères. Julie est, après Clémence, celle qui a donné les jugements les plus sévères. C'est aussi celle qui a le moins d'expérience avec les niveaux avancés et le moins d'expérience en général. Clémence a accordé les notes plus uniformes. Chez les quatre autres enseignantes, plus expérimentées avec le niveau 5, nous constatons une plus grande nuance dans les notes accordées.

Dans la prochaine section, nous présenterons les commentaires émis par les enseignantes sur leur feuille de notes et lors de l’entrevue semi dirigée, au sujet de la performance de chaque étudiant. Nous les avons répertoriés à l’aide d’un tableau bâti selon la définition de la compétence communicative que nous avons retenue.

#### 4.1.1 Évaluation de l’étudiant 1

Comment les cinq enseignantes ont-elles perçu la performance de l’étudiant 1? Nous présenterons leurs commentaires dans la partie suivante. Nous commencerons avec la première enseignante, Julie.

##### 4.1.1.1 Comment Julie a évalué l’étudiant 1

**Tableau 4.2**  
**Commentaires de Julie au sujet de l’étudiant 1**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie		
	Précision	<i>Au niveau du vocabulaire, les deux, c’était très bien.</i>	<i>L’adjectif «vieille» était à la mauvaise place. Il manquait le sujet dans ses phrases. Il manquait de vocabulaire.</i>
	Aisance	<i>Il parlait plus Ça coulait</i>	<i>C’est plus au niveau du rythme, ses phrases étaient très saccadées. Il cherchait dans sa tête les idées.</i>
Capacité stratégique	Stratégies de réussite		<i>Il utilisait souvent «quelque chose». Il ne connaissait le sujet (du premier jeu de rôle).</i>
	Stratégies d’évitement		
Savoir textuel		<i>Il était plus enchaîné.</i>	<i>Les idées ne se suivaient pas. Il n’a pas vraiment répondu à la question de l’autre.</i>
Compétence interactive			<i>Il écoutait moins sa partenaire.</i>
Critères non linguistiques			

Julie a accordé la note de 73% à l'étudiant 1. Elle a évalué séparément les deux jeux de rôle, lui accordant 66% pour le premier et 80% pour le deuxième. À cause de cela, certaines remarques semblent contradictoires : l'étudiant ne démontre pas selon Julie les mêmes forces et les mêmes faiblesses dans les deux jeux de rôles. Julie fait un commentaire en général plus négatif que positif au sujet de la compétence de cet étudiant. Ses remarques concernent chacun des aspects de la compétence communicative, sauf la phonologie. En outre, elle commente ainsi son premier jeu de rôle : *Au niveau de la compréhension, j'avais de la difficulté à le suivre, à le comprendre. [...] Je pense que c'est parce qu'il ne connaissait pas le sujet, il était moins préparé.* Elle mentionne trois fois qu'elle avait de difficulté à le comprendre.

#### 4.1.1.2 Comment Éline a évalué l'étudiant 1

Tableau 4.3  
Commentaires d'Éline au sujet de l'étudiant 1

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie		<i>« Papier » prononcé /papjer/ au lieu de /papje/.</i> <i>« Un » prononcé /un/ au lieu de /œ/</i> <i>« sport » prononcé /espor/ au lieu de /spor/.</i> <i>« expérience » prononcé /esperjās/ au lieu de /experjās/.</i> <i>« peut » prononcé /po/ au lieu de /pø/.</i>
	Précision	<i>Son vocabulaire est très riche</i>	<i>« Femme vieille » au lieu de « vieille femme ».</i> <i>« Besoin de l'aide » au lieu de « besoin d'aide ».</i> <i>« Problèmes motrices » au lieu de « moteurs ».</i> <i>« Votre » au lieu de « vos ».</i> <i>« Chaque deux semaines » au lieu de « aux deux semaines ».</i> <i>« Il a fait » au lieu de elles ont fait.</i> <i>« La bière » au lieu de « de la bière ».</i> <i>« Jamais je n'ai regardé ça » au lieu de je</i>

			<i>n'ai jamais vu ça. «Tu dire» au lieu de «tu dis». Il a de la difficulté à trouver ses mots, il ne connaît pas bien les règles du jeu.</i>
	Aisance	<i>Il a développé beaucoup d'idées, il a dit beaucoup de choses. C'est surtout lui qui parlait.</i>	
Capacité stratégique	Stratégies de réussite		<i>Il va parler d'un sport qu'il ne connaît pas très bien.</i>
	Stratégies d'évitement		
Savoir textuel		<i>Très bien articulé.</i>	
Compétence interactive			
Critères non linguistiques			

Dans le cas de l'étudiant 1, Élane a accordé une note de 88% pour le premier jeu de rôle et de 72% pour le deuxième. Cela lui donne une note globale de 80%. Selon Élane, ce qui fait que l'étudiant 1 obtient une meilleure note pour le premier jeu de rôle que pour le deuxième, c'est qu'il fait moins d'erreurs de syntaxe, de grammaire et de phonétique. Le relevé exhaustif des erreurs produites par l'étudiant est frappant, dans l'évaluation de cette enseignante. Elle justifie sa note par la gravité et le nombre d'erreurs entendues et aussi par ce qu'elle a relevé de positif, surtout en ce qui concerne l'utilisation adéquate des acquis propres au niveau. Élane juge que le message de l'étudiant est plus clair dans le premier jeu de rôle que dans le deuxième. Elle trouve que dans les deux situations, son vocabulaire est excellent. L'enseignante ne fait pas de commentaires au sujet du savoir textuel et de la compétence interactive.

#### 4.1.1.2 Comment Phuong a évalué l'étudiant 1

**Tableau 4.4**  
**Commentaires de Phuong au sujet de l'étudiant 1**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
	Phonologie		

Composante langagière	Précision	<i>Il se rattrape avec des structures simples et claires. Il n'a pas beaucoup de faiblesses.</i>	<i>Il a dit : «un papier» au lieu d'un c.v. «Femme vieille» au lieu de «vielle femme». On sent qu'il voulait montrer qu'il avait de bonnes structures, mais qu'au contraire, c'est là qu'il avait la faille, comme pour les accords des adjectifs, son vocabulaire, les prépositions. Il dit : «les femmes prend» au lieu de «prennent»; «votre papier» au lieu «vos papiers»; «vous avez l'expérience» au lieu de «vous avez de l'expérience»; «dans le matin» au lieu de «le matin»; «soins médicales» au lieu de «médicaux»; «elle ne peut de faire des choses» au lieu de «elle ne peut pas faire de choses»; Il fait des erreurs de conjugaison.</i>
	Aisance	<i>Il a donné beaucoup de détails, il a précisé son idée.</i>	
Capacité stratégique	Stratégies de réussite		
	Stratégies d'évitement		
Savoir textuel		<i>L'ensemble du message s'est bien passé.</i>	
Compétence interactive			<i>La première chose qui m'a choquée, c'est qu'il n'y a pas de contact entre eux. On dirait que chacun s'est composé un rôle. Il n'y a pas de rencontre.</i>
Critères non linguistiques			<i>Lui, il a l'air de celui qui contrôle. Tu sentais quelqu'un qui avait le pouvoir. On sent un côté forcé, on sent qu'il veut en mettre.</i>

Phuong a donné 80% à l'étudiant 1 comme note globale. Contrairement aux deux enseignantes précédentes, elle n'évalue pas les jeux de rôle séparément. D'après Phuong, les mêmes forces et les mêmes faiblesses se sont manifestées dans les deux situations : ce qu'elle avait constaté dans le premier jeu de rôle s'est répété dans le deuxième. Sur papier, l'enseignante a noté quelques erreurs qui relèvent toutes de la précision, mais au moment de

justifier sa note oralement, elle donne une appréciation globale. Elle dit que «*l'ensemble du message s'est bien passé*», que «*il n'a pas beaucoup de faiblesse*» et que «*il a compris les objectifs*». Comme Éline, cette participante trouve qu'il a donné «*beaucoup de détails*». Phuong fait une remarque négative au sujet de la compétence interactive de cet étudiant. Elle dit : «*ce qui m'a un peu choquée, c'est qu'il n'y a pas de contact entre eux*». Finalement, l'enseignante reproche à l'étudiant 1 son attitude, sa tendance à dominer et son côté «*forcé*».

#### 4.1.1.4 Comment Clémence a évalué l'étudiant 1

**Tableau 4.5**  
**Commentaires de Clémence au sujet de l'étudiant 1**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie		<i>Moins bonne phonétique que sa partenaire.</i>
	Précision		<i>Il fait des erreurs qui devraient être corrigées au niveau débutant. «La prochaine semaine» au lieu de «la semaine prochaine».</i>
	Aisance		<i>Il regardait ses notes, il était moins dégagé.</i>
Capacité stratégique	Stratégies de réussite	<i>Il était mieux préparé pour la première situation.</i>	<i>Il aurait pu parler d'une chose qu'il connaît.</i>
	Stratégies d'évitement		
Savoir textuel			<i>Elle lui pose des questions et il répond autre chose.</i>
Compétence interactive			<i>Chacun avait son scénario, mais ça ne répondait pas toujours à la même chose, ça ne concordait pas toujours.</i>
Critères non linguistiques			<i>Je trouve qu'il était moins solide qu'elle.</i>

Clémence a accordé la note globale de 68% à l'étudiant 1. Ce qui retient surtout son attention, c'est le fait qu'il a du mal à suivre la conversation, lors du premier jeu de rôle, mais

elle trouve quand même qu'il était meilleur que lors du deuxième. Clémence pense qu'il était «*mieux préparé*» pour cette situation du premier jeu de rôle. Quant au deuxième, elle dit qu'il a parlé d'un sujet qu'il ne connaissait pas, soit le hockey. La participante dit qu'il «*aurait pu parler d'une chose qu'il connaît*», qu'il «*aurait pu parler du soccer, qui est bien populaire*». Clémence a pris très peu de notes. Elle porte un jugement global qu'elle justifie en quelques phrases. Elle donne surtout des commentaires négatifs.

#### 4.1.1.5 Comment Marie-Ève a évalué l'étudiant 1

Tableau 4.6  
Commentaires de Marie-Ève au sujet de l'étudiant 1

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie		
	Précision		<i>Elle ne peut de faire au lieu de elle ne peut pas faire. Soins médicaux au lieu de médicaux. Dans le matin au lieu de «le matin». Vous avez l'expérience au lieu de l'expérience». Femme vieille au lieu de vieille femme. Votre papier au lieu de c.v Les femmes prend au lieu de prennent. Il aurait dû utiliser le passé composé au lieu du présent. Il fait des petites erreurs, mais c'est des détails.</i>
	Aisance	<i>Bon débit. Il parle beaucoup.</i>	
Capacité stratégique	Stratégies de réussite		
	Stratégies d'évitement		
Savoir textuel			
Compétence interactive			<i>Il prend la place de l'autre.</i>
Critères non linguistiques			

Marie-Ève a accordé la note globale de 87% à l'étudiant 1. Elle remarque qu'il est «*en contrôle avec son français*», que «*il a une aisance*» même si «*il fait certaines erreurs, mais c'est des petits détails*». Cette enseignante a perçu des erreurs de conjugaison, sans préciser lesquelles, quelques erreurs grammaticales, mais selon elle, ces erreurs ne l'empêchent pas d'obtenir une bonne note. Outre ces remarques sur la performance de l'étudiant 1, cette enseignante dit :

*Ce que je trouve très difficile d'évaluer comme ça, c'est que je les connais pas. Peut-être qu'ils auraient eu une note plus basse si je les avais connus, mais comme là je me sens pas hyper à l'aise de donner une note très basse parce que je leur laisse le bénéfice du doute, c'est dur pour moi de juger.*

Somme toute, Marie-Ève dit n'avoir remarqué que très peu d'erreurs et constaté que cet étudiant parle avec aisance. Elle le juge apte à suivre le niveau suivant.

#### **4.1.1.6 Bilan des commentaires sur l'étudiant 1**

Qu'est-ce qui a surtout retenu l'attention des cinq évaluatrices au sujet de l'étudiant 1? Sur les cinq enseignantes, quatre ont remarqué la même erreur de syntaxe, soit «*femme vieille*» au lieu de «*vieille femme*». Clémence, la seule qui n'ait pas relevé cette erreur, a tout de même mentionné une faute de syntaxe (qu'aucune autre n'a relevé d'ailleurs), soit l'inversion du nom et de l'adjectif dans le groupe nominal «*la prochaine semaine*». Le seul type d'erreur remarqué par les cinq enseignantes en est une d'inversion de l'adjectif et du nom.

Deux enseignantes font des remarques au sujet de la sous-composante phonologique. Élane dit qu'elle avait parfois de la difficulté à le comprendre à cause de ses fautes de prononciation, qu'elle a d'ailleurs relevées de façon exhaustive. Clémence, quant à elle, dit simplement que sa phonétique était moins bonne que celle de sa partenaire.

Notons que Phuong et Clémence font plusieurs commentaires au sujet de la compétence interactive de cet étudiant. Elles disent qu'il manque d'écoute et qu'il ne répond pas aux questions de sa partenaire. Phuong fait un commentaire à l'effet qu'il n'y a pas de contact entre eux. Ce manque d'interaction affecte négativement la note attribuée à l'étudiant 1. Élane et Marie-Ève, qui se sont concentrées sur le relevé des erreurs, n'ont pas fait de commentaire au sujet de la compétence interactive de cet étudiant. Marie-Ève est d'ailleurs celle qui a accordé la meilleure note à l'étudiant 1.

Trois enseignantes déplorent que l'étudiant ait choisi de parler d'un sujet (le hockey) qu'il connaissait mal dans le deuxième jeu de rôle, ce qui affecte de façon négative la bonne impression qu'elles avaient de l'étendue de son vocabulaire dans le premier jeu de rôle. En effet, toutes les enseignantes ont jugé que cet étudiant avait un bon vocabulaire lors du premier jeu de rôle, mais qu'il cherchait ses mots lors du deuxième..

#### 4.1.2 Évaluation de l'étudiante 2

Maintenant que nous avons résumé les commentaires sur l'étudiant 1, voyons ce qui a retenu l'attention chez sa partenaire, l'étudiante 2.

##### 4.1.2.1 Comment Julie a évalué l'étudiante 2

**Tableau 4.7**  
**Commentaires de Julie au sujet de l'étudiante 2**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie	<i>Elle faisait les liaisons entre les mots. Je ne me suis pas rendu compte c'était quoi sa langue maternelle.</i>	
	Précision	<i>Elle respectait la structure des phrases. Elle a vraiment de bonnes structures en français.</i>	

	Aisance	<i>Ça coulait bien.</i>	
Capacité stratégique	Stratégies de réussite	<i>Je n'avais pas l'impression qu'elle traduisait.</i>	
	Stratégies d'évitement		<i>Elle n'a pas beaucoup parlé.</i>
Savoir textuel		<i>Ça s'enchaînait bien.</i>	<i>Elle ne suivait pas avec l'idée de l'autre.</i>
Compétence interactive			<i>Elle aurait dû embarquer [dans la conversation].</i>
Critères non linguistiques			

Julie a accordé la note globale de 73% à l'étudiante 2, soit 80% pour le premier jeu de rôle et 66% pour le deuxième. La raison principale de cet écart est que l'étudiante, pendant le deuxième jeu de rôle, parle très peu. Mais selon Julie, sur le plan de la compétence langagière, cette étudiante s'en tire très bien. En ce qui concerne la phonétique, l'étudiante est jugée «*très compréhensible*», «*elle fait des liaisons entre les mots*». Julie perçoit d'une façon très positive le fait de ne pas être capable de deviner quelle est la langue maternelle de l'étudiante. L'absence d'accent étranger est reçue très favorablement.

De plus, le vocabulaire de cette étudiante, pour le premier jeu de rôle, est jugé très bien. Julie dit, à propos de sa maîtrise du sujet : «*Ça se voyait qu'elle avait travaillé cet aspect-là [le vocabulaire relié au travail] dans le programme du ministère et qu'elle l'avait bien acquis. Les mots qui ont rapport avec le travail.*»

En outre, Julie relève les «*vraiment bonnes structures en français*» de l'étudiante 2. En ce qui concerne son aisance, l'enseignante remarque que son discours «*coulait bien*». Ce qu'elle reproche par contre à l'étudiante, c'est le fait qu'elle «*n'embarque pas dans la deuxième situation*», que «*elle ne suit pas les idées de l'autre.*» Selon Julie, même si elle ne connaissait pas le sport dont il était question, le hockey, «*elle aurait dû embarquer au lieu de répéter qu'elle ne connaissait pas ça*». Au final, l'étudiante 2 obtient donc la même note que son partenaire. Au dernier instant, Julie avoue cependant :

*Même que je lui augmenterais sa note à l'étudiante 2, voyant le niveau, le groupe si on veut. Je ne sais même pas c'est quoi sa langue maternelle parce qu'elle a vraiment de bonnes structures en français. [...] Je changerais peut-être sa note.*

Elle laisse cependant la note inchangée.

#### 4.1.2.2 Comment Éline a évalué l'étudiante 2

**Tableau 4.8**  
**Commentaires d'Éline au sujet de l'étudiante 2**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie	<i>Elle fait de très bonnes liaisons. Elle fait la liaison entre les adjectifs. J'ai tout de suite remarqué ses bonnes liaisons.</i>	
	Précision	<i>Son vocabulaire est très riche.</i>	<i>Elle a dit : «Je peux vous laisser mon ancien employeur numéro de téléphone». Elle a utilisé l'imparfait à la place du passé composé. Elle a dit : «J'ai appris des cours» au lieu de «j'ai pris».</i>
	Aisance	<i>Elle a développé beaucoup d'idées, elle a dit beaucoup de choses.</i>	
Capacité stratégique	Stratégies de réussite		
	Stratégies d'évitement	<i>Elle n'a pas parlé beaucoup dans la deuxième situation, mais ce qu'elle a dit était bien.</i>	
Savoir textuel		<i>C'est très bien articulé. Elle est très bien structurée, la demoiselle.</i>	
Compétence interactive			
Critères non linguistiques			

Élaine a donné la note globale de 92% à l'étudiante 2. Pour la première situation, elle lui donne 92% et, à défaut d'avoir un échantillonnage suffisant dans la deuxième situation, elle lui accorde la même note : *«Je l'ai pas assez entendu. Je n'ai pas assez d'échantillonnage, mais ce qu'elle a dit, c'était bien. C'est surtout lui qui parlait.»* Pour Élaine, le fait de moins parler que son partenaire n'est pas un problème.

Sur les plans de la grammaire et de la phonologie, Élaine fait de bons commentaires. Ainsi, l'étudiante 2 a obtenu, selon la grille d'Élaine, le maximum de points pour «la facilité d'élocution». En ce qui concerne la précision, Élaine remarque seulement quatre erreurs : une de syntaxe, une dans le choix du temps de verbe, une de préposition et une de vocabulaire. Elle affirme que malgré ces erreurs, l'étudiante est *«très bien structurée»*. Pour la première situation, elle obtient une note légèrement supérieure à celle de son partenaire (92% contre 88%). La note de son partenaire chute à 72% dans la deuxième situation, à cause du plus grand nombre d'erreurs de phonétique et de syntaxe, comme nous l'avons expliqué plus haut. L'étudiante 2, quant à elle, n'est pas pénalisée par le fait de n'avoir pas parlé beaucoup dans le deuxième jeu de rôle, puisque sa note du premier jeu de rôle est extrapolée à la deuxième :

*Définitivement, elle est meilleure que lui (l'étudiant 1), même si je n'ai pas entendu plus. Les bonnes liaisons, ça joue beaucoup. Tout de suite, j'ai remarqué ses liaisons et aussi d'autres choses. Elle fait la liaison entre les adjectifs.*

#### 4.1.2.3 Comment Phuong a évalué l'étudiante 2

**Tableau 4.9**  
**Commentaires de Phuong au sujet de l'étudiante 2**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie	<i>Je l'ai adorée, parce qu'elle était claire. Élocution et phonétique claires.</i>	
	Précision	<i>Au niveau du vocabulaire, la syntaxe, la grammaire, la conjugaison, j'ai vu très peu de fautes. Bonne maîtrise de la grammaire. Une richesse de</i>	<i>J'ai vu une seule erreur : «en deux semaines» au lieu de «chaque deux semaines».</i>

		<i>vocabulaire. Elle précise. Sa structure est très solide.</i>	
	Aisance		
Capacité stratégique	Stratégies de réussite	<i>Les objectifs sont très clairs pour cette étudiante.</i>	
	Stratégies d'évitement		
Savoir textuel		<i>Elle est très pertinente au niveau des questions. La façon dont elle passe le contexte est posée dès le départ.</i>	
Compétence interactive		<i>Elle, elle le regarde et tranquillement, il s'en vient vers un contact visuel. Pour moi, ça dénote beaucoup.</i>	<i>La première chose qui m'a choquée, c'est qu'il n'y a pas de contact entre eux. On dirait que chacun s'est composé un rôle. Il n'y a pas de rencontre.</i>
Critères non linguistiques		<i>Je la trouve très... bonne présentation, elle maîtrise le sujet, elle est calme.</i>	

Phuong accorde la note globale de 93% à l'étudiante 2, en faisant le commentaire suivant : «Je l'ai adorée parce qu'elle était claire.». Cette enseignante dit que l'étudiante «*n'a pas de faiblesse réelle.*» et que ses forces se trouvent surtout dans «*tout ce qui est structure grammaticale*». Elle la trouve pertinente dans ses questions et dit qu'elle maîtrise le sujet. En ce qui concerne la compétence interactive, Phuong remarque le fait que les deux partenaires aient peu de contact entre eux.

#### 4.1.2.4 Comment Clémence a évalué l'étudiante 2

Tableau 4.10  
Commentaires de Clémence au sujet de l'étudiante 2

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie	<i>Sa phonétique est meilleure que celle de l'autre [l'étudiant 1].</i>	

	Précision		
	Aisance	<i>Elle est dégagée.</i>	
Capacité stratégique	Stratégies de réussite		
	Stratégies d'évitement		
Savoir textuel		<i>Elle a plus cerné, elle posait les questions. Elle dirigeait plus.</i>	
Compétence interactive		<i>Elle allait le chercher plus.</i>	
Critères non linguistiques			

Clémence a accordé la note globale de 72% à l'étudiante 2. Cette enseignante ne donne pas beaucoup de commentaires pour justifier sa note : elle ne relève pas d'erreurs et s'en tient à des observations globales. Elle dit que pour elle, 72% est une bonne note. Selon elle, l'étudiante 2 «*n'était pas fantastique*», mais elle était «*meilleure que son partenaire*» notamment sur le plan phonétique. Clémence trouve que l'étudiante 2 manifeste plus d'aisance que son partenaire, que contrairement à lui, elle est «*dégagée*», c'est-à-dire à l'aise, puisqu'elle n'a pas besoin de ses notes pour maintenir la conversation. Sur le plan du savoir textuel, l'étudiante 2 se démarque aussi : elle «*dirigeait plus*», c'est elle qui «*allait chercher son partenaire*», qui «*posait les questions*» Clémence trouve que l'étudiante 2 «*était plus solide*».

#### 4.1.2.5 Comment Marie-Ève a évalué l'étudiante 2

Tableau 4.11  
Commentaires de Marie-Ève au sujet de l'étudiante 2

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie		
	Précision		<i>«J'ai appris des cours» au lieu de «pris».</i>
	Aisance	<i>Elle est en confiance avec son français, elle a de l'aisance.</i>	<i>Dans la première situation, elle a des hésitations. Elle cherche ses mots.</i>

Capacité stratégique	Stratégies de réussite		
	Stratégies d'évitement		<i>Parle moins. Elle ne parle pas assez.</i>
Savoir textuel			
Compétence interactive			
Critères non linguistiques		<i>Elle est en contrôle avec son français</i>	

Marie-Ève a accordé la note globale de 83% à cette étudiante, soit 86% pour la première situation et 80% pour la deuxième. Pour elle, c'est une très bonne note. L'enseignante relève une seule erreur, de vocabulaire. Elle invoque deux raisons principales pour justifier sa note : l'étudiante 2 «*ne prend pas sa place dans la conversation*» et elle hésite au début du premier jeu de rôle. Elle dit que l'étudiante «*cherchait ses mots au début du dialogue*». Marie-Ève dit ne pas être en mesure de savoir si ces hésitations proviennent du stress ou d'une faible maîtrise de la langue, faute de connaître l'étudiante suffisamment. Somme toute, cette participante accorde une note plus élevée à l'étudiant 1 qu'à l'étudiante 2 dans la première situation et au global. Elle ne trouve pas, contrairement aux autres, que cette étudiante se démarque sur le plan phonétique, mais la trouve hésitante alors que les quatre autres la jugeaient solide, du moins plus solide que son partenaire.

#### 4.1.2.6 Bilan de l'évaluation de l'étudiante 2

L'étudiante 2 est celle qui a obtenu la meilleure moyenne parmi les quatre étudiants et c'est surtout grâce aux évaluations d'Élaine et de Phuong, les enseignantes les plus expérimentées avec le niveau 5. Trois enseignantes sur cinq lui ont accordé la meilleure note des quatre étudiants. Presque unanimement, elle est perçue comme étant celle qui a la meilleure compétence langagière : ses structures syntaxiques, son vocabulaire et sa bonne prononciation des phonèmes la placent dans une classe supérieure. Cependant, pour deux enseignantes (Julie et Marie-Ève), le fait qu'elle ne parle pas beaucoup dans le deuxième jeu de rôle joue contre elle. En effet, sa maîtrise de la compétence langagière est contrebalancée

par ses faiblesses sur le plan stratégique et textuel. Au contraire, pour Phuong et pour Clémence, elle démontrait une meilleure compétence textuelle et stratégique que son partenaire puisqu'elle démontrait, selon ces dernières, qu'elle contrôlait mieux la situation. Pour Phuong, c'est l'attitude dominatrice et peu collaboratrice de l'étudiant 1 qui est la cause de la faible implication de l'étudiante 2. Pour Clémence, même si l'étudiant 1 parlait plus, il était évident que l'étudiante 2 avait une meilleure maîtrise de la langue. Pour Élane, le fait de ne pas s'imposer dans la conversation n'était pas une mauvaise chose : la bonne impression laissée lors de la première situation n'est pas altérée par la faible participation dans la deuxième situation.

#### 4.1.3 Évaluation de l'étudiant 3

Nous venons de présenter les commentaires au sujet des étudiants de la première dyade. Nous verrons dans la partie suivante ce que les enseignantes ont dit de la deuxième dyade, en commençant par l'étudiant 3.

##### 4.1.3.1 Comment Julie a évalué l'étudiant 3

**Tableau 4.12**  
**Commentaires de Julie au sujet de l'étudiant 3**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie	<i>Il a des intonations plus québécoises. Il a des accents plus québécois.</i>	
	Précision	<i>Il utilisait beaucoup le pronom relatif. Il faisait des phrases complexes avec plusieurs verbes bien utilisés. Il a utilisé le passé composé et l'imparfait.</i>	
	Aisance		<i>Dans la 2<sup>e</sup> situation, il cherchait plus ses mots.</i>
	Stratégies de		

Capacité stratégique	réussite		
	Stratégies d'évitement		<i>Il parlait moins dans la deuxième situation.</i>
Savoir textuel		<i>Il a posé de bonnes questions.</i>	
Compétence interactive		<i>Il écoutait plus sa partenaire.</i>	

Julie a accordé sa meilleure note à l'étudiant 3, soit 80%. Elle lui donne 87% pour le premier jeu de rôle et 73% pour le deuxième. Ses premières observations concernent sa compétence langagière. Julie remarque tout de suite qu'il «*utilisait beaucoup le pronom relatif*», qu'il faisait des «*phrases complexes*» avec «*plusieurs verbes bien utilisés*», au passé composé et à l'imparfait dans le premier jeu de rôle. Donc, l'enseignante voit beaucoup de précision et d'aisance, puisque l'étudiant utilise adéquatement une grande quantité de tournures complexes. Elle trouve qu'il «*cherchait plus ses mots dans la deuxième situation*», que «*il parlait moins*». Sur le plan phonétique, cette participante le trouve «*facile à comprendre*». Elle note aussi qu'il a «*des intonations plus québécoises, une fluidité, des accents plus québécois, des lettres qu'il va dire à la québécoise*». Julie dit que «*son oreille est plus développée*». Sur le plan textuel, elle soutient qu'il a su «*poser de bonnes questions*» à sa partenaire. L'enseignante rapporte le fait que l'étudiant 3 a parlé de la défunte équipe de hockey des Nordiques de Québec, expatriée à Denver en 1995. Il a dit que c'était son équipe favorite avant son déménagement. Julie tire donc la conclusion que l'étudiant 3 est au Québec depuis plus de dix ans (ce qui est d'ailleurs le cas).

#### 4.1.3.2 Comment Élane a évalué l'étudiant 3

**Tableau 4.13**  
**Commentaires d'Élane au sujet de l'étudiant 3**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie	<i>Il fait très bien ses liaisons. Son accent est plus agréable. Il prononce bien.</i>	<i>Il ne parle vraiment pas assez fort. Il faut vraiment tendre l'oreille pour le comprendre.</i>
	Précision	<i>Au niveau du vocabulaire, son message est très clair.</i>	<i>J'ai entendu : «faire une décision» au lieu de «prendre».</i>

		<i>Monsieur a de meilleures structures de madame. Lui, c'est très bien. 2 fois, il avait bien son pronom relatif. Il utilise le passé composé et l'imparfait. Il a mieux son présent qu'elle (l'étudiante 4).</i>	<i>Il a eu de la difficulté : «nous rencontrerez» il a mal conjugué son verbe. Il a dit : «on va payer en comptant» au lieu de «comptant». Il a fait deux erreurs avec la place de son adverbe.</i>
	Aisance	<i>C'est sûr qu'il va avoir une meilleure note pour la facilité d'élocution. Il dit beaucoup de choses. Son rythme est bon.</i>	
Capacité stratégique	Stratégies de réussite		
	Stratégies d'évitement		
Savoir textuel		<i>Je pense qu'il est mieux structuré.</i>	
Compétence interactive		<i>Ils s'entendent bien : lorsqu'elle a dit : «Combien ça coûte pour payer». C'est pas clair. Pourtant, il a compris. C'est quand même particulier.</i>	
Critères non linguistiques			

Élaine accorde la même note à l'étudiant 3 qu'à l'étudiant 1, c'est-à-dire 80%. Elle lui donne 80% pour les deux situations. Sur le plan de la compétence langagière, elle remarque, comme Julie, l'utilisation adéquate de pronoms relatifs, du présent, du passé composé, de l'imparfait et du conditionnel. Selon cette enseignante, il fait de bonnes liaisons et il ne commet pas les erreurs de prononciation typiques des hispanophones. Pour cette participante, «*son accent est agréable*». Elle note aussi sa bonne compétence interactive : en effet, Élane pense que la partenaire de l'étudiant 3 est souvent très difficile à comprendre, mais que malgré ça, «*ils s'entendent bien*». Elle soutient, dans la situation 1, une des questions qu'a posée l'étudiante 4, «*un employeur ne l'aurait pas comprise*», mais que

«*pourtant, l'étudiant 3 l'a compris.*» Elle ajoute que «*c'est quant même particulier*». Le seul aspect négatif au sujet de cet étudiant, selon Éline, c'est son volume trop faible. L'enseignante croit qu'il ne prononce pas assez fort, qu'il devrait mieux articuler, que «*il faut vraiment tendre l'oreille pour le comprendre*» et pour elle, l'intelligibilité du message en souffre beaucoup.

#### 4.1.3.3 Comment Phuong a évalué l'étudiant 3

Tableau 4.14  
Commentaires de Phuong au sujet de l'étudiant 3

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie		
	Précision		<i>Au niveau de la conjugaison, on voit : ils sont disparaître au lieu de disparus. Il y a des erreurs de conjugaison à corriger.</i>
	Aisance	<i>Ce que j'aime chez lui, c'est sa fluidité. Il y a une fluidité dans sa communication, les mots viennent simplement.</i>	
Capacité stratégique	Stratégies de réussite		
	Stratégies d'évitement		
Savoir textuel			
Compétence interactive		<i>Il y a une intention d'échange. Pour moi, c'est important. J'aime son écoute. Il essaie de supporter l'autre pour qu'elle puisse s'exprimer. Il y a un contact simple.</i>	
Critères non linguistiques		<i>J'aime sa simplicité.</i>	

Phuong accorde sa deuxième meilleure note à l'étudiant 3, c'est-à-dire 87%. Comme pour les étudiants précédents, elle évalue les deux situations d'un seul bloc. Au sujet de la précision, elle relève «*des erreurs de conjugaison à corriger*». Au sujet de l'aisance, elle note aussi sa «*fluidité*», elle estime que «*les mots viennent simplement*». Phuong remarque aussi qu'il pose «*beaucoup de questions*». Donc, tant son aisance que sa compétence textuelle sont soulignées. Elle trouve que l'étudiant s'exprime avec des «*phrases toutes simples*». L'enseignante remarque aussi «*son écoute*» puisqu'il soutient sa partenaire : «*il essaye de la supporter pour qu'elle puisse s'exprimer*».

#### 4.1.3.4 Comment Clémence a évalué l'étudiant 3

**Tableau 4.15**  
**Commentaires de Clémence au sujet de l'étudiant 3**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie		
	Précision	<i>Ses structures étaient meilleures que celles de l'autre gars [l'étudiant 1]. Il est allé dans l'imparfait.</i>	
	Aisance		
Capacité stratégique	Stratégies de réussite		
	Stratégies d'évitement		
Savoir textuel			
Compétence interactive			
Critères non linguistiques			

Clémence accorde la note de 70% à cet étudiant. Elle justifie sa note en disant qu'elle «*pense que ce gars-là (l'étudiant 3) est meilleur que l'autre gars (l'étudiant 1)*». Au sujet de sa compétence langagière, elle dit, sans préciser lesquelles, que ses «*structures étaient meilleures [que celle de l'étudiant 1]*», qu'il «*est allé dans l'imparfait*». Pour cette

enseignante, 70% est une bonne note, c'est d'ailleurs sa deuxième meilleure note (elle a accordé 72 à l'étudiante 2).

#### 4.1.3.5 Comment Marie-Ève a évalué l'étudiant 3

**Tableau 4.16**  
**Commentaires de Julie au sujet de l'étudiant 3**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie	<i>Il a son accent latino, mais ça enlève pas ta compréhension.</i>	<i>J'entendais pas tout le temps bien.</i>
	Précision	<i>Il utilise plein de temps de verbes. Son vocabulaire est pas limité.</i>	<i>Je toujours me souviens. Ils sont disparaître. Il fait des petites erreurs, mais c'est pas des erreurs graves.</i>
	Aisance	<i>Il est en contrôle. Le débit est bon.</i>	
Capacité stratégique	Stratégies de réussite		
	Stratégies d'évitement		
Savoir textuel			
Compétence interactive			
Critères non linguistique			

Comme Julie, Marie-Ève a réservé sa meilleure note à l'étudiant 3, soit 93%. Elle lui accorde cette note pour les deux jeux de rôle. Ce sont des remarques au sujet de sa compétence langagière qui justifient sa note. Au sujet de la précision, l'enseignante estime qu'il «*utilise plein de temps de verbes*» (sans mentionner lesquels), elle dit que «*son vocabulaire n'est pas limité*». Elle ne voit que deux erreurs : une de syntaxe et une de conjugaison (la même que Phuong), mais ces erreurs n'altèrent pas sa bonne prestation en ce qui concerne sa précision. Marie-Ève apprécie aussi son bon débit, son aisance à s'exprimer.

#### 4.1.3.6 Bilan de l'évaluation de l'étudiant 3

L'étudiant 3 s'en tire donc presque aussi bien que l'étudiante 2, avec une moyenne presque équivalente, de 82% (contre 83% pour l'étudiante 2). Sa compétence langagière est soulignée par toutes les enseignantes. Quatre enseignantes (Julie, Élane, Clémence et Marie-Ève) font des commentaires au sujet de sa bonne utilisation de plusieurs temps de verbes. Julie, Élane et Marie-Ève observent l'utilisation de pronoms relatifs. Rythme, fluidité, débit : quatre enseignantes remarquent son aisance. Trois enseignantes parlent de son accent. Une le qualifie de «*agréable*», une autre dit qu'il est «*à la québécoise*» et la troisième dit que «*son accent latino n'enlève rien à la compréhension*». Trois enseignantes soulignent la qualité de ses questions et deux d'entre elles vont jusqu'à complimenter sa compétence interactive, puisqu'il aide sa partenaire, jugée plus faible que lui.

#### 4.1.4 Évaluation de l'étudiante 4

Nous venons de voir l'ensemble des commentaires du troisième étudiant. Nous verrons dans la partie suivante ce que les enseignantes ont pensé de sa partenaire, l'étudiante 4.

##### 4.1.4.1 Comment Julie a évalué l'étudiante 4

**Tableau 4.17**  
Commentaires de Julie au sujet de l'étudiante 4

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie		<i>La phonétique était très anglophone. Elle prononçait les «s».</i>
	Précision		<i>Elle fait des erreurs de niveau débutant. Elle se mélangeait entre un verbe et un nom. Elle dit «de le» au lieu de «du». Elle ne tenait pas compte du</i>

			<i>genre. Il lui manquait beaucoup de mots. Il y avait des structures anglaises.</i>
	Aisance		<i>C'était très détaché.</i>
Capacité stratégique	Stratégies de réussite	<i>Elle a fait un effort pour parler. Elle communique bien. Elle s'exprime quand même, on la comprend.</i>	<i>On entendait beaucoup d'anglicismes. Ça paraissait qu'elle traduisait de l'anglais mot à mot.</i>
	Stratégies d'évitement		
Savoir textuel			<i>Elle a de la difficulté à formuler ses questions. Elle ne répondait pas bien à des questions simples.</i>
Compétence interactive			
Critères non linguistique		<i>C'est une personne dynamique. Elle a beaucoup d'idées.</i>	

La dernière étudiante est celle qui a obtenu la moins bonne note de la part de Julie, soit 67%. L'enseignante lui accorde le même pointage pour les deux situations. Au sujet de la précision, elle dit:

*Elle s'exprime bien, sauf qu'elle fait des erreurs de niveau débutant. Beaucoup d'erreurs que présentement, avec mon niveau 1, on est en train de voir. Comme demander «combien vous payez» au lieu de «combien ça coûte», ou «Quel est votre détesté?» Elle se mélangeait entre un verbe et un nom. Elle fait des erreurs aussi de genre: «de le»: de le annonce» le monsieur et le madame, des choses de niveau débutant, 1, 2, 3.*

Sur le plan du vocabulaire, Julie trouve que «*elle cherche ses mots*». Son discours apparaît «*très détaché*». De plus, sa phonétique est perçue comme très anglophone en partie parce que l'étudiante prononce les consonnes muettes à la fin des mots. Son savoir textuel est jugé faible aussi, parce que l'étudiante n'arrive pas à répondre à des questions simples et qu'elle n'arrive pas à formuler des questions de base. Julie note que «*elle traduisait beaucoup trop*». Par contre, l'enseignante dit que cette étudiante «*parlait beaucoup, elle fait un effort*» et que

«Elle s'exprime bien.» Julie remarque que l'étudiante 4 «a beaucoup d'idées» que «elle veut tellement en dire que ça sort tout croche» que «c'est une personne dynamique».

#### 4.1.4.2 Comment Éline a évalué l'étudiante 4

**Tableau 4.18**  
**Commentaires d'Éline au sujet de l'étudiante 4**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie	<i>En général, elle parle plus fort que lui, alors on la comprend mieux. Elle prononce clairement. Elle a de l'intonation, elle va aller nous chercher avec son intonation.</i>	<i>Elle ne fait pas ses liaisons, donc, ça donne plus un accent.</i>
	Précision	<i>Elle mérite son 70% parce qu'elle a bien son passé composé. Elle a de l'imparfait. Elle l'utilise à 2 reprises. Le présent : elle en a eu 1 de bon.</i>	<i>Elle a dit : «À quel jour vous préférez commencer me» et «combien ça coûte pour payer». Il lui manque son «d'» Au niveau de la syntaxe, elle pose pas bien ses questions. Elle a de la difficulté avec la phrase interrogative. Elle mêle «malade» et «maladie». Elle perd des points au niveau de la structure. Elle mélange le féminin et le masculin. Elle a manqué son présent à 2 reprises.</i>
	Aisance		<i>C'est pas enchaîné, ça manque de fluidité. Elle avait beaucoup d'hésitations. Elle hésite avant de poser ses questions.</i>
Capacité stratégique	Stratégies de réussite	<i>Mademoiselle s'auto corrige, ce qui est très bien. Avec sa gestuelle, son regard, elle va capter l'attention. Même si ses</i>	<i>«Payement» : elle utilise la langue maternelle, en l'occurrence l'anglais.</i>

		<i>structures sont déficientes, ça compense.</i>	
	Stratégies d'évitement		
Savoir textuel			
Compétence interactive			
Critères non linguistiques		<i>Elle est expressive.</i>	

Élaine accorde la note globale de 70% à cette étudiante. Pourtant, en compilant les notes de sa première situation (72%) et de sa deuxième (64%), nous obtenons une moyenne de 68%. Mais Élane tenait à accorder la note «symbolique» de 70% en disant «*qu'elle la méritait*». Elle commente longuement sa compétence langagière. Elle note de nombreuses lacunes syntaxiques, lexicales, des erreurs de genre, mais aussi une bonne maîtrise du passé composé et de l'imparfait. Au sujet de l'aisance, l'enseignante remarque que l'étudiante 4 a du mal à faire ses liaisons, un manque d'enchaînement et de fluidité, ce qui lui «*donne un accent*». Elle relève aussi de nombreuses hésitations. Au sujet de la phonologie, Élane apprécie le fait que cette étudiante fasse un effort pour parler clairement, qu'elle parle fort, même si ce qu'elle prononce est teinté d'un fort accent anglophone. Élane remarque aussi de grandes faiblesses syntaxiques, puisque cette étudiante n'arrive pas à formuler des questions simples et sur le plan textuel, et qu'à quelques reprises, elle n'arrive pas à répondre à des questions simples. L'enseignante croit que «*sa gestuelle et son regard vont capter l'attention*», que «*même si ses structures sont déficientes, ça compense*». Selon cette enseignante, l'étudiante arrive à faire passer un message clair. De plus, Élane juge favorablement le fait que l'étudiante 4 «*s'auto corrige*». Cela dénoterait que «*elle est à un bon niveau, un niveau avancé*».

#### 4.1.4.3 Comment Phuong évalue l'étudiante 4

**Tableau 4.19**  
**Commentaires de Phuong au sujet de l'étudiante 4**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
	Phonologie		<i>Elle a un accent</i>

Composante langagière	Précision		<i>anglophone très fort.</i> <i>Le masculin et le féminin : elle ne maîtrise pas.</i> <i>«Hockey joueur» au lieu de «joueur de hockey».</i> <i>On voit un manque de vocabulaire, un manque de précision.</i> <i>«Malade» au lieu de «maladie».</i>
	Aisance		
Capacité stratégique	Stratégies de réussite		<i>«Combien ça coûte pour payer?» : elle traduit de l'anglais.</i>
	Stratégies d'évitement		
Savoir textuel			<i>On voit pas où elle veut aller. C'est ni focussé, ni pointé.</i>
Compétence interactive		<i>Elle a le sens de la communication. La transmission et la réception d'information : elle a ça.</i>	
Critères non linguistiques			<i>C'est beaucoup plus dans l'émotion.</i>

Phuong donne la note globale de 80% à cette étudiante. Comme ses collègues, Phuong remarque «un manque de vocabulaire et un manque de précision». Elle note que «elle cherche ses mots» et que «on ne sait pas trop où elle veut aller». Sur le plan phonologique, elle remarque un «accent anglophone très fort», que «sa phonétique n'est pas très juste». Comme les deux enseignantes précédentes, Phuong trouve que cette étudiante communique bien malgré tout. Donc, malgré des lacunes importantes, l'étudiante s'en tire très bien puisque Phuong lui accorde la note de 80%. Il s'agit quand même de la note la plus faible accordée par cette enseignante, note partagée avec l'étudiant 1.

## 4.1.4.4 Comment Clémence a évalué l'étudiante 4

**Tableau 4.20**  
**Commentaires de Clémence au sujet de l'étudiante 4**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie		
	Précision		<i>Elle faisait des erreurs, mais d'anglophone. J'ai vu plus d'erreurs ici.</i>
	Aisance		
Capacité stratégique	Stratégies de réussite	<i>De temps en temps, elle avait le tour d'aller nous chercher. Elle utilisait le geste, elle nous avait avec ça.</i>	
	Stratégies d'évitement		
Savoir textuel			
Compétence interactive			
Critères non linguistiques		<i>Le charme aussi. On voyait qu'elle savait quand elle faisait des erreurs. On voyait qu'elle le savait et ça la choquait de pas avoir le mot, elle s'apercevait de ses erreurs.</i>	

Clémence attribue 70% à l'étudiante 4, soit la même note que son partenaire. En guise d'unique commentaire au sujet de la compétence langagière, elle dit que «*elle faisait pas mal d'erreurs*», mais que l'étudiante «*essayait quand même d'aller nous chercher. Elle utilisait le geste et le charme aussi.*». Autre aspect positif, comme Élane, elle remarque que l'étudiante «*s'apercevait de ses erreurs*» que «*ça la choquait*». Elle ne va pas jusqu'à dire, comme Élane, que l'étudiante se corrige, mais du moins, qu'elle est consciente de ses faiblesses. Donc, l'étudiante mérite selon elle la même note que son partenaire et une meilleure note que l'étudiant 1, à qui elle avait accordé 68%, malgré une plus grande compétence langagière.

## 4.1.4.5 Comment Marie-Ève a évalué l'étudiante 4

**Tableau 4.21**  
**Commentaires de Marie-Ève au sujet de l'étudiante 4**

Composantes	Sous-composantes	Remarques positives	Remarques négatives
Composante langagière	Phonologie		<i>Problème avec les liaisons Prononce le «t» final de «moment» Le français, quand tu prononce trop d lettres, ça change. Ses fins de mots : »je préféré». Quand c'est trop mal prononcé, on comprend pas.</i>
	Précision		<i>De le annonce. Malade au lieu de maladie. Vocabulaire limité. Elle fait beaucoup trop d'erreurs. Ses structures sont pas bonnes. Ses verbes sont pas conjugués.</i>
	Aisance	<i>C'est pas qu'elle parle lentement, son débit est bon.</i>	
Capacité stratégique	Stratégies de réussite		
	Stratégies d'évitement		
Savoir textuel			<i>Elle répond pas bien aux questions.</i>
Compétence interactive			
Critères non linguistique			

La dernière enseignante, pour sa part, n'accorde pas la note de passage à l'étudiante 4. Elle lui donne 60% pour la première situation et 46% pour la deuxième, pour un total de 53%. Elle justifie sa note principalement de cette façon :

*Quand on lui pose une question de base : «Quel est ton joueur préféré? Elle répond «Les Canadiens». Donc, selon moi, ça passe pas. 16/30. Ça pourrait être 17/30, mais de toute façon, ça passe pas, donc que je lui mets 16. C'est tout en fait : ses structures sont pas bonnes, ses verbes sont pas conjugués, la prononciation.*

Pour Marie-Ève, ses faiblesses sont trop grandes pour qu'elle puisse réussir son examen.

#### **4.1.4.6 Bilan de l'évaluation de l'étudiante 4**

Globalement, l'étudiante 4 s'en sort assez bien, même si elle obtient la moins bonne moyenne des quatre étudiants. Toutes remarquent ses lacunes sur le plan langagier, mais pour quatre des cinq enseignantes, sa capacité stratégique, sa compétence interactive et des critères non linguistiques compensent pour ses faiblesses. Élane donne une bonne note à l'étudiante, surtout à cause de sa maîtrise du passé composé et de l'imparfait. Phuong et Clémence accordent une note très satisfaisante à cette étudiante.

Nous venons de présenter un résumé des commentaires émis par les enseignantes au sujet de la performance des quatre étudiants. Dans la partie suivante, nous présenterons les résultats obtenus lors de la deuxième partie des entrevues, qui concernait les pratiques évaluatives des cinq participantes et les critères personnels auxquels elles disent accorder de l'importance lorsqu'elles évaluent.

## **4.2 Présentation des résultats de la deuxième partie des entrevues**

Comme nous l'avons mentionné dans le cadre théorique, soit le chapitre II, notre entrevue s'est déroulée en deux temps. D'abord, nous avons demandé aux enseignantes d'évaluer les étudiants et de commenter leur performance. Ensuite, nous avons demandé aux enseignantes de nous parler de leurs critères d'évaluation personnels et de la façon dont elles s'y prennent en général pour évaluer leurs étudiants. Le tableau 4.22 résume l'ensemble des critères évoqués par les participantes.

Nous leur avons aussi demandé de nous parler de l'importance qu'elles accordent à chacun de ces critères, et de leur attribuer une note sur une échelle de un à cinq. Le tableau

4.22 regroupe les réponses de toutes les enseignantes. Au total, elles ont identifié huit catégories générales qui constituent autant de critères d'évaluation. La colonne de gauche présente les critères et les colonnes subséquentes montrent la note, de un à cinq, que les enseignantes accordent aux critères qu'elles ont mentionnés. Le tiret «-» signifie que l'enseignante n'a pas mentionné le critère. La colonne de droite, titrée «total», indique combien de participantes ont mentionné ce critère. La ligne du bas montre le nombre de critères différents mentionnés par les participantes.

Notons que dans trois cas (Julie, Éline, Simon), la note concernant la phonologie a été mise entre parenthèse. La raison est la suivante : ces trois enseignantes n'ont pas différencié la phonologie de la phonétique. Pour elles, il s'agit d'un même critère. Nous avons donc accordé la même note à ces deux critères. Par contre, pour Clémence et Marie-Ève, phonétique et phonologie sont deux critères distincts. C'est pourquoi nous avons privilégié la présentation différenciée de ces critères, qui sont, somme toute, différents. En outre, Éline a mentionné le critère très général «grammaire» lors de l'entrevue. Comme ce critère englobait deux aspects mentionnés par les autres enseignantes (accords, conjugaison) nous l'avons décomposé en sous aspects afin de mieux le représenter. Finalement, Julie, Phuong, Clémence et Marie-Ève ont mentionné un critère d'évaluation que nous n'avons pas cru bon d'ajouter au tableau. Il s'agit de l'utilisation adéquate des éléments linguistiques appris pendant le cours. Pour ces enseignantes, pendant l'examen, l'étudiant doit démontrer qu'il a retenu ce qui a été à l'étude pendant le cours. Mais comme il s'agit, selon elles, essentiellement de démontrer sa connaissance du vocabulaire, de la conjugaison, de la phonétique et de la syntaxe, nous avons omis d'inclure ce critère dans le tableau, puisqu'il aurait été redondant. Contrairement aux autres enseignantes, Éline ne mentionne pas ce critère, mais elle l'intègre dans sa grille de correction, puisque les éléments linguistiques propres au niveau se retrouvent dans différentes catégories de sa grille.

**Tableau 4.22**  
**Critères des enseignantes**

	<b>Julie</b>	<b>Élaine</b>	<b>Phuong</b>	<b>Clémence</b>	<b>Marie-Ève</b>	<b>Total</b>
1. Respect de la consigne	3/5	-	5/5	3/5	-	3
2. Clarté du message	-	3/5	4/5	-	-	2
3. Phonétique	4/5	3/5	3/5	4/5	4/5	5
4. Phonologie (enchaînements et liaisons)	(4/5)	(3/5)	(3/5)	4/5	3/5	5
5. Aisance (débit)	4/5	3/5	3/5	4/5	3/5	5
6. Syntaxe	3/5	4/5	4/5	5/5	-	4
7. Verbes (Conjugaison, expression de temps et concordance)	4/5	3/5	4/5	5/5	4/5	5
8. Vocabulaire	4/5	2/5	3/5	5/5	4/5	5
Nombre de critères mentionnés	7	8	8	7	5	

Dans la partie suivante, nous verrons comment les enseignantes définissent les critères qu'elles ont choisis et comment se déroule l'évaluation dans leurs classes.

#### 4.2.1 Critères de Julie

Julie nous a dit être «assez traditionnelle dans [sa] correction». Elle se base sur cinq critères principaux pour évaluer. Le premier est la phonologie, qui englobe la phonétique, le débit, l'intonation et les liaisons. Ce critère est important selon elle parce que «ça peut jouer tellement sur la compréhension». C'est aussi le critère «qui fait la différence entre l'écrit et l'oral». À titre d'exemple, elle parle des asiatiques, qui, souvent, «connaissent le mot, [qui] savent comment ça s'écrit, mais [qui] sont pas capables de le prononcer». Pour elle, ce critère est le plus «technique», voire «mécanique».

Comme deuxième critère, elle mentionne la conjugaison. Elle dit accorder beaucoup d'importance aux verbes. Pour elle, un étudiant qui néglige de conjuguer ses verbes et qui se limite à utiliser l'infinitif au niveau avancé, va arriver à se faire comprendre, mais aura toujours un discours «de bébé», un discours très «pauvre». Elle veut que ses élèves «soient

*capables d'utiliser les verbes étudiés en classe*». Pour elle, le verbe est le «*noyau de la phrase*».

En troisième, elle mentionne la réutilisation du vocabulaire appris pendant la session. Elle juge ce vocabulaire comme étant très important, puisque selon elle, il va aider à la «*fluidité des idées*». La richesse du vocabulaire «*va faire la différence entre une très bonne note et une note moyenne*», c'est aussi, selon Julie, ce qui permet à l'élève de produire des phrases plus complexes.

En quatrième vient la syntaxe. C'est un critère auquel elle accorde un peu moins d'importance qu'aux trois premiers, puisque selon elle, un étudiant peut arriver à se faire comprendre avec des verbes et du vocabulaire, sans avoir une très bonne syntaxe. Pour elle, la syntaxe est «*la colle qui vient mettre ça ensemble*».

Comme dernier critère, elle signale la cohérence du message. Pour elle, ce critère est représenté par la capacité de l'étudiant à répondre adéquatement à la question d'examen. Il s'agit donc du plus «*académique*» des critères : un étudiant peut s'exprimer très bien, mais ne pas avoir bien compris la question et avoir un discours plus ou moins cohérent. Mais la cohérence ne se limite pas à l'adéquation de la réponse à la question d'examen, mais aussi à la capacité de suivre le fil de la conversation de façon logique, la fluidité des idées.

Julie nous a confié que l'évaluation finale «*n'[était] pas une fin en soi*». Elle dit observer ses élèves tout au long de la session et se servir de l'examen «*seulement pour confirmer*» ses observations. De plus, elle ne suit pas toujours le prototype d'examen et elle modifie les jeux de rôle pour les rendre plus «*réalistes, plus dynamiques*», en conformité avec les besoins de ses étudiants. En outre, elle fait passer à ses étudiants un examen oral de mi-session, soit un exposé oral. Il vaut pour un tiers de la note finale à l'oral.

Au sujet de la grille de la DGFA (1994), elle ne s'en sert pas. Elle s'en est fait une qu'elle utilise pour évaluer l'interaction orale de ses étudiants (voire appendice E). Il s'agit d'une grille analytique composée de quatre critères, plutôt cohérents avec les aspects mentionnés plus haut : le message, la grammaire, le vocabulaire et la phonétique qui valent cinq points chacun. L'addition des points lui donne un total sur 20, qu'elle ramène à la

pondération exigée par le niveau. Toutefois, pendant notre rencontre, Julie n'a pas utilisé sa grille personnelle pour évaluer les étudiants. Elle a pris des notes sur une feuille réservée à cette fin et a donné une note globale qu'elle a justifiée à l'aide des commentaires qu'elle avait écrits.

#### 4.2.2 Critères d'Élaine

Comme Julie, Élane n'utilise pas la grille du MELS (1994) pour évaluer ses étudiants. Elle a confectionné une grille analytique à l'aide des critères généraux du ministère et en s'inspirant de son expérience à l'université (voir appendice F). Sa grille renferme donc cinq critères : le message, les structures, la grammaire, le vocabulaire et la facilité d'élocution. Élane trouve sa grille *«plus simple»* que celle du MELS et plus facile à utiliser. Par exemple, dans les descripteurs de la grille du MELS, nous pouvons lire : «Plusieurs éléments linguistiques sont personnels». Élane commente : *«Ça me dit rien, un «élément linguistique personnel». Les éléments linguistiques, pour moi, c'est la grammaire, les verbes. Ce langage-là, je le comprends mieux.»*

De plus, elle trouve la grille du MELS trop générale. Pour elle, une grille analytique détaillée est plus simple d'utilisation.

Élane a une façon d'évaluer qui est différente des autres enseignantes. Dans le critère «grammaire» de sa grille, elle fait la liste des éléments grammaticaux qu'elle s'attend à retrouver lors de la présentation, que ce soit l'utilisation d'un temps ou d'un mode verbal particulier, des pronoms relatifs ou des pronoms personnels par exemple. Quand elle entend l'étudiant utiliser cet élément grammatical, elle coche la case appropriée. Elle se justifie en disant : *«Je ne peux pas lui [son étudiant] dire seulement ses erreurs, faut lui dire ses bons coups aussi. Si je vois un passé composé, je vais lui en parler.»*

Donc, en plus de noter, comme les autres le font, les erreurs commises par les étudiants, elle note aussi les éléments grammaticaux qu'ils ont systématisés. Pendant notre entrevue, Éleine a utilisé sa grille pour évaluer les étudiants.

Pour elle, le critère le plus important est la syntaxe. L'intelligibilité du message dépend en grande partie de ce critère. La grammaire, la facilité d'élocution et la clarté du message viennent ensuite. Par grammaire, Éleine entend la conjugaison, la concordance des temps, le genre, les prépositions et les déterminants et dans le critère «facilité d'élocution», elle rassemble les éléments prosodiques, la phonétique et le débit. Cette enseignante a inclus dans sa grille le critère «message clair». Le dernier critère d'Éleine est le vocabulaire. C'est celui auquel elle accorde le moins d'importance. Elle dit que *«les mots sont importants»*, que *«tu peux utiliser des mots, sans utiliser de verbes et [que] les gens vont te comprendre»*, mais selon elle, *«la structure est plus importante que le vocabulaire»*. Notons qu'avec une grille analytique où chaque critère vaut le même nombre de point, comme celle qu'Éleine utilise, le fait d'accorder plus ou moins d'importance à un critère ne se reflète pas dans la pondération. Dans la grille de cette enseignante, chaque critère vaut cinq points, le vocabulaire comme la grammaire.

#### 4.2.3 Critères de Phuong

Phuong, contrairement à Éleine et à Julie, se dit à l'aise avec la grille du MELS. Elle préfère donner une note globale plutôt qu'une note détaillée. Cependant, comme Julie, elle modifie les prototypes d'examen pour les adapter aux particularités de ses étudiants. Elle a utilisé la grille du MELS (1994) pour évaluer les étudiants lors de l'entrevue. De plus, comme Julie, en mi-session, cette enseignante demande à ses étudiants de faire un exposé oral, qui compte pour la moitié de la note finale.

Phuong est l'enseignante qui a énuméré le plus grand nombre de critères d'évaluation. D'abord, l'étudiant doit «répondre aux objectifs de l'examen». C'est selon elle le critère le plus important. Elle dit :

*Mon objectif principal, c'est s'ils ont bien compris la consigne. S'ils ont mal compris, c'est déjà à côté. Parce que même s'ils ont une bonne structure, même s'ils maîtrisent la conjugaison, tout ça, t'as pas bien compris le sens et pour moi, le sens est manquant.*

Contrairement à Julie et à Éline, Phuong dit évaluer le message général de façon globale plutôt qu'analytique. Elle justifie sa note en invoquant par la suite des critères précis. La syntaxe, la conjugaison et la concordance des temps sont ceux qu'elle juge les plus importants. Pour elle, ces critères relèvent «des lois qui gouvernent la langue, des principes sous-jacents», c'est «ce qui se passe au niveau de l'intelligence, les liens qui se font.» Vient ensuite le vocabulaire, dont la richesse apporte «du raffinement», de la «précision», sans être de première importance. L'avant-dernier critère mentionné est la phonétique. Selon elle, la phonétique, «c'est l'aspect extérieur, c'est esthétique» et «mécanique». Elle dit :

*Pour moi, c'est pas la prononciation des phonèmes qui m'importe. Si la personne a son accent, ça me dérange pas du tout, c'est à part. Mais attention : quand ça change le sens, par exemple, crevette et cravate : «Je porte une belle crevette», là, ça va pas du tout.*

Finalement, cette enseignante dit accorder une importance capitale au fait que l'étudiant soit capable de «transférer ses connaissances dans plusieurs domaines». Pour le vérifier, elle dit que «dans un examen [elle va] toujours donner l'alternative» à l'étudiant, c'est-à-dire une deuxième mise en situation (comme le prévoit la DGFA).

#### 4.2.4 Critères de Clémence

Comme Julie et Éline, Clémence utilise une grille analytique qu'elle s'est confectionnée pour évaluer ses étudiants (voire appendice G). Elle nous a dit que le contenu de la grille était adapté aux objectifs du niveau, c'est-à-dire qu'elle s'attend à retrouver pendant la production orale les éléments linguistiques enseignés pendant le cours. Comme Julie et Éline, Clémence nous a confié qu'elle accordait plus d'importance à ses observations quotidiennes qu'à l'examen final. Elle estime qu'après huit semaines, elle connaît ses étudiants. Elle dit :

*Je les observe souvent, je les observe tout le temps. Je le sais : ces trois-là sont super bons, ceux-là sont moyens, ceux-là en arrachent. Ça joue un rôle très important, ces observations.*

Les critères d'évaluation de Clémence sont au nombre de quatre. Comme pour Julie et Phuong, l'étudiant doit d'abord produire un discours conformément à la tâche demandée. Elle dit : *«c'est rare qu'ils [les étudiants] se trompent là-dessus.* Ce critère est plus relié à la compréhension qu'à la production orale, mais elle en tient compte. Ce à quoi Clémence accorde le plus d'importance, c'est à la capacité de l'étudiant à reproduire correctement dans l'examen oral les temps de verbes, les structures de phrases et le vocabulaire appris pendant le cours. Ensuite, elle accorde de l'importance au rythme (temps d'exécution et fluidité de la phrase) et finalement, à la phonétique. Dans cette dernière catégorie, elle englobe la prononciation de la finale des verbes. Pour elle, un verbe dont la finale n'est pas bien prononcé (par exemple :«je marché» au lieu de «je marche») relève d'une erreur de phonétique plutôt que de conjugaison, puisque l'étudiant peut écrire correctement le verbe, mais oublie que le «e» finale ne se prononce pas en français. Ce critère tient compte aussi des interférences avec la langue maternelle, du fait que l'accent de l'étudiant affecte le message et joue sur son intelligibilité. Clémence établit une distinction claire entre l'aisance à communiquer et la phonétique.

#### **4.2.5 Critères de Marie-Ève**

Comme Julie, Éline et Clémence, Marie-Ève s'appuie davantage sur les observations qu'elle a faites pendant la session que sur l'examen pour évaluer ses étudiants. Elle dit : *«Un examen comme ça, c'est juste une forme pour l'évaluation parce qu'on voit plus ce qui se passe dans la classe qu'à l'examen final.»* D'ailleurs, pendant notre entrevue, elle dit avoir trouvé difficile d'évaluer des étudiants qu'elle ne connaissait pas. Par exemple, quand une étudiante cherchait ses mots, elle disait ne pas savoir si elle devait attribuer ses hésitations au stress ou à des lacunes linguistiques. En conséquence, elle ne savait pas si elle devait pénaliser ou non les hésitations de cette étudiante, qui par ailleurs, lui avait paru *«en contrôle avec son français»*. De plus, comme Julie et Phuong, Marie-Ève ajuste l'examen aux

particularités de ses étudiants. Si elle voit qu'un étudiant n'est pas particulièrement inspiré par un jeu de rôle, elle va essayer d'orienter le jeu de rôle de façon à lui permettre d'exprimer plus d'idées. En outre, elle dit :

*Souvent, les étudiants se préparent juste à ça, l'examen final, et j'aime pas ça. Donc, souvent, avant l'examen, je leur pose des questions : Ça va ? Es-tu stressé ? Qu'est-ce que tu as fait hier ? Qu'est-ce que vous allez faire après l'examen ? Et là, des fois, ça prend une tout autre direction et c'est sur ça que j'évalue.*

Finalement, comme Phuong, Marie-Ève utilise la grille du MELS pour évaluer ses étudiants. Elle dit la trouver pratique, simple. Elle dit aimer son aspect global et la trouve facile à utiliser.

Marie-Ève nous a dit s'appuyer sur quatre critères principaux pour évaluer ses étudiants : le vocabulaire, les verbes, la prononciation et le débit. Pour elle, le vocabulaire et les verbes sont les deux critères qui ont le plus d'impact sur le sens de la phrase : une bonne maîtrise de ces deux aspects permet de se faire comprendre. En outre, selon cette enseignante, une bonne prononciation est garante d'une production réussie. Elle dit : «*Quand c'est trop mal prononcé, on ne comprend pas.*». Cependant, avoir un accent, en autant que cet accent ne nuise pas à la compréhension, n'est pas négatif selon elle. C'est quand l'accent fait qu'elle ne comprend pas ce que les étudiants disent qu'il y a un problème. Selon Marie-Ève, l'important, c'est «*de se comprendre*». Aussi, cette enseignante est sensible à la qualité des liaisons. Elle croit qu'un étudiant qui connaît bien son vocabulaire et sa conjugaison va être en mesure de le démontrer par des liaisons adéquates. C'est pourquoi elle dit, en parlant de ses critères, que «*tout est relié quelque part, il y a toujours un lien, c'est difficile de découper. C'est pas séparé.*» Comme Clémence, Marie-Ève fait une distinction entre la phonétique et le débit, son dernier critère. Celui-ci est jugé moins important que les trois premiers critères. Elle le met «*dans une classe à part*» parce que c'est celui qui a le moins d'incidence sur le sens de la phrase.

La présentation des critères d'évaluation des enseignantes et des résultats de l'évaluation nous permet de constater d'une part que les différentes composantes de la

compétence communicative ont une valeur qui varie selon les évaluateuses et, d'autre part, qu'elles ne remarquent pas toutes les mêmes aspects du discours des étudiants. Dans la partie suivante, l'analyse des données, nous tenterons de mettre en évidence les éléments de la compétence de communication les plus significatifs dans l'évaluation.

## CHAPITRE V

### INTERPRÉTATION DES RÉSULTATS

Dans ce chapitre, nous interpréterons les résultats que nous avons présentés dans le chapitre précédent. Nous discuterons des résultats obtenus lors de la première partie de nos entrevues, c'est-à-dire la partie qui portait sur l'évaluation des jeux de rôle, en les reliant aux commentaires recueillis lors de la deuxième partie de l'entrevue, qui portait sur les critères personnels d'évaluation des participantes. Il sera intéressant de comparer les résultats des deux catégories (ce que les participantes font et ce qu'elles disent qu'elles font), afin de voir s'ils sont concordants. À la fin de cette étape, nous pourrons faire ressortir les composantes de la compétence communicative qui ont le plus d'incidence sur l'évaluation.

Nous analyserons ensuite les commentaires des enseignantes au sujet de leurs pratiques évaluatives, pour voir quelle importance elles accordent à l'examen oral de fin de session de leurs étudiants.

Nous avons vu dans la section précédente que les notes accordées aux étudiants variaient beaucoup. Mais malgré les jugements divergents des cinq participantes, l'analyse des commentaires des participantes formulés au sujet de chacune des composantes de la compétence communicative permet de dégager des constantes : certains aspects semblent avoir une grande influence sur le résultat des examens oraux. Dans la première partie de ce

chapitre, nous tenterons de mettre ces aspects en évidence. Nous relèverons, dans chaque composante (la composante langagière, la compétence stratégique, le savoir textuel, la compétence interactive et les critères non-linguistiques) ce qui contribue à faire augmenter ou diminuer la note des étudiants. De plus, nous tenterons d'expliquer les commentaires de nos participantes en les reliant aux renseignements démographiques recensés par notre questionnaire.

## **5.1 Composante langagière**

Selon le modèle que nous avons choisi, celui de Fulcher (2003), la première composante de la compétence communicative est la composante langagière. Elle se divise en trois sous-composantes : la phonologie, la précision et l'aisance. Dans la section suivante, nous interpréterons les commentaires des cinq enseignantes au sujet de ces trois sous-composante, pour faire ressortir ceux qui semblent influencer la note de façon significative.

### **5.1.1 La phonologie**

Comme nous l'avons vu dans le chapitre II, le cadre théorique, la sous-composante appelée «phonologie» regroupe la prononciation des phonèmes (ou phonétique), les liaisons, l'intonation et le volume (Fulcher, 2003). Dans la littérature, il apparaît que la phonologie est la sous-composante jugée avec le plus de sévérité par les locuteurs natifs (Brown, 1995). Aussi, selon Pollitt et Muray (1996) et Fulcher (2003), la phonologie serait le grand critère qui ferait la différence entre un étudiant de niveau débutant, intermédiaire et avancé. Nos résultats semblent confirmer ce que nous avons trouvé dans la littérature puisque les étudiants qui obtiennent les meilleures notes sont ceux qui reçoivent les meilleurs commentaires en regard de cette composante. Toutefois, les enseignantes ne jugent pas avec la même sévérité les différents accents. En outre, elles semblent accorder beaucoup d'importance aux liaisons et un aspect comme le volume de la voix peut jouer un rôle majeur dans le résultat des étudiants.

### 5.1.1.1 La prononciation des phonèmes

La façon de prononcer les phonèmes a une influence majeure sur le résultat des étudiants, puisque ceux dont l'accent semble le plus proche de la langue cible, l'étudiante 2 et l'étudiant 3, obtiennent les meilleures notes. Les cinq enseignantes nous ont dit qu'elles appréciaient la façon dont l'étudiante 2 et l'étudiant 3 prononcent les phonèmes. Selon Julie, l'accent de l'étudiante 2 est presque imperceptible puisqu'elle dit: *«Je ne me suis même pas rendu compte c'était quoi sa langue maternelle»* et il est perçu comme étant *«agréable»* (Élaine), voire *«plus québécois»* (Julie) en ce qui concerne l'étudiant 3. L'utilisation de l'adjectif «québécois» par cette enseignante pour définir l'accent de l'étudiant 3 semble dénoter, pour Julie, son opinion très favorable de cet étudiant, puisqu'elle lui accorde sa meilleure note. Cela veut dire que pour obtenir une bonne note à l'examen, l'étudiant doit chercher à reproduire la prononciation de la langue cible, ce qui semble naturel et qui est confirmé par la littérature.

Au sujet de l'étudiant 3, Élaine fait la remarque suivante : *«C'est un latino, mais ses liaisons, il les fait avec le «z»: «vous avez», c'est pas «vous savez»* (l'enseignante nous a dit que les hispanophones avaient tendance à assourdir cette consonne). Par cette remarque, Élaine montre qu'elle sait que la langue maternelle a une grande influence sur la phonologie et que l'étudiant 3 est d'autant plus méritant de s'être «débarrassé» d'une habitude fortement ancrée en lui. Cela peut signifier que si un étudiant arrive à corriger des erreurs de prononciation typiques de sa langue maternelle, la perception qu'en ont certains évaluateurs s'en trouve d'autant plus favorable, même s'il reste un accent (l'enseignante perçoit que la langue maternelle de l'étudiant 3 est l'espagnol).

Toutefois, la perception de l'accent n'est pas la même chez toutes les participantes. Nous le voyons en comparant les commentaires formulés par Élaine, Clémence et Marie-Ève, au sujet de l'étudiant 1, un hispanophone, et de l'étudiante 4, une anglophone. Élaine relève la tendance de l'étudiant 1, qui est Mexicain, à prononcer des consonnes finales normalement muettes, ou encore à mal prononcer des sons avec lesquels *«les latinos ont de la difficulté»*. Au sujet de cet étudiant, Clémence dit : *«Plus faible. Moins bonne phonétique»*.

*[que sa partenaire, l'étudiante 2]. [...] Le gars, la phonétique, c'était moins bon.»* Ces commentaires se traduisent par une note plutôt sévère accordée à l'étudiant 1 par Éline (sa deuxième note la plus sévère, soit 80%) et Clémence (sa note la plus sévère, 68%). À l'opposé, Marie-Ève juge que les erreurs de prononciation de l'étudiant 1 sont sans gravité et qu'elles ne nuisent pas à sa compréhension. Elle dit : *«Il a son accent latino, mais ça n'enlève pas ta compréhension.»* Elle lui accorde d'ailleurs sa deuxième meilleure note, 87%). Nous voyons donc que la perception de l'accent n'est pas la même pour les trois évaluatrices.

Mais ce n'est pas tout. La différence s'accentue encore entre Marie-Ève, Clémence et Éline dans leur perception de l'accent de l'étudiante 4. Au sujet de cette étudiante, Marie-Ève dit :

*Elle prononce les lettres finales [l'enseignante remarque que l'étudiante prononce le «t» de «moment»] [...] Ben beau pas prononcer, le français, quand tu prononces trop de lettres, ça change. [...] Elle, ses fins de mots : «je préféré» [au lieu de je préfère], ça passe pas.*

Il est intéressant de constater que pour Marie-Ève, les erreurs de phonétique de l'étudiant 1, qui avait lui aussi tendance selon Éline à prononcer les consonnes finales muettes et à prononcer le «e» final comme un «é» n'ont pas la même gravité que pour l'étudiante 4 : elle est très sévère au sujet de la prononciation de l'étudiante 4, (elle lui accorde d'ailleurs sa pire note, 53 %) et plutôt clémente avec l'étudiant 1. Pourquoi? Nous pourrions trouver une explication dans les renseignements démographiques concernant les participantes. Contrairement à Éline et Clémence, Marie-Ève parle espagnol. De plus, son conjoint est hispanophone. Il se peut que cette enseignante soit habituée à l'accent mexicain de l'étudiant 1, qu'il entrave moins sa compréhension et qu'elle le juge avec moins de sévérité. Cela peut signifier que pour un évaluateur, la connaissance de la langue maternelle d'un étudiant peut se traduire par une meilleure perception de son accent et en bout de ligne, une meilleure évaluation.

### 5.1.1.2 Les liaisons

Mais en plus de leur accent, comment se distinguent les deux meilleurs étudiants? Il semble que le respect des règles phonologiques propres au français contribue à donner une bonne impression aux enseignantes : ces deux étudiants, et particulièrement l'étudiante 2, sont reconnus par Julie et Éline pour bien faire la liaison entre les mots : «*Elle faisait les liaisons entre les mots.*» (Julie); «*J'ai tout de suite remarqué ses bonnes liaisons.*» (Éline). Pourquoi cet aspect a-t-il autant d'importance? Une explication serait que le respect des règles phonologiques, qui peut se traduire entre autres par l'enchaînement entre les mots, reflète la connaissance des règles propres à la morphologie flexionnelle. Parfois, la limite entre la phonologie et la précision grammaticale est poreuse : par sa façon de prononcer les mots, l'étudiant peut révéler ses failles sur le plan grammatical et à l'opposé, par sa prononciation adéquate, il peut montrer ses forces sur le plan grammatical. La phonologie refléterait la connaissance grammaticale. D'ailleurs, lors de la deuxième partie de notre entrevue, qui portait sur les critères de correction personnels des enseignantes, le critère «liaison» a été invoqué par quatre enseignantes sur cinq : Julie, Éline, Clémence et Marie-Ève. Selon cette dernière, la connaissance de l'orthographe lexicale et grammaticale se démontre par des liaisons adéquates. C'est pourquoi les liaisons semblent donc déterminantes sur la note de l'étudiant. Outre les liaisons, il reste un dernier aspect de la sous-composante, le volume, qui au moins pour une enseignante revêt beaucoup d'importance. Nous le verrons dans la section suivante.

### 5.1.1.3 L'intensité vocale

L'intensité vocale (appelé «volume» par Julie et Éline) se définit par la force et la portée de la voix (Le Huche et Allali, 2001). C'est un aspect de la phonologie qui peut paraître anodin, parce qu'il n'est généralement pas relié à une habileté cognitive, mais plutôt à un trait de personnalité. Julie en parle en ces termes :

*Moi, je n'évalue pas du tout le volume. Chaque personnalité est différente. Il y a des timides qui n'aiment pas parler fort. Je vais essayer de me rapprocher d'eux. Si dans sa langue maternelle il est comme ça, tu peux pas...*

Pourtant, notre expérimentation démontre que pour au moins une enseignante sur cinq (Élaine), il peut jouer un rôle crucial sur la note. Nous le voyons dans l'évaluation de l'étudiant 3 et de l'étudiante 4. Selon la grille analytique d'Élaine, cette dernière obtient seulement un point de moins que l'étudiant 3 dans la partie phonologie, uniquement parce qu'elle parle fort et avec intonation : *«Elle parle plus fort que monsieur (l'étudiant 3) et on la comprend mieux.»* D'ailleurs, la seule remarque négative au sujet de la phonologie de l'étudiant 3 vient d'Élaine :

*Lui, il parle juste pas assez fort. Il prononce bien, fait de bonnes liaisons, mais c'est pas assez fort. Il dit beaucoup de choses, son rythme est bon. Il parle vraiment pas assez fort, il n'a pas de volume, c'est la seule chose. Il faut vraiment tendre l'oreille pour le comprendre.*

Pour Élaine, le fait de ne pas parler fort est vraiment problématique puisque l'étudiant 3, au sujet duquel elle fait, au total (pour toutes les composantes), plus de remarques positives et moins de remarques négatives qu'au sujet de l'étudiant 1, s'en tire avec la même note que lui, soit 80%. Elle justifie sa note en disant : *«Quand tu prononces pas bien, tu parles pas assez fort, on comprend mal. Les gens se désintéressent et les gens ne te comprennent pas.»*

Donc, pour cette enseignante, la faible intensité vocale aussi préjudiciable que les mauvaises liaisons ou l'accent étranger marqué. Mais c'est la seule des cinq enseignantes qui soulève ce critère. Toutefois, il n'est pas impossible que la faible intensité vocale de l'étudiant 3 n'ait pas eu d'influence négative dans la perception des autres enseignantes.

#### **5.1.1.4 Synthèse de la phonologie**

Finalement, nos conclusions au sujet de la phonologie rejoignent en grande partie ce que Brown (1995), Pollitt et Muray (1996) et Fulcher (2003) rapportent, soit que cette sous-composante a une grande importance pour les évaluateurs. Les liaisons et l'accent sont

déterminants sur la note et un critère comme l'intensité vocale peut être significatif pour certains enseignants. Somme toute, les deux étudiants qui obtiennent la note globale la plus basse sont ceux qui ont eu le plus de commentaires négatifs au sujet de la phonologie. Moins l'accent étranger était perceptible, plus la note accordée par les évaluateuses augmentait. D'ailleurs, les cinq enseignantes ont mentionné, en deuxième partie d'entrevue, que la phonologie, sans être le critère le plus déterminant, était important. Leur évaluation semble être conforme à ce qu'elles ont dit à ce sujet.

### **5.1.2 La précision**

La précision, qui regroupe selon la typologie de Fulcher (2003) le vocabulaire, la grammaire, la conjugaison et la syntaxe, est une sous-composante qui joue un grand rôle sur le résultat de l'évaluation de la production orale. Notre recension des écrits a révélé que la précision, notamment en ce qui concerne les erreurs affectant ses différents aspects, est le critère auquel les enseignants accordent le plus d'attention (Brown, 1995; Fulcher, 2003). Notre expérimentation semble le confirmer. Parmi les sept critères invoqués par les enseignantes en deuxième partie d'entrevue, trois relèvent de la précision : la syntaxe, le vocabulaire et la conjugaison. Il n'est pas surprenant que les deux étudiants qui obtiennent les meilleures moyennes se distinguent en regard de cette sous-composante : d'abord, par le très faible nombre d'erreurs détectées, notamment chez l'étudiante 2, puis, par l'utilisation répétée de structures recherchées, surtout par l'étudiant 3. Nous interpréterons dans la section suivante les commentaires des participantes au sujet des aspects de cette sous-composante qui semblent déterminants sur le score final, soit le nombre et la gravité des erreurs, la conjugaison et la syntaxe.

### 5.2.2.1 Le nombre et la gravité des erreurs

Selon les enseignantes, c'est l'étudiante 2 qui fait le moins d'erreurs (mais c'est aussi celle qui parle le moins). Au total, les cinq enseignantes ensemble ne détectent que quatre erreurs dans son discours. Julie dit que *«elle a vraiment de bonnes structures en français»*. Même constat chez Phuong : *«Sa structure est très solide»*. L'étudiant 3, lui non plus, ne commet pas beaucoup d'erreurs (les enseignantes en dénombrent six). Ce sont aussi les deux étudiants qui se méritent les meilleures notes, ce qui nous porte à croire que le nombre d'erreurs détectées joue un rôle déterminant dans l'évaluation. À l'opposé, l'étudiant 1 commettrait, selon les enseignantes, 17 erreurs (quatre fois plus que l'étudiante 2 et trois fois plus que l'étudiant 3) et l'étudiante 4 en commettrait 18. La nature de ces erreurs est variée : conjugaison, syntaxe, accord en genre et impropriétés lexicales. Toutefois, même si le nombre d'erreurs commises a sûrement une influence sur la note, ce qui semble encore plus déterminant est la gravité des erreurs, gravité par ailleurs bien relative : nous le constatons en comparant les commentaires de l'étudiant 1 et de l'étudiante 4 formulés par Clémence et Marie-Ève. Pour Marie-Ève, même si l'étudiant 1 commet des erreurs, il mérite une bonne note :

*[...] Je leur [à l'étudiants 1 et à l'étudiante 2] ai donné une bonne note, selon moi, c'est très bon, 13/15 pour l'homme parce que oui, il a une aisance, oui, il fait certaines erreurs, mais c'est des petits détails. [...] Ce sont des petites choses qui selon moi, ne sont pas importantes pour le niveau 5.*

Toujours selon Marie-Ève, l'étudiante 4, pour sa part, ne mérite pas de passer : *«Selon moi, la femme passe pas. Ses structures sont pas bonnes, ses verbes sont pas conjugués, ... ça passe pas»*.

À l'opposé, pour Clémence, l'étudiante 4 mérite une meilleure note que l'étudiant 1. Même si elle remarque que l'étudiante 4 faisait beaucoup d'erreurs : *«Elle faisait des erreurs pas mal, moi je trouvais, mais des erreurs d'anglophone»*, pour elle, l'étudiant 1 est plus faible parce qu'il fait des erreurs *«qui devraient être corrigées au niveau 1 : « La prochaine semaine». Des affaires que tu vois tout de suite. C'est dans le niveau 1»*. Nous relevons donc

une totale contradiction entre le commentaire de Clémence et celui de Marie-Ève, puisque pour Marie-Ève, les erreurs de l'étudiant 1 sont «*des petits détails*», des «*petites erreurs*» et pour Clémence, elles sont inacceptables, même au niveau 1. Leur perception de la gravité des erreurs de l'étudiante 4 est aussi en contradiction : pour Marie-Ève, ses erreurs sont inacceptables et pour Clémence, ce sont des erreurs qu'une anglophone fait normalement, ce qu'elle semble excuser. Nous pourrions expliquer le fait que Marie-Ève juge les erreurs de l'étudiant 1 sans trop de sévérité (moins de sévérité que Clémence) pour les mêmes raisons qu'elle trouvait son accent acceptable : comme cette enseignante parle espagnol et vit avec un hispanophone, il est possible qu'elle ait développé plus d'indulgence à l'égard des erreurs commises par ce groupe linguistique. Par contre, il n'y a rien dans les renseignements que nous avons au sujet de Clémence qui expliquerait pourquoi elle semble juger les erreurs typiquement anglophones avec moins de sévérité. Cette enseignante parle anglais, mais c'est aussi le cas de Marie-Ève. Aucune des deux n'utilise l'anglais à la maison.

Au sujet de la gravité des erreurs, nous pouvons faire un autre constat intéressant. Élane et Phuong, les deux enseignantes les plus généreuses (ce sont elles qui ont accordé les meilleures notes globalement) sont celles qui ont le plus d'expérience avec le niveau 5. Ces deux enseignantes expérimentées se font probablement moins d'illusions que les trois autres au sujet de la précision. Elles savent peut-être mieux que les autres ce à quoi il est légitime de s'attendre de la part d'un étudiant de niveau 5. Clémence, qui a enseigné presque exclusivement à des débutants, se montre surprise quand elle entend une erreur que ses étudiants de niveau 2 font. Julie et Marie-Ève aussi ont tendance à qualifier certaines erreurs d'erreurs de débutants. Pour ces enseignantes, moins expérimentées avec le niveau 5, il peut paraître aberrant d'entendre des erreurs qu'un étudiant moins avancé commet lui aussi. Mais il s'avère que certaines erreurs «de débutant» peuvent se fossiliser et se retrouver dans le discours d'un locuteur de niveau très avancé (Lennon, 1991). Cela ne veut pas dire que le locuteur a stagné à un niveau débutant.

Outre son expérience, il y a peut-être une autre explication à la mansuétude de Phuong. Nous avons vu dans le chapitre III, «Présentation des résultats», que cette dernière est celle pour qui les faiblesses concernant la précision linguistique ont moins de gravité.

Cela est particulièrement évident dans sa façon d'évaluer l'étudiante 4. Cette enseignante lui accorde la même note globale qu'à l'étudiant 1, même si ce dernier était perçu comme étant supérieur sur le plan de la précision, puisque les enseignantes détectent chez lui moins d'erreurs affectant la précision. Peut-être cela s'explique-t-il par le fait que c'est la seule des cinq participantes qui n'a pas le français comme langue maternelle (sa langue maternelle est le vietnamien). Bien que cette enseignante, quand nous l'avons interrogée, parlait un français tout à fait irréprochable (voire légèrement teinté d'un accent français), il est possible qu'elle soit plus indulgente que les autres au sujet de la précision, ayant peut-être eu à faire des efforts plus grands que ses collègues pour maîtriser le français. Toutefois, la littérature ne confirme pas totalement cette hypothèse : dans les recherches que nous avons lues, nous notons parfois plus d'indulgence et parfois plus de sévérité de la part d'un évaluateur dont la langue maternelle est différente de celle qu'il évalue (Orr, 2000; Pollit et Murray, 1996; Brown, 1995).

### 5.2.2.2 La conjugaison

Un autre aspect de la précision invoqué par les enseignantes comme critère déterminant lors de l'évaluation est la conjugaison. Or, en plus de se distinguer par le nombre réduit d'erreurs qu'il commet, l'étudiant 3 impressionne Julie, Éline et Marie-Ève par son utilisation répétée de l'imparfait et du passé composé (la conjugaison). Même Clémence, qui dit ne pas connaître le programme du niveau 5, estime que l'étudiant 3 mérite une meilleure note que l'étudiant 1 (avec lequel elle le compare) parce que «*il est allé dans l'imparfait*». L'utilisation de temps de verbes variés semble donc être un gage de réussite à l'examen oral.

Les enseignantes sont donc sensibles à la conjugaison, mais la perçoivent de manière très différente dans le discours des étudiants. Nous voyons cette différence en comparant le commentaire d'Éline avec celui de Marie-Ève, au sujet de l'étudiante 4. Éline soutient :

*Le passé composé reste stable. Elle le connaît bien, elle l'utilise. Elle dit : «j'étais», l'imparfait, le conditionnel aussi. L'imparfait elle l'a bien aussi : «Quand j'étais*

*jeune», elle l'a dit à deux reprises, c'est bien. [...] Elle aurait 70%, elle le mérite son 70% parce qu'elle a bien son passé composé.*

À l'opposé, Marie-Ève dit: «*Ses verbes sont pas conjugués*». En réalité, l'étudiante 4, même si elle faisait plusieurs erreurs de conjugaison, montrait à certains moments, qu'elle était quand même capable de conjuguer adéquatement ses verbes. Il semble que pour Marie-Ève, la correction se fasse de façon négative (les erreurs s'accumulent) et que pour Élane, une bonne utilisation compense pour une erreur. Cela démontre une fois de plus la subjectivité de l'évaluation : même un critère somme toute assez clair comme la conjugaison, dont les enseignantes disent qu'il est très important, n'est pas repéré de façon univoque dans le discours des étudiants. Cela appuie une fois de plus les résultats de Orr (2002), qui révélaient que les évaluateurs ne portent pas tous attention aux mêmes parties du discours.

### 5.2.2.3 La syntaxe

D'autres remarques de Phuong et de Julie illustrent bien les contradictions que nous relevons dans les commentaires des enseignantes, plus particulièrement au sujet de la syntaxe, notamment en ce qui concerne la complexité des phrases. Pour Phuong, l'étudiant 3 s'exprimait avec «*des phrases toutes simples*» et pour Julie, ce même étudiant était supérieur aux autres parce qu'il s'exprimait avec des phrases complexes :

*Je trouvais qu'il utilisait beaucoup le pronom relatif avec de belles structures de phrases, des phrases complexes : «J'ai besoin de quelqu'un qui...» Des phrases avec plusieurs verbes bien utilisés.*

Il est curieux que les deux enseignantes invoquent des raisons contraires pour expliquer la bonne impression que l'étudiant 3 leur laisse. Peut-être que pour Phuong, l'impression de simplicité que lui donne l'étudiant 3 vient du fait qu'il commet peu d'erreurs affectant la précision. Pour cette enseignante, la simplicité se traduirait par l'absence d'erreurs.

#### 5.2.2.4 Synthèse de la sous-composante de la précision

Notre expérimentation révèle donc que les enseignantes sont sensibles à la précision linguistique puisque généralement, elles détectent les erreurs des étudiants, mais ce qui frappe, c'est la grande différence d'importance qu'elles leur accordent. Marie-Ève semble être celle pour qui l'erreur a le plus d'importance, mais il est frappant de voir que les erreurs des étudiants 1 et 3 (les deux hispanophones) sont jugées moins graves que celles de l'étudiante 4. Comme nous l'avons déjà mentionné, peut-être que cette enseignante, dont le conjoint est hispanophone, accorde moins d'importance aux erreurs typiques de ce groupe linguistique. À ce sujet, il est intéressant de noter que Julie, dont le conjoint aussi est hispanophone, et Marie-Ève ont accordé leur note la plus élevée aux deux étudiants mexicains. Il est possible que le fait de parler la langue des étudiants évalués ait rendu les participantes moins sensibles à leurs lacunes. Il est aussi possible que ces étudiants aient attiré la sympathie de ces enseignantes. Inconsciemment, Julie et Marie-Ève ont peut-être favorisé les étudiants qui s'expriment d'une façon similaire à leur conjoint. Toutefois, notre recherche ne peut pas le confirmer : il faudrait concevoir une expérimentation qui contrôlerait mieux ces variables pour le conclure.

De plus, une enseignante comme Élane, qui a tendance à noter toutes les erreurs de façon systématique, mais qui note aussi ce qu'elle appelle les «bons coups», comme l'utilisation de temps de verbes variés, a tendance à être plus indulgente dans sa correction, comme si un «bon coup» rachetait une erreur. Cela est perceptible dans son évaluation de l'étudiante 4. Grâce à son utilisation des temps du passé, elle a «annulé» des erreurs. À l'opposé, Marie-Ève, apparaît moins sensible aux «bons coups», comme s'ils étaient effacés par les erreurs. Pour Clémence, la justification semble plus impressionniste, puisqu'elle ne note pas les erreurs des étudiants. Cela s'explique par le fait qu'elle dit peu connaître le programme du niveau 5 : elle n'a pas de repères précis sur lesquels appuyer son jugement. Elle confesse d'ailleurs : «*J'ai pas les objectifs du 5 devant moi, c'est pour ça l'histoire.*» En l'absence de tels repères, elle fixe ses notes de façon un peu arbitraire et les justifie en invoquant d'autres critères que la précision. Cette absence de repères explique peut-être aussi

l'uniformité des notes accordées par Clémence (68%, 70%, 70% et 72%). Ne sachant pas précisément ce qu'elle devait évaluer, elle a eu tendance à percevoir les jeux de rôle de façon un peu linéaire.

Nous venons de voir que la précision est une sous-composante importante dans l'évaluation. Dans la section suivante, nous verrons que l'aisance, la dernière sous-composante de la composante langagière, joue aussi un rôle crucial.

### 5.1.3 L'aisance

L'aisance, comme nous l'avons vu dans le «Cadre théorique», le chapitre II, concerne le rythme, la fluidité de la parole et la quantité d'information véhiculée. Nous avons vu aussi que la précision a une influence sur l'aisance : le manque de précision affecte généralement l'aisance (Germain et Netten, 2002; Chambers, 1997; Kromos et Dénes, 2004). Il semblerait logique que les étudiants qui obtiennent les meilleurs commentaires au sujet de la précision obtiennent aussi les meilleurs commentaires au sujet de l'aisance, mais nos résultats ne le confirment pas tout à fait. Dans les sections suivantes, nous analyserons les commentaires formulés au sujet du rythme et de la fluidité et ensuite, au sujet de la quantité d'information.

#### 5.1.3.1 Le rythme et la fluidité

Contrairement aux résultats d'une recherche de Kromos et Dénes (2004), la perception des enseignantes au sujet de la fluidité du message n'est pas unanime, du moins en ce qui concerne l'étudiante 2 et l'étudiante 4. En effet, l'étudiante 4, est perçue par Julie et Éléine comme ayant le moins d'aisance à communiquer. Julie dit : *«C'était très détaché»* et Éléine : *«C'est pas enchaîné, ça manque de fluidité». Elle avait beaucoup d'hésitations»*. Mais pour sa part, Marie-Ève, qui s'est montrée très dure avec l'étudiante 4 au sujet de la phonologie et de la précision, a été plus clémente en ce qui concerne son aisance. Elle

dit: «*C'est pas qu'elle parle lentement, son débit est bon*». Nous voyons que les commentaires au sujet de l'aisance de l'étudiante 4 sont contradictoires : Julie et Élane disent qu'elle hésite et Marie-Ève prétend que son débit est bon. Comment expliquer ces différences de perception? Peut-être que Marie-Ève, qui avait fait surtout des remarques négatives au sujet de la performance de cette étudiante voulait finalement relever quelque chose de positif. Même contradiction entre Marie-Ève et Julie au sujet de l'aisance de l'étudiante 2. Julie dit à son sujet : «*Ça coulait bien*» et Marie-Ève, pour sa part, remarque des «*hésitations*» et qu'elle «*cherche ses mots*». Ces commentaires contradictoires donnent encore une fois l'impression que les enseignantes n'ont pas retenu les mêmes parties dans le discours de l'étudiante 2. Au début du premier jeu de rôle, l'étudiante 2 semblait, selon Marie-Ève, hésitante et un peu timide. Il est possible que cette enseignante ait retenu surtout cette partie du discours de l'étudiante. Cela serait cohérent avec ce que nous avons vu plus tôt au sujet de la façon dont Marie-Ève avait tendance à retenir surtout les erreurs des étudiants, et corriger de façon négative : pour cette enseignante, les erreurs ne s'effacent pas. La mauvaise impression laissée par un début de jeu de rôle hésitant aurait perduré dans l'esprit de Marie-Ève et aurait empêché l'étudiante 2 de faire bonne figure, même si pendant le reste du jeu de rôle, elle parlait avec plus d'aisance. Cette divergence pourrait aussi s'expliquer par le fait que l'attention des évaluateurs varie et qu'ils ne se souviennent pas tous des mêmes détails.

### 5.1.3.2 La quantité d'information

Nous avons vu que l'aisance ne concerne pas seulement le rythme du discours. Cette sous-composante englobe aussi la quantité d'information que le locuteur arrive à véhiculer. Habituellement, plus l'étudiant arrive à véhiculer d'informations, meilleure sera la perception des évaluateurs à son sujet (Fulcher, 2003). Mais notre expérimentation révèle que ce n'est pas toujours le cas. En effet, il est intéressant d'analyser les commentaires émis au sujet de l'étudiant 1 et de sa partenaire, l'étudiante 2. Pour Julie l'étudiant 1, même s'il démontre moins de précision et une moins bonne phonétique que l'étudiante 2, mérite une note aussi bonne que la sienne parce qu'il parle plus qu'elle. Même attitude de la part de Marie-Ève :

elle accorde beaucoup d'importance à la quantité d'information. L'étudiant 1, qui a parlé plus que sa partenaire, obtient une meilleure note qu'elle. Elle justifie ainsi :

*Pourquoi la fille (l'étudiante 2) a plus bas? Je trouve ça un peu imbécile à dire, c'est que son deuxième examen, elle parle pas assez. Je peux pas l'évaluer, c'est toujours l'homme qui parle.*

Ces deux enseignantes sont donc sensibles à la quantité d'informations transmises par les étudiants. À l'inverse, pour Éline et Phuong, le fait que l'étudiante 2 ne parle pas beaucoup dans le deuxième jeu de rôle ne dérange pas. Éline dit :

*Elle n'a pas parlé beaucoup dans le deuxième jeu de rôle. [...] Elle, je ne l'ai pas assez entendu. J'ai pas assez d'échantillonnage mais, ce qu'elle a dit, c'était bien. [...] Définitivement, elle est meilleure que lui, même si j'ai pas entendu plus.*

Éline, au lieu de pénaliser l'étudiante, se fie sur ce qu'elle a entendu dans le premier jeu de rôle et extrapole sa bonne note du premier jeu de rôle à celle du deuxième. Phuong non plus ne pénalise pas l'étudiante 2, qui s'exprime moins longuement, bien au contraire. L'apparente indulgence de Phuong et d'Éline peut s'expliquer par leur expérience. Elles repèrent avec peut-être plus d'acuité que Julie et Marie-Ève, les deux enseignantes les moins expérimentées, les forces de cette étudiante et elles se permettent d'extrapoler leurs observations. Julie et Marie-Ève, peut-être moins sûres d'elles, se sont fiées au critère de la quantité pour juger de l'habileté de l'étudiant 1 à s'exprimer. Ce critère est rapporté par Kronos et Dénes (2004) comme étant plus facilement identifiable. Il serait donc un point de repère «rassurant», en quelque sorte, du moins pour les évaluateuses moins expérimentées. Au sujet de l'évaluation de l'étudiante 2, il est intéressant de se rappeler que Lisa, peu après avoir commenté la performance de cette étudiante, nous avait confié qu'elle augmenterait peut-être sa note, en y pensant bien. Cette jeune enseignante, comme les deux plus expérimentées, aurait donc flairé elle aussi la plus grande compétence communicative de l'étudiante 2.

Une autre observation au sujet de l'aisance qui semble en contradiction avec ce que nous avons trouvé dans la littérature, est que pour Phuong, non seulement la plus grande quantité d'information véhiculée par l'étudiant 1 ne s'est pas traduite par une note plus élevée

que sa partenaire, mais au contraire : le fait qu'il ait parlé beaucoup lui a nuï. Phuong dit de lui :

*On sent un côté trop forcé. Il veut en mettre, on le sent. [...] On sent qu'il voulait montrer qu'il avait de bonnes structures, mais au contraire, c'est là qu'il avait la faille.*

Nous pourrions croire, par cette remarque, que pour Phuong, l'étudiant 1 a voulu camoufler ses faiblesses en parlant beaucoup. De plus, à cause de cette attitude, il a donné l'impression qu'il voulait dominer sa partenaire: «*Lui, il a l'air de celui qui contrôle*». Aux yeux de Phuong, c'est donc à cause de lui que l'étudiante 2 n'a pas pu parler. D'une certaine façon, elle explique le manque d'aisance de l'étudiante 2 par le comportement de son partenaire. Cette sévère critique du comportement de l'étudiant 1 peut s'expliquer par les critères retenus par Phuong en deuxième partie d'entrevue. Pour cette enseignante, l'important lors d'un jeu de rôle est de montrer son intention de communiquer, d'échanger. L'étudiant 1 aurait failli à ce critère, en voulant parler plus que sa partenaire, ce qui s'est traduit par un jugement sévère de la part de Phuong.

### 5.1.3.3 Synthèse de la sous-composante de l'aisance

Bref, le débit semble être un critère important lors de l'évaluation, puisque l'étudiante 4, celle qui obtient la note la plus basse, est perçue par quatre enseignantes sur cinq comme celle dont cet aspect est le moins bien maîtrisé parmi les quatre étudiants. En outre, en ce qui concerne la quantité d'information, pour Julie et Marie-Ève, il semble préférable de s'exprimer plus librement, quitte à commettre plus d'erreurs, que de ne pas s'exprimer, puisqu'elles pénalisent l'étudiante 2 sous prétexte qu'elle parle moins que son partenaire. Mais pour Élane et Phuong, vaut mieux parler moins et «mieux». Clémence, pour sa part, ne fait pas de commentaire au sujet de l'aisance, sauf une courte remarque, au sujet de l'étudiante 2, qu'elle trouve «*plus dégagée*» que son partenaire. Ces trois enseignantes, les plus expérimentées (même Clémence, peu expérimentée avec le niveau 5, mais qui enseigne depuis 19 ans), semblent donc moins influencées que les deux plus jeunes par la quantité

d'information. Nous pensons que leur expérience en évaluation leur permet de porter un jugement plus rapidement, ce qui explique qu'elles n'aient pas besoin d'un échantillon très vaste pour évaluer la compétence communicative d'un étudiant.

## 5.2 Capacité stratégique

Nous avons dit dans le chapitre II, le Cadre théorique, que la capacité stratégique est décrite dans la littérature comme une façon de compenser pour les lacunes dans les autres aspects de la compétence communicative et comme l'habileté à planifier son discours en fonction de ses connaissances et des consignes de l'examen (Canale et Swain, 1980; Bachman, 1990; Bachman et Palmer, 1996; Fulcher, 2003). Notre expérimentation tend à démontrer que l'utilisation de telles stratégies peut influencer les évaluatrices. Dans les sections suivantes, nous verrons que les enseignantes ont eu tendance à noter positivement l'utilisation par l'étudiante 4 de stratégies de réussite linguistiques et non linguistiques et que la stratégie de planification, utilisée judicieusement par la deuxième dyade, a été perçue favorablement par certaines enseignantes.

### 5.2.1 Les stratégies de réussite

L'utilisation de stratégies, pour compenser des lacunes dans la composante langagière, peut améliorer la perception des évaluateurs et permettre à un étudiant d'obtenir une bonne note. L'étudiante 4, qui est reconnue unanimement comme celle qui a le plus de faiblesses dans la composante langagière est aussi celle chez qui les enseignantes relèvent le plus d'utilisation de stratégies de réussite. Elle utilise différentes stratégies linguistiques : *«Elle fait un effort pour parler. [...] Elle s'exprime quand même»*. (Julie); *«Elle a le sens de la communication.»* (Phuong) et :

*Mademoiselle s'auto-corrige, ce qui est très bien. Elle a fait une erreur et elle s'est reprise tout de suite, ce qui est très bien. Donc, elle est à un bon niveau, un niveau avancé.* (Élaine).

Elle fait aussi usage des stratégies non linguistiques : *«Avec sa gestuelle, son regard, elle va capter l'attention. Même si ses structures sont déficientes, ça compense.»* (Élaine) À son tour, Clémence remarque : *«Elle utilisait le geste, elle nous avait avec ça».*

Nous voyons que Julie, Élaine, Phuong et Clémence notent positivement l'utilisation de stratégies de réussite, linguistiques ou non, chez cette étudiante. Cela se traduit, chez Clémence et Phuong surtout, par une évaluation somme toute assez bonne des jeux de rôle de cette étudiante. Marie-Ève, pour sa part, ne relève aucune de ces stratégies et c'est aussi la seule qui n'accorde pas la note de passage à cette étudiante. Sa note (53%) est beaucoup plus sévère que celle des quatre autres enseignantes. Cela veut dire que l'utilisation de stratégies ou du moins, la démonstration que l'étudiant fait un effort pour communiquer, peut, pour certaines enseignantes, compenser pour les faiblesses dans la composante langagière et se traduire par une bonne note, finalement.

### **5.2.2 Stratégie de planification du discours**

L'utilisation d'une autre stratégie, rapportée par Bachman et Palmer (1996), celle de la planification du discours en fonction des consignes de l'examen et de ses connaissances, peut aussi se révéler judicieuse. Clémence et Élaine mentionnent la faiblesse de l'étudiant 1 au sujet de l'utilisation de cette stratégie. Elles déplorent qu'il ait parlé du hockey alors qu'il ne connaissait visiblement pas ça. Il se peut que Clémence, qui a donné sa note la plus faible à cet étudiant, ait été très sensible à ce sujet. Elle dit :

*Il aurait pu parler d'une chose qu'il connaît. Il était obligé de faire des préjugés. Il aurait pu parler du soccer, qui est bien populaire.*

Pour Éline, ce mauvais choix se traduit par des lacunes lexicales et un plus grand nombre d'erreurs que dans le premier jeu de rôle, et par une note plus basse. Elle dit :

*Monsieur va parler d'un sport qu'il ne connaît pas très bien. Il essaie de lui expliquer. Il fait plus d'erreurs.*

Selon les directives de la DGFA (1994), l'étudiant 1 n'était pas obligé de parler du hockey. Il aurait pu choisir tout autre sport. Son choix plus ou moins éclairé lui a valu des remarques négatives qui ont eu une influence certaine sur sa note. Sa partenaire, l'étudiante 2, ne connaissait pas non plus ce sport, et cela s'est manifesté par son peu d'engagement dans la conversation. Ce choix lui a donc été néfaste. À l'inverse, la deuxième dyade s'est mise en valeur par ce choix, puisque tous deux savaient visiblement de quoi ils parlaient. Pendant le jeu de rôle, ils ont nommé plusieurs joueurs des Canadiens et des Maple Leafs. L'étudiante 4 a donné l'impression qu'elle avait «*beaucoup d'idées*» l'étudiant 3 a même parlé du temps où il était partisan des Nordiques de Québec, ce qui a fait dire à Julie :

*Je pense que ça fait longtemps qu'il est au Québec. J'ai l'impression. Il va avoir des intonations plus québécoises, une fluidité, des accents plus québécois. Son oreille est plus développée aussi (Julie).*

*Pourquoi tu penses que ça fait longtemps qu'il est ici? (Sara)*

*Il a parlé des Nordiques, alors, je sais que ça fait longtemps qu'il est ici. Ça n'influence pas ma note le fait qu'il connaisse les Nordiques, parce que ce que j'évalue, c'est comment il s'exprime.*

Même si elle affirme que le fait que l'étudiant connaisse les Nordiques n'a pas d'influence sur sa note, il se peut que la perception de Julie au sujet de cet étudiant, duquel elle disait qu'il avait un accent «*plus québécois*», l'ait incité à lui accordé une meilleure note. Peut-être aussi que cette remarque sur la défunte équipe de hockey lui a laissé, inconsciemment, une opinion favorable au sujet de la connaissance de l'étudiant 3 de la société québécoise et a contribué à augmenter son score.

### 5.2.3 Synthèse de la composante stratégique

L'analyse des commentaires des enseignantes nous permet de croire que l'utilisation de stratégies linguistiques et non linguistiques de réussite et de stratégies de planification du discours influence le résultat. L'étudiante 4, perçue comme étant la plus faible sur le plan langagier, a tout de même obtenu une bonne note grâce à sa capacité stratégique. Il semble aussi que par le choix du sujet, la deuxième dyade a réussi à se démarquer et à faire une bonne impression et que l'inverse s'est produite pour les deux premiers étudiants. Cette composante de la compétence communicative a semblé jouer un rôle important dans l'évaluation, au moins pour quatre enseignantes sur cinq. Pour un étudiant, il peut être rassurant de savoir qu'il peut faire une bonne impression, même s'il commet des erreurs affectant la composante langagière. Mais outre ces deux composantes, une troisième semble jouer un rôle important dans l'évaluation. Il s'agit du savoir textuel, dont nous parlerons dans la section suivante.

### 5.3 Savoir textuel

Selon Bachman (1990), le savoir textuel concerne la cohésion (l'organisation du discours) et la cohérence (le lien avec le but du message). Un étudiant démontre ce savoir par l'utilisation de connecteurs logiques ou encore par le respect des règles implicites et explicites de la conversation. Notre expérimentation tend à démontrer que cette composante peut grandement influencer le résultat d'un étudiant.

Il est clair que pour Julie, cette composante est très importante. Cela est perceptible dans son évaluation de l'étudiant 1. Elle lui donne une note plus faible dans son premier jeu de rôle en justifiant : *«J'avais de la difficulté à le comprendre, il n'avait pas... je ne voyais pas son cheminement dans sa tête. Les idées ne se suivaient pas»*. Elle poursuit :

*Pour le deuxième [jeu de rôle], il manquait de vocabulaire, il utilisait souvent «quelque chose», sauf que son discours était plus clair. Il parlait plus, s'exprimait plus, il était plus enchaîné [...] ça se suivait plus. Pour la première situation, il cherchait, changeait de sujet, ça allait d'une chose à l'autre.*

Julie réitère plus loin le fait qu'elle accorde de l'importance à la cohésion lorsqu'elle dit : «*Il faut que les étudiants fassent des liens pendant leur interaction: il a dit ça, alors je vais lui poser ça comme question, qui est en lien avec ça, etc.*». Nous voyons que pour Julie, le savoir textuel peut être encore plus important que la composante langagière, puisque même si l'étudiant maîtrisait moins bien le vocabulaire, lors du deuxième jeu de rôle, elle le trouvait plus compréhensible parce qu'elle voyait où il allait. Rappelons-nous que c'était exactement le contraire pour Éline, qui avait trouvé le premier jeu de rôle de cet étudiant supérieur au deuxième, justement parce qu'il maîtrisait mieux le vocabulaire et qu'il faisait moins d'erreurs. C'est donc dire que la cohésion n'a pas la même importance pour toutes les enseignantes.

Clémence remarque aussi les lacunes de l'étudiant 1 au sujet de la cohérence lors du premier jeu de rôle. Elle dit au sujet de cet étudiant :

*Elle [l'étudiante 2] lui pose une question et il répond autre chose. Ça ne concordait pas toujours. [...] Je trouve qu'il était moins solide qu'elle. Elle a plus cerné, elle posait les questions, elle dirigeait plus. Pourtant, c'est lui qui l'engageait.*

Le commentaire de Clémence fait ressortir le manque de cohésion et de cohérence du message de l'étudiant 1 dans son premier jeu de rôle et il met en lumière les forces de l'étudiante 2. C'est elle qui semble maîtriser le mieux les règles de la conversation, puisqu'elle pose les questions et dirige plus l'entretien. Au contraire, l'étudiant 1 ne s'est pas conformé aux règles que son rôle lui imposait : il aurait dû diriger l'entretien parce qu'il était l'employeur.

Julie et Clémence semblent être plus sensibles que les autres à la cohérence et à la cohésion du discours, puisqu'elles sont les deux seules enseignantes qui font des remarques à ce sujet. Nous pourrions expliquer cette sensibilité par le fait que pendant qu'elles visionnaient les jeux de rôles, ces dernières n'ont pas interrompu leur écoute pour noter des erreurs sur la feuille prévue à cette fin. Julie y a noté quelques erreurs seulement et quelques remarques au sujet de la performance globale des étudiants. Clémence, pour sa part, n'a presque rien noté du tout, excepté le score final sur 100. Peut-être que ces deux enseignantes, qui n'étaient pas aussi préoccupées que les autres par la recension des erreurs, ont prêté une

plus grande attention au message en général. Notre expérience ne peut cependant pas le confirmer, mais il serait intéressant poursuivre des recherches afin de déterminer si la prise de notes des évaluateurs pendant le jeu de rôle influence le type de commentaire qu'ils font au sujet de la compétence communicative.

#### 5.4 Compétence interactive

Nous avons vu que, selon Fulcher (2003), la compétence interactive concerne la façon dont le locuteur agit pour faire avancer la conversation. Elle dépasse la compétence stratégique et le savoir textuel parce qu'elle implique simultanément au moins deux personnes. La place de cette composante au sein de la compétence communicative ne fait pas l'unanimité chez les chercheurs, comme nous avons pu le voir dans le chapitre II, le «Cadre théorique», étant donné qu'elle relève souvent plus de la psychologie (elle concerne les traits de personnalité et l'attitude des locuteurs) que de la linguistique. Toutefois, l'analyse des commentaires des enseignantes nous montre que cette composante est souvent invoquée pour justifier leur note. Ce n'est pas surprenant, puisque le jeu de rôle permet à cette compétence de se manifester (Fulcher, 2003; Leblanc, 2002).

Ce sont les étudiants 3 et 4 qui ont récolté le plus de commentaires positifs au sujet de cette composante. Plus particulièrement, l'étudiant 3 semble plaire aux évaluatrices par son écoute et par le support qu'il donne à sa partenaire plus faible que lui. Phuong dit: *«Il essaie de la (l'étudiante 4) supporter pour qu'elle puisse s'exprimer.»* Éleine va dans le même sens : *«Lorsqu'elle (l'étudiante 4) dit : «Combien ça coûte pour payer», c'est pas clair. Pourtant, il a compris. C'est quand même particulier»* Pour sa part, Julie commente :

*Il écoutait sa partenaire. Ce qui peut influencer ma note, c'est que l'étudiant 3, j'avais l'impression qu'il écoutait plus sa partenaire que l'étudiant 1. Dans une interaction, c'est pas seulement de vomir son discours individuel, mais si la personne me pose une question, y répondre. Interagir.*

Toujours au sujet de l'interaction de ces deux étudiants, Phuong note : *Déjà, il y a une intention d'échange. Pour moi, c'est important.* Éline va dans le même sens: *Ils s'entendent bien.*

Nous voyons donc que l'attitude de l'étudiant 3, très coopératif, a aidé sa partenaire à obtenir une bonne note. Cela veut dire que lors d'un jeu de rôle, il suffirait parfois à un étudiant de manifester de l'empathie pour s'assurer une opinion favorable de la part de son évaluateur et ainsi, améliorer son score.

À l'opposé, la première dyade, et en particulier l'étudiant 1, récolte de la part de Julie, Phuong et Clémence des commentaires négatifs au sujet de cette composante, à commencer par Clémence qui déplore que : *«Chacun avait son scénario.»* Phuong aussi remarque leur manque d'interaction et elle est particulièrement sévère avec l'étudiant 1. Elle dit :

*La première chose qui m'a un peu choquée, c'est qu'il n'y a pas de contact entre eux. On dirait que chacun s'est composé un rôle. Lui, il a l'air de celui qui contrôle et je le sens au niveau du rapport. Elle, elle le regarde et tranquillement, il s'en vient vers un contact visuel. Pour moi, ça dénote déjà beaucoup.*

Par ces remarques, nous constatons que Phuong accorde beaucoup d'importance à la compétence interactive des étudiants. En fait, elle semble y accorder autant d'importance qu'à leur compétence langagière puisque l'étudiant 1, bien que n'ayant pas de faiblesses langagières réelles selon Phuong (elle dit : *«il n'a pas beaucoup de faiblesses»*), s'est mérité la même note que l'étudiante 4, qui démontrait pourtant de grandes lacunes sur le plan langagier. Au final, ce qui semble avoir diminué la note de l'étudiant 1 pour Phuong, c'était sa faible écoute de sa partenaire et son attitude dominatrice. L'étudiante 2 n'a pas été pénalisée sur ce plan, parce que selon Phuong, c'est à cause de son partenaire peu coopératif qu'elle n'a pas pu interagir beaucoup.

Pour Julie cependant, l'attitude de l'étudiant 1 n'excuse pas le manque d'interaction de l'étudiante 2. Selon elle, lors du deuxième jeu de rôle, elle était trop effacée. Elle dit :

*Elle ne suivait pas avec l'idée de l'autre. Elle aurait dû embarquer. Elle faisait juste dire : «je ne sais pas, je ne connais pas ce sport-là». C'est tout ce qu'elle disait.*

Contrairement à Phuong, Julie pénalise l'étudiante 2 pour sa faible implication dans le deuxième jeu de rôle. Elle ne lui donne pas d'excuse. Cela transparait dans la note qu'elle lui accorde, qui est beaucoup plus faible que celle de son premier jeu de rôle et plus faible que celle de son partenaire pour ce jeu de rôle.

Finalement, nous voyons donc que le fait de démontrer de l'empathie lors d'un jeu de rôle (comme l'étudiant 3) peut se traduire par une bonne note et que le fait de ne pas montrer d'enthousiasme pour participer à une conversation (comme l'étudiante 2) peut avoir pour conséquence de diminuer son score. La compétence interactive, quoiqu'elle ne figure pas dans les modèles actuels de la compétence communicative, pourrait peut-être y trouver une place, du moins dans l'évaluation des jeux de rôles, puisqu'elle s'y manifeste et qu'elle contribue à faire avancer le dialogue, ce qui est finalement le but véritable de l'interaction orale. D'ailleurs, nos évaluatrices se sont montrées en général sensibles à l'utilisation de cette habileté.

### **5.5 Critères non linguistiques**

Des recherches sur l'évaluation de la compétence langagière à l'oral montrent que les critères non linguistiques ont une influence significative sur le score (Brown, 1995; Pollitt et Murray, 1996; Orr, 2002). Notre expérimentation semble le confirmer, mais uniquement en ce qui concerne le score de l'étudiante 4.

C'est presque exclusivement pour commenter la performance de l'étudiante 4 que les enseignantes invoquent des critères non linguistiques. En plus des commentaires concernant sa compétence interactive (qui ne sont pas linguistiques) et de quelques commentaires contenant sa capacité stratégique, comme le geste et le regard, les enseignantes disent :

*C'est une personne dynamique. Elle a beaucoup d'idées [...] Avoir une étudiante comme ça, c'est quand même le fun. (Julie);*

*La demoiselle anglophone, elle a des problèmes avec ses structures, mais elle est expressive, elle va capter l'attention. (Élaine);*

*C'est beaucoup plus dans l'émotion. (Phuong);*

*Elle utilisait le charme aussi, ce côté-là. En les regardant, je me disais que dans le fond, dans la vraie vie, il faut faire plus souvent des choses comme ça. (Clémence).*

Les commentaires concernant l'étudiante 4 sont intéressants : selon les participantes, elle est visiblement plus faible que les trois autres sur le plan strictement langagier, mais ses évaluatrices semblent tenir à dire du bien d'elle. Seule Marie-Ève reste insensible aux forces para linguistiques de l'étudiante 4 : elle ne fait aucun commentaire sur le sujet et elle est la seule enseignante qui ne lui accorde pas la note de passage. Difficile d'expliquer pourquoi elle est moins sensible que les autres sur cet aspect. Ni son âge ni son expérience ne semblent fournir de justification. Pour cette enseignante, il semble que la maîtrise du code linguistique prime. Son attention est résolument tournée vers la composante langagière. Cette attitude est tout à fait légitime : plusieurs auteurs démontrent que pour différents enseignants, la compétence communicative, c'est d'abord la maîtrise de la composante langagière (Fulcher, 1993; Orr, 2002; Pollit et Murray, 1996; Wigglesworth, 1993). Mais pour les quatre autres enseignantes, il apparaît que les critères extra linguistiques peuvent être aussi importants que les critères linguistiques.

## **5.6 Synthèse des commentaires des enseignantes**

Nous avons analysé les commentaires des cinq enseignantes en regard de la compétence communicative des étudiants. Finalement, que pouvons-nous en retenir? Quelles composantes de la compétence communicative ont la plus grande influence sur la note des étudiants? Si les enseignantes, en deuxième partie d'entrevue, avaient mentionné des critères personnels d'évaluation appartenant surtout à la composante langagière, il apparaît

finalement que leur jugement repose également sur bien d'autres choses. À l'exception de Marie-Ève, pour qui la composante langagière est importante au point où elle ne fait pratiquement aucun commentaire au sujet des autres composantes, toutes les autres enseignantes sont influencées par d'autres aspects de la compétence communicative. Pour Élane, la précision, en particulier l'utilisation de temps de verbes variés, et la phonologie, notamment le volume, priment, mais des lacunes concernant ces aspects peuvent être rachetées par la capacité stratégique. Pour Julie et Clémence aussi, les trois aspects de la composante langagière sont importants, mais la cohésion et la cohérence du message semblent l'être autant. Pour Phuong, la compétence interactive semble avoir une influence déterminante sur la note, au moins autant que la composante langagière.

Notre analyse nous permet constater que les enseignantes portées à noter systématiquement les erreurs (la précision), comme Marie-Ève, sont moins sensibles aux autres aspects de la compétence communicative. Contrairement à ce que nous pourrions croire, les enseignantes peuvent accorder une bonne note à un étudiant faible sur le plan langagier si ce dernier maîtrise les autres aspects de la compétence communicative. Quand un étudiant démontre des forces sur le plan langagier, comme l'étudiante 2, les enseignantes formulent moins de commentaires sur le reste des composantes. À l'inverse, quand un étudiant montre des faiblesses dans la composante langagière, comme l'étudiante 4, il mérite généralement beaucoup de commentaires au sujet des autres composantes et s'il manifeste des forces sur le plan stratégique et interactif, il peut obtenir une note aussi bonne qu'un étudiant qui maîtrise mieux la composante langagière (phonologie, précision et aisance). Finalement, nous voyons que lorsqu'une enseignante «a décidé» qu'un étudiant méritait de passer, elle justifie sa note en invoquant des critères importants pour elle. Parfois, les remarques qu'elle fait contredisent celles d'une autre enseignante : nous avons vu que le même étudiant pouvait susciter des commentaires opposés. Cela démontre que l'évaluation est extrêmement subjective. Le tableau 5.1 traduit les commentaires des enseignantes en «conseil» formulés à l'endroit des étudiants concernant les aspects de chaque composante qui nous ont semblé les plus susceptibles d'influencer leur note.

**Tableau 5.1**  
**Synthèse des remarques**

Composantes	Sous-composantes	À faire	À éviter
Composante langagière	Phonologie	Faire ses liaisons ; Parler fort ; Parler avec intonation ; Bien prononcer les phonèmes.	Prononcer les consonnes finales muettes, comme le «s» et le «t».
	Précision	Utiliser le passé composé et l'imparfait; Utiliser les pronoms relatifs;  De façon générale, réutiliser le contenu linguistique vu en classe.	Placer l'adverbe et l'adjectif au mauvais endroit dans la phrase ; Mal conjuguer ses verbes au présent ; Ne pas contracter les articles : «de le» à la place de «du».
	Aisance	Parler beaucoup.	Hésiter.
Capacité stratégique	Stratégies de réussite	Parler d'un sujet connu ; Faire des gestes ; Montrer sa volonté de se faire comprendre en s'engageant dans la conversation.	Utiliser la syntaxe et le vocabulaire de la langue maternelle.
	Stratégies d'évitement		Ne pas essayer de parler, même si le sujet est peu familier.
Savoir textuel		Faire des liens entre les différentes étapes de son discours ; Tenir compte du discours de son partenaire pour construire son discours.	Répondre de façon erronée aux questions de son partenaire.
Compétence interactive		Démontrer de l'empathie envers un partenaire plus faible; Montrer qu'on écoute son partenaire.	Parler plus que son partenaire ; Ne pas regarder son partenaire.
Critères non linguistiques		Montrer de l'enthousiasme.	

Mais outre les composantes de la compétence communicative, d'autres critères sont invoqués par les enseignantes pour justifier la note qu'elles accordent à leurs étudiants en fin de session. Évaluer des étudiants inconnus et évaluer leurs propres étudiants semblent deux activités très différentes. Nous allons voir, dans la partie suivante, que la performance des étudiants lors de l'examen final n'a souvent pas une grande influence sur le score final. En effet, les commentaires des cinq enseignantes révèlent que dans la réalité, la partie est souvent jouée d'avance.

Dans la partie suivante, nous verrons que deux facteurs : la réutilisation du contenu enseigné en classe et les observations générales des enseignantes pendant la session peuvent avoir plus d'influence sur la note de l'étudiant que sa réussite du jeu de rôle.

### 5.7 Réutilisation des acquis propres au niveau

D'abord, les enseignantes tiennent à retrouver dans la performance d'examen de leurs étudiants ce qui a été enseigné pendant le cours. Elles s'attendent à percevoir les éléments linguistiques couverts par le programme dans la performance de leurs étudiants. Toutes les enseignantes nous ont dit que le critère le plus important, pour elle, était la capacité à réutiliser la matière vue en classe. Cela a d'ailleurs amené Éline à créer une grille d'évaluation qui intègre dans ses descripteurs les notions grammaticales et syntaxiques vues pendant la session. Cela amène aussi Julie et Marie-Ève à modifier l'épreuve prescrite par la DGFA (1994), pour que cette épreuve soit plus conforme au contenu vu en classe. La remarque de Julie est éloquente à ce sujet :

*Des fois, je ne vais même pas suivre le prototype. [...] Si c'est un thème [celui prévu dans l'épreuve] qu'on a vu ensemble, je vais leur en parler, mais si c'est juste comme ça, du jour au lendemain, je m'attends pas à ce qu'ils soient capables de me faire un discours.*

C'est d'ailleurs la raison pour laquelle Clémence a dit avoir eu tant de mal à évaluer les quatre étudiants lors de notre expérimentation: ayant donné le niveau 5 seulement une fois en 19 ans de carrière, elle ne savait pas sur quoi se baser pour les noter : «Ça m'embêtait d'évaluer tes étudiants parce que je n'avais pas les objectifs du niveau 5.». Donc, le premier

critère d'évaluation, selon les cinq enseignantes, est l'utilisation correcte des acquis propres au niveau, enseignés pendant la session, ce qui semble tout à fait logique. Cette attente légitime en contexte scolaire est rapportée dans la littérature:

Overall achievement at the end of a course should be assessed with reference primarily to course objectives because they are the most inclusive statement of what students are expected to have learned. (Genesee et Upshur, 1996, p.19)

Mais si les cinq enseignantes nous disent de façon unanime que ce qu'elles recherchent avant tout, c'est l'atteinte des objectifs linguistiques (critère qui concerne essentiellement la composante langagière) du niveau 5, enseignés pendant la session, comment se fait-il que même celles qui connaissent les objectifs linguistiques du niveau 5 n'arrivent pas à un consensus dans l'évaluation des quatre étudiants? Nous pensons qu'il y a deux raisons principales à cela. La première, c'est que la grille d'évaluation de la DGFA (1994) ne tient absolument pas compte du contenu linguistique sujet à l'évaluation : rappelons que les enseignants utilisent la même grille pour tous les niveaux. Il revient à l'enseignant de choisir parmi les éléments linguistiques du programme ceux qui seront recherchés lors de l'évaluation. Or, nous avons vu que ce choix varie selon les priorités des enseignantes.

La deuxième raison, c'est que même lorsque les enseignantes connaissent les objectifs linguistiques du niveau, elles ne les perçoivent pas de la même façon dans le discours des étudiants. Nous l'avons vu dans la perception de l'utilisation des temps de verbes du passé par l'étudiante 4 : pour une enseignante, elle les utilisait bien et pour une autre, l'étudiante ne conjuguait pas ses verbes.

En outre, l'analyse des résultats révèle que même si les enseignantes disent que les critères qui sont importants pour elles relèvent presque exclusivement de la composante langagière, nous voyons que lors de l'évaluation, toutes les autres composantes de la compétence communicative influencent leur jugement. Globalement, l'attitude de l'étudiant peut avoir autant d'importance que sa compétence strictement langagière.

Mais au-delà de ces considérations sur la façon d'évaluer les étudiants lors de l'examen de production orale, les enseignantes nous ont révélé que finalement, ce n'est pas tant la performance des étudiants à l'examen qui importe, mais leur performance tout au long de la session. La prochaine section mettra ce fait en lumière.

### 5.8 Les observations pendant la session

Les enseignantes nous ont révélé que les observations qu'elles font de leurs étudiants tout au long de la session ont une grande influence sur leur note à l'examen oral. Il se dégage des commentaires des enseignantes que l'examen final joue un rôle plutôt symbolique puisque les enseignantes ont déjà une idée assez précise de leur score, avant même que l'étudiant soit testé formellement. À ce sujet, Éleine dit : *Quand je corrige l'oral, je fais 3 piles : mes meilleurs, les moyens et les faibles et ça reste pas mal comme ça.* Clémence va dans le même sens :

*Je les observe souvent, de toute façon, je les observe tout le temps. Je le sais, ces trois-là sont super bons, ceux là sont moyens, ceux là en arrachent. Ça joue un rôle très important, ces observations.*

Dans le même ordre d'idée, pendant l'entrevue, Clémence nous a parlé d'une procédure administrative en cours à la CSDM (c'est aussi le cas à la Commission scolaire Marguerite-Bourgeoys), qui vise à faciliter l'organisation scolaire. Avant la fin de la session, la direction demande aux enseignants de faire des prévisions, c'est-à-dire de prédire le nombre d'échecs et de réussites dans leur groupe, afin de former le plus tôt possible les groupes de la session suivante. Donc, environ une semaine avant les examens, les groupes de la session suivante sont déjà formés et ne changent à peu près pas. C'est dire à quel point les observations des enseignants sont prises au sérieux. L'examen sert à confirmer ces observations et à déterminer la note finale de l'étudiant. Clémence commente cette mesure en disant :

*C'est peut-être aussi bien. J'ai remarqué que des fois, tu hésites [à faire passer un étudiant] et à chaque fois que je me suis dit que j'allais le faire passer pareil, c'était une erreur de ma part. Quand tu hésites, c'est mieux de pas les faire passer.*

Les commentaires des enseignantes illustrent bien qu'en bout de ligne, ce n'est pas tant l'examen final que le jugement global de l'enseignant au fil de la session qui importe. Les étudiants sont évalués par des enseignants et des enseignantes qui les connaissent, puisqu'ils ont passé huit semaines à les observer et à interagir avec eux. Nous pourrions croire que le contact prolongé avec les étudiants permet aux enseignants de juger adéquatement leur compétence réelle. Mais est-ce vraiment le gage d'une plus grande objectivité? Selon une étude de Glisan et Foltz (1998), il semblerait que non. Cette étude, qui consistait à vérifier si les intuitions des enseignants au sujet de la compétence en production orale de leurs étudiants étaient confirmées par les résultats des évaluations nationales, a montré que les enseignants, surtout les moins expérimentés, se trompent très souvent au sujet de la compétence réelle de leurs étudiants. Les chercheurs démontrent qu'il existe effectivement une corrélation entre ce qui se passe en salle de classe et ce qui se passe à l'examen national, mais qu'une bonne performance en classe n'est pas nécessairement garante d'une bonne performance à cet examen, qui n'est pas relié au programme scolaire et qui est évalué par des inconnus. C'est donc dire que malgré ce que les enseignantes pensent, le fait de connaître leurs étudiants ne les rend pas nécessairement plus objectives pour les évaluer.

## **5.9 Synthèse du chapitre**

Finalement, suite à l'analyse des données, quels constats pouvons-nous faire? D'abord, il est clair que les enseignantes n'accordent pas toutes la même importance aux erreurs et qu'elles ne remarquent pas les mêmes aspects du discours des étudiants. Des facteurs personnels, tels que l'expérience, la connaissance de la langue parlée par l'étudiant ou encore la langue maternelle influencent sûrement l'évaluation de la composante langagière.

En outre, il est évident que les grilles ne sont pas garantes de l'objectivité, qu'elles soient globales ou analytiques. Les enseignantes n'utilisent pas toutes les mêmes grilles et même si c'était le cas, nous voyons qu'elles ont tendance à d'abord fixer une note et ensuite à organiser la pondération pour arriver à cette note.

De plus, des critères autres que linguistiques tels que la coopération, l'empathie manifestée envers son partenaire et des traits de personnalité tels que le charme et l'extraversion, ont une très grande influence sur le jugement global de la prestation orale d'un étudiant, pouvant même faire en sorte qu'il obtienne une excellente note, malgré des lacunes importantes sur le plan langagier.

Finalement, l'évaluation est grandement influencée par les attentes des enseignantes et peu par la réussite de la tâche prévue par le Ministère. Elle est souvent la confirmation des observations en salle de classe, donc, en grande partie jouée avant même d'avoir lieu.

## CONCLUSION

L'évaluation de la compétence communicative à l'oral en français, langue seconde, a toujours été une préoccupation pour les chercheurs en linguistique et en didactique des langues. Elle évolue en suivant le rythme des différents courants de pensée en linguistique. L'évaluation s'adapte au style d'enseignement dans lequel elle s'insère et elle reflète les valeurs des instances qui y ont recours. Dans tous les cas, elle nécessite une définition claire et opérationnelle du construit à évaluer. Mais comme l'évaluation de la compétence communicative à l'oral est affaire de perception et de conviction personnelle des évaluateurs, la subjectivité qu'elle appelle est inévitable.

Dans ce mémoire, qui visait à découvrir les critères d'évaluation invoqués par cinq enseignantes pour justifier la note attribuée aux étudiants, nous avons mené une expérimentation qui a placé les enseignantes en situation d'évaluation. Au terme de cette expérimentation, nous avons vu qu'en contexte scolaire, notamment à l'éducation des adultes à la Commission scolaire de Montréal, le choix des critères d'évaluation des cinq participantes se résumait à huit : 1) le respect de la consigne, 2) la clarté du message, 3) la phonologie, 4) la phonétique, 5) l'aisance, 6) la syntaxe, 7) la conjugaison et 8) le vocabulaire. Tous ces critères, excepté le respect de la consigne et la clarté du message (qui relèvent de la composante textuelle (cohérence et cohésion)), relèvent de la composante langagière. Toutefois, nous avons constaté que dans les faits, lors de l'évaluation, les enseignantes tiennent compte aussi de la capacité stratégique de l'étudiant, de sa compétence

interactive et de critères non linguistiques. L'analyse des résultats nous a permis de constater que les enseignantes accordent une importance variable à chacun de ces critères. De plus, les enseignantes nous ont révélé que dans la pratique, ce sont surtout leurs observations en classe et la réutilisation par les étudiants du contenu linguistique enseigné pendant la session qui influençaient leur décision.

En outre, notre expérimentation a révélé l'absence de constance entre les cinq évaluatrices, leur jugement au sujet du même étudiant variant beaucoup. Ceci peut paraître somme toute relativement inquiétant.

### **Limites de la recherche**

Une des principales limites de notre recherche réside dans l'interprétation et la codification des commentaires. Nous avons choisi de codifier les commentaires selon leur appartenance à une composante de la compétence communicative, afin de déterminer quelle composante a le plus d'impact lors de l'évaluation. Or, bien que nous ayons étudié plusieurs modèles de la compétence communicative (celui de Canale et Swain, 1980 ; de Bachman et Palmer 1996 et de Fulcher, 2003), nous n'avons pas pu dégager un consensus dans la définition de chaque composante de la compétence communicative. Les différents auteurs consultés la définissent de façon différente et il n'est pas facile de tracer les limites de chacune des composantes de cette compétence. Nous avons dû adapter un modèle pour y classer les commentaires des enseignantes. Non seulement il n'a pas été simple de choisir un modèle de la composante, mais nous avons décidé de façon quelque peu subjective dans quelle composante classer les commentaires recueillis. Parfois, un commentaire pouvait appartenir à deux composantes. Ceci a rendu la codification difficile et spécifique à nos composantes.

Une autre limite réside évidemment dans le nombre restreint de participantes et dans l'étroitesse de l'échantillon. Cinq enseignantes, ce n'est certainement pas représentatif de l'ensemble du corps enseignant. De plus, l'échantillonnage, limité à la production orale de

quatre étudiants, n'a certainement pas permis de faire ressortir un nombre représentatif de commentaires. En outre, si les étudiants évalués avaient été débutants, les commentaires des enseignantes auraient sûrement porté sur d'autres composantes de la compétence communicative.

### **Retombées**

Malgré ses limites, nous croyons que notre recherche peut avoir des retombées importantes pour les enseignants, les directions d'école et finalement, les étudiants. D'abord, elle permettra aux enseignants de comparer leurs perceptions à celles de leurs collègues et ainsi, à réévaluer l'importance de leurs critères d'évaluation et peut-être à en adopter de nouveaux. Cette recherche, qui révèle l'absence de consensus entre les évaluatrices, montre que l'évaluation, un acte isolé qui repose sur un jugement très subjectif, peut prendre des directions opposées, encore plus que les enseignants pourraient le croire. Observer des faits aussi éloquents que ceux que révèle notre recherche pourrait inciter les enseignants à repenser leur façon de faire.

L'absence de constance entre les évaluatrices, révélée par notre expérimentation, pourrait alarmer les directions d'école, responsables de diplômer les étudiants. Nous voyons que la note qui figure à leur dossier scolaire ne veut pas dire grand chose, puisqu'elle résulte en grande partie d'un jugement qui varie beaucoup d'un enseignant à l'autre. Il serait peut-être temps pour les directions des centres d'éducation des adultes de prendre des mesures afin d'assurer une plus grande constance entre les évaluateurs, afin de redonner de la crédibilité au diplôme octroyé aux étudiants. La première de ces mesures pourrait être la construction d'une grille d'évaluation, qui reposerait sur une définition claire du construit que nous voulons évaluer. Si des composantes comme la compétence interactive ou la capacité stratégique sont jugées importantes par les enseignants, il faudrait que la grille d'évaluation leur réserve une place, ce qui n'est pas le cas présentement. En outre, les directions pourraient décider de donner une formation sur la façon d'interpréter la grille aux enseignants. D'ailleurs, les enseignants réclament une grille plus claire et une formation en docimologie.

De plus, nous croyons que les étudiants pourront bénéficier de notre recherche, en prenant conscience de l'importance de certains critères d'évaluation. Notre recherche révèle que les enseignants sont particulièrement sensibles à certains aspects de la compétence communicative et le fait de connaître ces aspects pourrait aiguiller les étudiants sur ce qu'ils devraient travailler en priorité afin d'améliorer leur compétence communicative à l'oral.

Finalement, nous pensons qu'au delà du monde scolaire, notre recherche pourrait intéresser tout chercheur en linguistique désireux de mieux comprendre ce que sont la compétence communicative et l'évaluation de cette compétence.

## APPENDICE A

### GRILLE DU MELS

#### Expression d'un message oral

LAN-1002-4 à LAN-4048-4	LAN-S071-4 et LAN-S072-4	<b>Description du niveau d'habileté de l'expression</b>
20	15	<p>Message complet, bien organisé et facile à suivre, qui contient toutes les idées essentielles. Consigne bien respectée. Structures justes, vocabulaire approprié L'élève s'exprime avec aisance et a un débit naturel. En réutilisation, les éléments linguistiques sont personnels et appropriés.</p>
16	12	<p>Message bien organisé et facile à suivre, qui contient la plupart des idées essentielles. Consigne bien respectée. Quelques erreurs de structures, vocabulaire généralement approprié. L'élève s'exprime avec une certaine facilité. En réutilisation, la majorité des éléments linguistiques sont personnels et souvent appropriés.</p>
12	9	<p>Message significatif, bien que limité. Consigne assez bien respectée. Erreurs de structures assez fréquentes, vocabulaire limité ou redondant. L'élève hésite parfois et a besoin d'une aide occasionnelle de la part de l'interlocuteur ou de l'interlocutrice. En réutilisation, plusieurs des éléments linguistiques sont personnels bien que l'on y retrouve également des éléments identiques au texte de la situation. Ils sont parfois inappropriés.</p>
8	6	<p>Message incomplet ou manquant souvent de logique. Consigne partiellement respectée. Structures déficientes, vocabulaire sans recherche. L'élève hésite souvent et a besoin d'une aide fréquente de la part de l'interlocuteur ou de l'interlocutrice. En réutilisation, certains des éléments linguistiques peuvent être personnels, bien que l'on y retrouve plusieurs éléments identiques au texte de la situation. Ils sont souvent inappropriés.</p>
4	3	<p>Message incomplet, qui nécessite un effort de concentration de la part de l'interlocuteur ou de l'interlocutrice. Consigne peu respectée. Structures déficientes, vocabulaire limité, recours occasionnels à la langue maternelle. L'élève hésite souvent et longtemps et a besoin d'une aide constante de la part de l'interlocuteur ou de l'interlocutrice. En réutilisation, la majorité des éléments linguistiques sont identiques au texte de la situation d'origine. Ils sont très souvent inappropriés.</p>
0	0	L'élève ne parvient pas à communiquer un message ayant un sens.

## APPENDICE B

### QUESTIONNAIRE SOCIODÉMOGRAPHIQUE

#### Page 1

**S'il vous plaît, veuillez remplir la section suivante :**

Prénom et nom: \_\_\_\_\_

Adresse/rue : \_\_\_\_\_

Ville : \_\_\_\_\_

Code postal : \_\_\_\_\_

Numéro de téléphone : \_\_\_\_\_

Courriel : \_\_\_\_\_

Lieu(x) d'enseignement actuel(s) : \_\_\_\_\_

\_\_\_\_\_

Code du participant : \_\_\_\_\_

#### Page 2

**S'il vous plaît, veuillez remplir les sections suivantes.**

1. Sexe :  Femme  Homme

2. Âge :  18-23 ans  24-29 ans  30-39 ans  40-49 ans  50-59 ans  60 ans+

3. Langue maternelle (apprise en premier et toujours parlée) :

Français  Anglais

Autre : \_\_\_\_\_

4. Autres langues comprises et parlées :

Français  Anglais

Autre(s) : \_\_\_\_\_

5. Langue(s) parlée(s) à la maison :

Français       Anglais

Autre(s) : \_\_\_\_\_

Code du participant : \_\_\_\_\_

Page 3

**6. Années d'expérience en enseignement (pour chaque catégorie) :**

<b>Français langue seconde (primaire/secondaire) :</b>					
<input type="checkbox"/> Aucune	<input type="checkbox"/> 0-3 ans	<input type="checkbox"/> 3-5 ans	<input type="checkbox"/> 5-8 ans	<input type="checkbox"/> 8-11 ans	<input type="checkbox"/> 11 ans +
En quelle(s) année(s)? : _____					
<b>Français langue seconde aux adultes :</b>					
<input type="checkbox"/> 0-4 ans	<input type="checkbox"/> 5-9 ans	<input type="checkbox"/> 10-14 ans	<input type="checkbox"/> 15-19 ans	<input type="checkbox"/> 20-24 ans	<input type="checkbox"/> 25 ans +
En quelle(s) année(s)? : _____					
<b>Nombre approximatif de semestres d'expérience en enseignement du français, langue seconde aux adultes, par niveau :</b>					
Préalable : _____	Niv. 1 : _____	Niv. 2 : _____	Niv. 3 : _____		
Niv. 4 : _____	Niv. 5 : _____	Niv. 6 : _____	Niv. 7 : _____		
<b>Français langue maternelle (primaire/secondaire) :</b>					
<input type="checkbox"/> Aucune	<input type="checkbox"/> 0-3 ans	<input type="checkbox"/> 3-5 ans	<input type="checkbox"/> 5-8 ans	<input type="checkbox"/> 8-11 ans	<input type="checkbox"/> 11 ans +
En quelle(s) année(s)? : _____					
<b>Français langue maternelle aux adultes :</b>					
<input type="checkbox"/> Aucune	<input type="checkbox"/> 0-3 ans	<input type="checkbox"/> 3-5 ans	<input type="checkbox"/> 5-8 ans	<input type="checkbox"/> 8-11 ans	<input type="checkbox"/> 11 ans +
En quelle(s) année(s)? : _____					
<b>Autre type d'enseignement (précisez, s'il vous plaît) :</b> _____					
_____					
<input type="checkbox"/> 0-2 ans	<input type="checkbox"/> 3-5 ans	<input type="checkbox"/> 6-8 ans	<input type="checkbox"/> 9-11 ans	<input type="checkbox"/> 12-14 ans	<input type="checkbox"/> 15 ans +
En quelle(s) année(s)? : _____					

7. Nombre total d'années d'expérience en enseignement : \_\_\_\_\_

8. Année d'obtention du brevet : \_\_\_\_\_

9. Champ(s) d'enseignement: \_\_\_\_\_

Code du participant : \_\_\_\_\_

**Page 4**

**10. Lieu(x) d'enseignement:**

- Organisme communautaire  
 Ministère de l'Immigration et des Communautés culturelles (MICC)  
 Commissions scolaires  
 Cégep  
 Université  
 École privée  
 Autre(s) \_\_\_\_\_

**11. Années d'expérience dans chaque type d'établissements :**

<b>Organisme communautaire :</b>					
<input type="checkbox"/> Aucune	<input type="checkbox"/> 0-3 ans	<input type="checkbox"/> 4-7 ans	<input type="checkbox"/> 7-10 ans	<input type="checkbox"/> 11-14 ans	<input type="checkbox"/> 15 ans +
En quelle(s) année(s)?: _____					
<b>Ministère de l'Immigration et des Communautés culturelles :</b>					
<input type="checkbox"/> Aucune	<input type="checkbox"/> 0-3 ans	<input type="checkbox"/> 4-7 ans	<input type="checkbox"/> 7-10 ans	<input type="checkbox"/> 11-14 ans	<input type="checkbox"/> 15 ans +
En quelle(s) année(s)?: _____					
<b>Commission scolaire :</b>					
<input type="checkbox"/> 0-3 ans	<input type="checkbox"/> 4-6 ans	<input type="checkbox"/> 7-9 ans	<input type="checkbox"/> 10-12ans	<input type="checkbox"/> 13-15 ans	<input type="checkbox"/> 16 ans+
En quelle(s) année(s)?: _____					
<b>Cégep :</b>					
<input type="checkbox"/> Aucune	<input type="checkbox"/> 0-3 ans	<input type="checkbox"/> 4-7 ans	<input type="checkbox"/> 7-10 ans	<input type="checkbox"/> 11-14 ans	<input type="checkbox"/> 15 ans +
En quelle(s) année(s)?: _____					
<b>Université :</b>					
<input type="checkbox"/> Aucune	<input type="checkbox"/> 0-3 ans	<input type="checkbox"/> 4-7 ans	<input type="checkbox"/> 7-10 ans	<input type="checkbox"/> 11-14 ans	<input type="checkbox"/> 15 ans +
En quelle(s) année(s)?: _____					

**École privée :**

Aucune    0-3 ans    3-5 ans    5-8 ans    8-11 ans    11 ans +  
En quelle(s) année(s)? : \_\_\_\_\_

**Autre(s) :** \_\_\_\_\_

0-3 ans    4-6 ans    7-9 ans    10-12ans    13-15  
ans   16 ans+

En quelle(s) année(s)? : \_\_\_\_\_

**Merci beaucoup de votre coopération.**

Code du participant : \_\_\_\_\_

## APPENDICE C

### FEUILLE DE NOTES

#### Fiche d'évaluation

Code du participant ou de la participante : \_\_\_\_\_

Examen oral 1

Étudiant 1

Note sur 15

Notes personnelles

Aide-mémoire

Commentaires,

Relevé des erreurs :

Étudiante 2

Note sur 15

Notes personnelles

Aide-mémoire

Commentaires,

Relevé des erreurs :

## **APPENDICE D**

### **QUESTIONNAIRE D'ENTREVUE SEMI DIRIGÉE**

#### **Première partie : l'évaluation des étudiants**

1. Quelle note avez-vous accordée à l'étudiant x ?
2. Pourquoi avez-vous accordé cette note ?
3. De façon générale, comment avez-vous trouvé son examen ?
4. Quelles étaient ses forces ?
5. Quelles étaient ses faiblesses ?
6. Quelles erreurs avez-vous remarquées ?
7. En comparaison avec l'étudiant x, comment était l'étudiant y ? (répéter la question pour chaque paire)

#### **Deuxième partie : les pratiques évaluatives et les critères d'évaluation personnels**

1. De façon générale, comment vous y prenez-vous pour évaluer vos étudiants ?
2. Utilisez-vous la grille du MELS ? Pourquoi ?
3. Utilisez-vous les prototypes d'examen ? Pourquoi ?
4. Quels sont les critères d'évaluation les plus importants pour vous ?
5. Qu'entendez-vous par le critère de x ?
6. Quelle importance accordez-vous au critère x, sur une échelle de 1 à 5 ? Pourquoi ?

## APPENDICE E

### GRILLE DE JULIE

#### Présentation individuelle

Sujet: \_\_\_\_\_

<b>Message:</b> consignes, effort mis sur la présentation	1	2	3	4	5	/5
<b>Grammaire</b>	1	2	3	4	5	/5
<b>Total:</b>						/10

#### Rencontre individuelle

Situation: \_\_\_\_\_

<b>Message:</b> consignes, cohérence, réutilisation des acquis	1	2	3	4	5	/5
<b>Grammaire:</b> syntaxe, verbes	1	2	3	4	5	/5
<b>Vocabulaire:</b> richesse, choix, acquis	1	2	3	4	5	/5
<b>Phonétique:</b> débit, prononciation, intonation	1	2	3	4	5	/5
<b>Total:</b>						/20

**APPENDICE F**

**GRILLE D'ÉLAINE**

Expression d'un Message Oral  
(25%)

<b>Domaine :</b> <b>Tâche :</b> <b>Item :</b>	<hr/> <hr/> <hr/>					
<b>Légende des cotes</b> 5 - excellent 4 - très bien - 3 - bien 2 - passable 1 - à améliorer	5	4	3	2	1	Erreurs
<b>1. Message clair</b> - pertinence du contenu - nombre de renseignements - développement des idées						
<b>2. Structures justes</b> - Syntaxe: phrases, mots liens - pronoms relatifs - ordre des mots						
<b>3. Grammaire</b> - conjugaison des verbes, - temps, concordance, - déterminants, prépositions - pronoms compléments - expressions de temps						
<b>4. Vocabulaire</b> - riche et non-redondant, - n'emploi pas la langue maternelle ni les calques						
<b>5. Facilité d'élocution</b> - prononciation, débit, rythme - liaison, enchaînement - n'a pas d'hésitation, ni besoin d'aide						

Total : /25

Commentaires :

## APPENDICE G

### GRILLE DE CLÉMENCE

<b>Grille d'évaluation</b>		
1. Compréhension du message : est-ce que la réponse est conforme à la tâche demandée ?		<b>/ 5</b>
2. Expression du message par des actes de parole : voir les objectifs du niveau: temps de verbes, structures des phrases, vocabulaire.		<b>/20</b>
3. Rythme : temps d'exécution, fluidité de la phrase.		<b>/5</b>
4. Phonétique : la finale des verbes, les interférences avec la langue maternelle maîtrisée ou non, jusqu'ou son accent interfère le message et devient + ou - compréhensible.		<b>/10</b>
Note finale :	<b>/40</b>	<b>/40</b>

## RÉFÉRENCES

- Allal, Linda. 2000. «Acquisition et évaluation des compétences en milieu scolaire». Voir Dolz, et Ollagnier. 2000.
- Alptekin, Cem. 2002. «Towards Intercultural Communicative Competence in ELT». *ELT Journal*, vol. 56, no 1, p. 57-64.
- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. 1999. *Standards for Educational and Psychological Testing*. Washington, 98 p.
- Abddallah-Preteille, Martine. 1996. «Compétence culturelle, compétence interculturelle. Pour une anthropologie de la communication». *Le Français dans le monde*, no 11, (janvier), p. 28-38.
- Bachman, Lyle. 1990. *Fundamental Considerations in Language Testing*, Oxford University Press, 408 p.
- Bachman, Lyle et Adrian S. Palmer. 1982. «The Construct Validation of Some Components of Communicative Proficiency». *TESOL Quarterly*, vol.16, no 4, p. 449-465.
- Bachman, Lyle et Adrian S. Palmer. 1996. *Language Testing in Practice*. Oxford University Press, 377 p.
- Bachman, Lyle, Brian Lynch et Maureen Mason. 1995. «Investigating Variability in Tasks and Rater Judgements in a Performance Test of Foreign Language Speaking». *Language Testing*, vol. 12, no 2, p. 239-257.
- Beacco, Jean-Claude. 2002. «Synthèse : Sur les fonctions de la notion de compétence en langue en didactique des langues». Voir Castellotti, et Py. 2002.
- Berrier, Astrid. 1991. «L'évaluation de l'oral : quelles questions?». *The French Review*, vol. 64, no 3, p. 476-486.
- Biber, Douglas. 1998. *Variation Across Speech and Writing*. Cambridge : Cambridge University Press, 320 p.
- Biemans, H., Nieuwenhuis, L., Poell, R., Mulder, M. and Wesselink, R. (2004). «Vocational and Business Education and training in Europe; Qualifications and the World of Work». bwp@Issue No. 7, [http://www.bwpat.de/7eu/bimans\\_etal\\_1\\_bwpat7.shtml](http://www.bwpat.de/7eu/bimans_etal_1_bwpat7.shtml)

- Bissonnette, Steve et Mario Richard. 2001. *Comment construire des compétences en classe*. Montréal : Chenelière/McGraw-Hill, 138 p.
- Bolton, Sibylle. 1987. *Évaluation de la compétence communicative en langue étrangère*. col. LAL, Paris : Hatier-Credif, 143 p.
- Boutin, Gérald et Louise Julien. 2000. *L'Obsession des compétences. Son impact sur l'école et la formation des enseignants*. Montréal : Éditions Nouvelles, 107 p.
- Bronckart, Jean-Paul. 1977. *Théories du langage : une introduction critique*. Bruxelles : P. Mardaga, 361 p.
- Brown, Annie. 2005. «Book Review. Fulcher G. Testing second language speaking», *Language Testing*, vol.22, no 2, p.234-241.
- Brown, Annie. 2003. «Interviewer Variation and the Co-Construction of Speaking Proficiency». *Language Testing*, vol. 20, no 1, p. 1-25.
- Brown, Anne. 1995. «The Effect of Rater Variables in the Development of an Occupation-Specific Language Performance Test». *Language testing*, vol. 12, no 2, p.1-15.
- Byram, Michael et Geneviève Zarate. 1998. «Définitions, objectifs et évaluation de la compétence socioculturelle». *Le Français dans le monde*, (juin) p. 70-96.
- Brumfit, Christopher. 1980. «Ideology, Communication, and Learning to Use English». *English Language Teaching Journal*, vol. 34, no 3, p. 169-172
- Bygate, Martin, Peter Skehan et Merrill Swain (Dir.). 2001. *Researching pedagogic tasks. Second language learning, teaching and testing*. Essex, England: Pearson Education, 258 p.
- Cadre européen commun de référence. 2005.  
[http://www.dfes.gov.uk/languages/DSP\\_languagesladder.cfm](http://www.dfes.gov.uk/languages/DSP_languagesladder.cfm)
- Canale, Michael et Merrill Swain. 1980. « Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing». *Applied Linguistics*, vol. 1, no 1, 1-47.
- Canale, Michael. 1983. «On Some Dimension of Language Proficiency». Voir Oller, 1983.
- Carroll, John B. 1961. «The Nature of Data, or How to Choose a Correlation Coefficient». *Psychometrika*, vol. 26, no 4, p. 347-372.
- Carroll, John B. 1983. «Psychométric Theory and Language Testing». Voir Oller, 1983.

- Castellotti, Véronique et Bernard Py (Dir.). 2002. *La Notion de compétence en langue*. Lyon : ENS Éditions, 123 p.
- Chalhoub-Deville, Micheline. 1995. «Deriving Oral Assessment Scales Across Different Tests and Rater Group». vol. 12, no 1, p. 17-33.
- Chalhoub-Deville, Micheline et Carolyn. Turner. 2000. «What to look for in ESL Admission Tests : Cambridge Certificate Exams, IELTS and TOEFL». *System*, vol. 28, p. 523-539.
- Chalhoub-Deville, Micheline. «Performance assessment and the Components of the Oral Construct Across Different Tests and Rater Groups». voir Milanovic et Saville, 1996.
- Chambers, Francine. 1997. «What do we Mean by Fluency?». *System*, Vol. 25, no 4, p. 535-544.
- Chomsky, N.1965. *Aspects of the Theory of Syntax*. Cambridge: M.I.T. Press, 251 p.
- Cumming, Alister. 1994. «Does Language Assessment Facilitate Recent Immigrant's Participation in Canadian Society?». *TESL Canada Journal*, vol. 11, no 2, 117-133.
- Cumming, Alister et Richard Berwick. 1996. *Validation in language testing*. Clevedon: Modern Language in Practice, 258 p.
- Cumming, Alister. 1996. «Introduction: The Concept of Validation in Language Testing». Voir Cumming et Berwick, 1996.
- Cushing Weigle, Sara et Brian Lynch. 1996. «Hypothesis Testing in Construct Validation». Voir Cumming et Berwick, 1996.
- Cziko, Gary. 1983. «Psychometric and Edumetric Approches to language testing». Voir Oller, 1983.
- Center for Canadian Language Benchmarks. 2005. <http://www.language.ca/>
- Dolz, Joaquim et Bernard Schneuwly. 1998. *Pour un enseignement de l'oral: initiation aux genres formels à l'école*. Paris : ESF, 211 p.
- Dolz, Joaquim et Edmée Ollagnier. 2000. *L'Énigme de la compétence en éducation*. Paris-Bruxelles : DeBoeck, 124 p.
- Dolz, Joaquim. 2002. «L'énigme de la compétence en éducation. Des travaux en sciences de l'éducation revistés». Voir Castellotti et Py. 2002.
- Farhady, Hossein. 1983. «The Disjonctive Fallacy Between Dicrete-Point and Integrative Tests». Voir Oller. 1983.

- Figueras, Neus. 2005. «Testing, testing everywhere, and not a while to think». *ELT Journal*, vol. 59, no 1, p.47-54.
- Fontenay, Hervé de. 1991. «Les techniques d'entrevue au service de l'évaluation de la performance orale». *Actes du colloque AQEFLS 1990. Bulletin AQEFLS*, vol. 12, no 3-4, 77-97.
- Fulcher, Glenn. 2000. «The 'Communicative' Legacy in Language Testing» *System*, vol. 28, 483-497.
- Fulcher, Glenn. 2003. *Testing second Language Speaking*. Harlow : Pearson Education, 288 p.
- Genesee, Fred et John A. Upshur. 1996. *Classroom-based Evaluation in Second Language Education*. New York : Cambridge University Press, 268 p.
- Germain, Claude, Joan Netten et Parvin Movassat. 2004. «L'évaluation de la production orale en français intensif : critères et résultats». *La Revue canadienne des langues vivantes*, vol. 60, no 3, 309-352.
- Germain, Claude et Joan Netten. 2002. «La précision et l'aisance en FLE/FL2 : définitions, types et implications pédagogiques». *Actes du colloque «La didactique des langues face aux cultures linguistiques et éducatives»*. *Marges linguistiques* : <http://www.marges-linguistiques.com>
- Glisan, Eileen et David A. Foltz. 1998. «Assessing Students' Oral Proficiency in an Outcome-Based Curriculum : Student Performance and Teacher Intuitions». *The Modern Language Journal*, vol. 82, no 1, p.1-18.
- Gronlund, Norman E. 1985. *Measurement and Evaluation in Teaching*. New York : Macmillan Publishing compagny, 540 p.
- Grice, Paul. 1979. Logique et conversation. *Communications*, no 30, p. 57-72.
- Harrison, Andrew. 1983. *A Language testing Book*. London : The Macmillan Press Limited, 144 p.
- Hinofotis, Frances. 1983. «The Structure of oral communication in an educational environment: a comparison of factor analytic». Voir Oller. 1983.
- Holec, Henri. 1980. «You Did Say 'Oral Interactive Discourse'? » *Applied Linguistics*, vol. 1, no 3, p. 189-200
- Hyland, Terry. 1997. «Reconsidering Competence». *Journal of Philosophy of Education*, vol. 31, no 3.

- Hymes, Dell. 1984. *Vers la compétence de communication*. col. LAL, Paris : Hatier-Credif, 220 p.
- Jennings, M., Fox J., Shohamy, E. Graves, B. 1999. «The effect of choice on test takers' scores in language tests». *Language Testing*, vol. 15, no 4.
- Jonnaert, Philippe. 2002. *Compétence et socioconstructivisme. Un cadre théorique*. Bruxelles : De Boeck.
- Kaplan, Isabelle. 1991. «L'Évaluation des compétences fonctionnelles en langue seconde», *Bulletin de l'AQEFLS*. No 1 p.124-137.
- Kormos, Judith et Mariann Dénes. 2004. «Exploring Measures and Perceptions of Fluency in the Speech of Second Language Learners». *System*, 32, p.145-164.
- Kramsch, C. (1986). «From Language Proficiency to Interactional Competence». *The Modern Language Journal*, Vol. 70, no 4, p. 366-372.
- Lado, Robert. 1962. *Language Testing. The Construction and Use of Foreign Language Tests. A Teacher's book*. London : Longmans, Green, 389 p.
- Laurier, M. (Dir.) 2003. *Évaluation et communication*. Outremont : Québecor, 306 p.
- Laurier, Michel et Denise Lussier. 2001. «Approches et pratiques en évaluation de la langue seconde au Québec». *Les langues modernes. Dossier : Évaluation et certification en langues, 2*, p. 62-70.
- Leblanc, Marie-Cécile. 2002. *Jeu de rôle et engagement*. Paris : L'Harmattan, 338 p.
- Leclerc, Jacques. 1989. *Qu'est-ce que la langue?* Laval : Mondia, 460 p.
- Le Huche, François et André Allali. 2001. *La Voix, tome 1. Anatomie et physiologie des organes de la voix et de la parole*. Issy les Moulineaux Cedex : Masson, 208 p.
- Lennon, Paul. 1991. «Error Elimination and Error Fossilisation : A Study of an Advances Learner in the L2 Community». *Review of Applied Linguistics*, vol. 93-94, (octobre), p.129-151.
- Leung, Constant. 2005. «Convivial Communication : Recontextualizing Communicative Competence». *International Journal of Applied Linguistics*, vol. 15, no 2, p.119- 144.
- Lumley, Tom et Tim F. McNamara. 1995. «Rater Characteristics and Rater Bias : Implications for Training». vol.12, no 1, p. 54-71.
- Lumley, Tom. 2002. «Assesment Criteria in a large-scale writing Test : what do they rallye mean to the Raters?». *Language Testing*, vol. 19, no 3, p.247-276.

- Lussier, Denise et Carolyn Turner. *Le Point sur l'évaluation en didactique des langues secondes*. Montréal : CEC, 1995, 257 p.
- McNamara, Tim. 1996. *Measuring Second Language Performance*. London : Longman, 344 p.
- Milanovic, Michael et Nick Saville (Dir.). 1996. *Performance Testing, Cognition and Assessment. Selected papers from the 15th Language Testing Research Colloquium*. Cambridge : Cambridge University Press, 313 p.
- Milanovic, Michael, Nick Saville, Alastair Pollitt et Annette Cook. 1996. «Developing Rating Scales for CASE : Theoretical Concerns and Analyses». Voir Cumming et Berwick. 1996.
- Mullen, Karen. 1978. «Direct Evaluation of Second-Language Proficiency : The Effects of Rater and Scale in Oral Interviews». *Language Learning*, vol. 28, no 2, p.301-308.
- North, Brian. 1995. «The Development of a Common Framework Scale of Descriptors of Language Proficiency Based on a Theory of Measurement». *System*, vol. 23, no 4, p. 445-465.
- North, Brian et Günther Schneider. 1998. «Scaling Descriptors for Language Proficiency Scales». *Language Testing*, vol. 15, p. 217-263.
- Norton, Julie. 2005. «The paired format in the Cambridge Speaking Tests». *Elt Journal*, vol. 59, no 4, p. 287-465.
- Oller, John W. 1983. *Issues in Language Testing research*. Rowley: John W. Oller, Jr, Editor, 387 p.
- Orr, Michael. 2002. «The FCE Speaking : Using Rater Reports to Help Interpret Test». *System*, vol. 30, p. 143-154.
- O'Sullivan, Barry. 2002 «Learners Acquaintanceship and Oral Proficiency Test Pair-Task Performance». *Language testing*, vol.19, no 3, (juillet), p. 277-295.
- Palmer, S. Adrian, Peter J. M. Groot et George A. Trosper. 1981. *The Construct Validation of Tests of Communicative Competence*. Washington: Teachers of English to Speakers of other Languages, 165 p.
- Perrenoud, Philippe. 1997. *Construire des compétences dès l'école*. Paris : ESF éditeur, 125 p.
- Pollitt, Alastair et Neil L. Murray. 1996. «What Raters Really Pay Attention to». Voir Milanovic et Saville. 1996.

- Québec, ministère de l'Éducation, Direction générale de la formation des adultes. 1994. *French second language*, 225 p.
- Québec, ministère de l'Éducation, Direction générale de la formation des adultes. 1998. *Français, langue seconde. Définition du domaine d'examen*, 13 p.
- Québec, ministère de l'Éducation, Direction générale de la formation des adultes. 1988. *Français, langue seconde. Guide d'élaboration des instruments d'évaluation sommative à l'éducation des adultes*, 140 p.
- Québec, ministère de l'Éducation, du Loisir et des Sports. 2006. *Programme de formation de l'école québécoise, enseignement secondaire, premier cycle*.
- Québec, ministère de l'Immigration et des Communautés culturelles. 2005. <http://www.immigration.quebec.gouv.qc.ca/francais/partenaires/Introduction.html>
- Québec, ministère de l'Immigration et des Communautés culturelles. 2005. *Tableaux sur l'immigration au Québec*.  
<http://www.immigrationquebec.gouv.qc.ca/francais/partenaires/Introduction.html>
- Québec, ministère de l'Immigration et des Communautés culturelles. 2005. *Tableaux dans le Plan d'immigration du Québec pour l'année 2006*, le gouvernement du Québec  
<http://www.immigration.quebec.gouv.qc.ca/francais/partenaires/Introduction.html>
- Romainville, Marc. 1996. «L'irrésistible ascension du terme «compétence» en éducation». *Enjeux*, vol. 38, no 37.
- Savignon, J. Sandra. 1972. *Communicative Competence: An Experiment in Foreign Language Teaching*. Philadelphia: The Center for Curriculum Development Inc, 116 p.
- Savignon, J. Sandra et Pavel V. Sysoyev. 2002. «Sociocultural Strategies for a Dialogue of Cultures». *The Modern Language Journal*, vol. 86, no 4 (hiver), p. 508-524.
- Saville, Nick et Peter Hargreaves. 1999. «Assessing Speaking in the Revised FCE». *ELT Journal*, 53, pp. 42-51.
- Shin, Sang-Keun. 2005. «Did they Take the Same Test? Examinee language proficiency and the structure of language tests». *Language Testing*, vol. 22, no 1, p. 31-57.
- Shohamy, Elana. 1994. «The validity of Direct Versus Semi-Direct Oral Tests». *Language Testing*, vol. 11, no 2, p. 99-123.
- Stroobants, M. 2000. «La qualification ou comment s'en débarrasser». Voir Dolz et Ollagnier.

- Tardif, Jacques. 2006. *L'évaluation des compétences*. Montréal : Chenelière Éducation. 361 p.
- Taylor, Lynda. 2000. «Approaches to Rating Scale Revision». *EFL Research Notes*, vol. 3 p. 14–16.
- Taylor, Lynda et Neil Jones. 2001. «Revising the IELTS Speaking Test». *EFL Research Notes* vol. 4, p. 9–12.
- Tseng, Yueh-Hung. 2002. «A Lesson in Culture ». *ELT journal*, vol. 56, no 1 (janvier), p. 11-21.
- Upshur, A. John et Carlyne E. Turner. 1995. «Constructing Rating Scales for Second Language Tests». *ELT Journal*, vol. 49, p. 3–12.
- Venema, James. 2002. «Developing Classroom Specific Rating Scales: Clarifying Teacher Assessment of Oral Communicative Competence». *JALT Testing and Evaluation SIG Newsletter*, vol. 6, no 1, p. 2-5.
- Wall et Alderson. 1996. «Examining Washback: The Sri Lankan Impact Study». Voir Cumming, et al, 1996, p.194-221
- Walter, F. Scott. 2004. «An Application of a Conversation Analysis to the Development of Test of Second-Language Pragmatic Competence». Thèse de doctorat, University of Illinois at Urbana-Champaign, 341 p.
- Wigglesworth, Gillian. 1993. «Exploring Bias Analysis as a Tool for Improving Rater Consistency in Assessing Oral Interaction». *Language Testing*, vol. 10, no 3, p. 305-336.