

Some pages of this thesis may have been removed for copyright restrictions.

If you have discovered material in Aston Research Explorer which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our [Takedown policy](#) and contact the service immediately (openaccess@aston.ac.uk)

A Genre-Based Approach to Speaking in EFL

Wilkins, Simon Gary

Doctor of Philosophy

Aston University

August 2018

© Simon Gary Wilkins, 2018, asserts his moral right to be identified as the author of this thesis.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.

Thesis Summary
 Aston University
 A Genre-Based Approach to Speaking in EFL
 Simon Gary Wilkins
 PhD Languages & Social Sciences
 August 2018

Speaking skills have been neglected in many language curricula and assessments at tertiary level, producing students who are viewed as structurally competent but communicatively deficient. An increasing emphasis on the communicative approach has given prominence to the importance of spoken language, but misinterpretations of communicative pedagogy and cultural resistance to these practices has meant that this approach has not always been successful. These issues are complicated further by the presentation of new approaches to language teaching as discrete packages that reject what has happened previously. A genre-based approach, derived from the principles of systemic functional linguistics, is offered as a comprehensive framework for incorporating a universal understanding of language teaching. This thesis asks in what ways a genre-based approach assists the development of Japanese students' speaking abilities. The theoretical underpinnings of the genre-based approach were used to design a syllabus focused on students' speaking skills. This syllabus intervention was applied to numerous English as a foreign language classes at the tertiary level in Japan. Both qualitative and quantitative data were collected and analysed during an action research process via an emic perspective, with qualitative data analysed inductively. Conclusions were based upon thorough triangulation of qualitative and quantitative data collected from a range of sources and at different times.

Findings of this study suggest that the genre-based approach provides a framework for addressing students' speaking skills in a principled and logical manner, allowing students to improve their skills in casual conversation, lengthening their spoken utterances, and reducing feelings of anxiety and frustration in speaking. This framework allows teachers to incorporate elements of both fluency and accuracy in their classrooms, whilst maintaining a primary focus on spoken communication. However, findings indicated that the efficacy of a genre-based approach relies heavily on a number of corresponding pedagogical factors. Firstly, a genre-based syllabus provides opportunities for integrating Assessment for Learning strategies, and it is vital to incorporate these strategies into syllabus design. The authenticity of texts is another key component of achieving desired results under such an approach. When selecting authentic texts, however, careful consideration is also needed to ensure that the notion of the "native speaker" is replaced with a concept of mutual intelligibility. Such cogitation is essential in order to bridge the gap between teacher expectations and student achievement. The necessity for student and teacher training in these various approaches, as well as their implementation in the classroom, demand considerable time and effort. Such an investment of resources must be considered before embarking on similar interventions, but findings suggest that this engagement is justified.

Keywords: systemic functional linguistics, action research, Assessment for Learning, authentic texts, mutual intelligibility

Table of Contents

Chapter 1: Introduction	9
1.1 Rationale	9
1.2 The University Context	11
1.2.1 The concept of “Native speaker”	12
1.3 The Japanese Context	13
1.4 Assessment	17
1.5 Focus of the Study	18
1.6 Research Design	19
1.6.1 Action research	20
1.6.2 Data collection	21
1.7 Thesis Outline	21
1.8 Summary	23
Chapter 2: Literature Review	24
2.1 Introduction	24
2.2 The Implementation of New Teaching Approaches	25
2.3 Trends in Language Teaching	24
2.3.1 The audiolingual method	27
2.3.2 Chomsky and linguistic competence	28
2.3.3 Grammar translation	28
2.3.4 Communicative language teaching	30
2.3.4.1 Problems with communicative language teaching	31
2.3.5 Assessment and entrance examinations	33
2.3.5.1 Alternative forms of assessment	35
2.3.6 Current issues in Japanese language teaching	36
2.3.6.1 The role of MEXT	37
2.3.6.2 Other considerations in the contemporary context	39
2.3.7 Summary of section 2.3	42
2.4 The Genre-Based Approach	43
2.4.1 Text and genre	43
2.4.2 Context of situation and context of culture	45
2.4.2.1 Context of culture	46
2.4.3 The role of grammar in a genre-based approach	48
2.5 From Theory to Practice	52
2.6 Register Theory	56
2.6.1 Field	56
2.6.2 Tenor	57
2.6.3 Mode	60
2.7 Cohesion, Theme and Rheme	63
2.8 Spoken and Written Language	64
2.9 Designing a Genre-Based Syllabus for Speaking	69
2.9.1 The role of grammar in a text-based syllabus	71
2.9.2 Cycles of language teaching and learning	73
2.10 Current Research	75
2.11 Alternative Approaches	76
2.12 Summary	78
Chapter 3: Methodology: Research Perspectives	80
3.1 Introduction	80
3.2 A Framework for Thinking about Research	80

3.2.1 Ontology	82
3.2.2 Epistemology	82
3.3 Research Perspectives	83
3.3.1 Positivism	83
3.3.2 Interpretivism	85
3.3.3 Critical theory	87
3.3.4 Feminism and post-structuralism	89
3.4 Why an Action Research Approach?	90
3.4.1 My research experiences	95
3.4.2 Limitations of the action research approach	98
3.4.3 Paradigms for reporting action research	99
3.4.4 Limitations of this study	100
3.5 The Participants	102
3.5.1 Students' spoken proficiency	104
3.6 Summary	106
Chapter 4: Cycle One: Planning and Action	108
4.1 Introduction	108
4.2 Planning: The Research Question	109
4.3 Planning: Syllabus Design	110
4.4 Ethical Considerations	117
4.5 Action: Data Collection	119
4.6 Qualitative Data Collection Points	120
4.6.1 Classroom documents	121
4.6.2 Lesson plans	121
4.6.3 Teacher and student journals	122
4.6.4 University-wide questionnaires	122
4.6.5 Anonymous feedback	122
4.7 Quantitative Data Collection Points	123
4.7.1 The assessment instrument	123
4.8 Reflection-in-Action	124
4.8.1 The model text	125
4.8.2 Turn-taking	125
4.9 Preparing for Data Analysis	127
4.9.1 Qualitative data analysis	127
4.9.2 Quantitative data analysis	134
4.9.2.1 The inclusion of Rasch analysis in action research	137
4.9.2.2 Software: Winsteps and Facets	141
4.10 Presenting the Results	142
4.11 Summary	142
Chapter 5: Cycle One: Observation and Reflection	143
5.1 Introduction	143
5.2 Observation: Undertaking Data Analysis	143
5.3 Qualitative Data Analysis	143
5.3.1 Student emotions	144
5.3.2 The model text	150
5.4 Quantitative Data Analysis	156
5.5 Post-Syllabus Student Dialogue	166
5.6 Reflection	168

Chapter 6: Cycle Two: Planning and Action	171
6.1 Planning: Syllabus Design	171
6.2 Assessment for Learning	171
6.2.1 Assessment for learning in the Asian context	172
6.2.2 Issues in assessment for learning	175
6.2.3 Integration of assessment for learning into the syllabus	176
6.3 The Model Text	180
6.3.1 Subtitled videos	183
6.3.2 Storyboarding	184
6.3.3 Turn-taking	184
6.3.4 Collaboration	185
6.3.5 Self-assessment	186
6.4 Ethical Considerations	190
6.5 Action: Data Collection	190
6.5.1 Data from collaborative researchers	190
6.5.2 Student self-assessment and reflection	192
6.6 Preparing for Data Analysis and Presentation	194
6.7 Summary	195
 Chapter 7: Cycle Two: Observation and Reflection	 196
7.1 Introduction	196
7.2 Student Emotions	196
7.3 Cross-Referencing Student Emotion with Quantitative Data	201
7.4 The Model Text	205
7.5 Cross-Referencing Qualitative and Quantitative Data on the Model Text	210
7.5.1 Model texts: Reflection-in-action	212
7.6 Assessment	213
7.6.1 Assessment Results	214
7.7 Cross Referencing with Qualitative Data	219
7.8 Impact on Teaching	222
7.7.1 Grammar teaching	223
7.7.2 Syllabus design	227
7.7.3 The teacher's role in the classroom	227
7.7.4 Issues with the syllabus	230
7.8 Post-Syllabus Student Dialogues	232
7.9 Reflection	234
7.9.1 Reflections on student emotion	234
7.9.2 Reflections on the model text	237
7.9.3 Reflections on assessment	238
7.9.4 Reflection on the impact on teaching	240
 Chapter 8: Final Reflections and Future Planning	 242
8.1 Introduction	242
8.2 Summary of Findings	242
8.3 The Research Question	246
8.4 Limitations of the Findings	248
8.5 Implications for Language Teaching	250
8.6 Future Planning	251

References	254
Appendices	264
Appendix 1: Activity to raise awareness of differences between genres	264
Appendix 2: Syllabus entry form	265
Appendix 3: Study plan distributed to students in the first class	267
Appendix 4: “Greetings” worksheet used to explore context and tenor	268
Appendix 5: Series of worksheets building the concept of “orientation”	269
Appendix 6: Temporal conjunctions worksheet	273
Appendix 7: Asking and answering questions worksheets	274
Appendix 8: Lesson plan with teacher’s notes	276
Appendix 9: Questions to elicit content for the teacher and student journals	277
Appendix 10: Teacher journal entry 4 – 29 April 2013	279
Appendix 11: Student journal (undated)	283
Appendix 12: Spoken assessment sheet	284
Appendix 13: Amended recount genre with criteria illustrated for students	285
Appendix 14: Data control file for Winsteps with example of raw data	286
Appendix 15: Example of analytic rubric negotiated with students	287
Appendix 16: Holistic rubric developed from the Common European Framework	288
Appendix 17: Example of text deconstruction via a storyboard	289
Appendix 18: Examples of worksheets designed to address difficulties in the micro-aspects of conversation	290
Appendix 19: Reflective questions for teachers	294
Appendix 20: Rubric for assessment during assessment exercise	295
Appendix 21: Example of analytic rubric negotiated with students	296
Appendix 22: Self-assessment sheet completed by student	297
Appendix 23: Cycle Two: post syllabus student recount texts	298

List of Tables

Table 2.1 Classroom types and their corresponding syllabus type. Adapted from Feez and Joyce (1998)	40
Table 2.2 The major differences between traditional and functional grammar, developed from Martin and Rothery (1993)	50
Table 2.3 Characteristics features of three genre types. Adapted from Butt et al. (2000)	54
Table 2.4 Process types identified by Butt et al. (2000)	55
Table 2.5 Mood adjuncts in functional grammar	60
Table 2.6 The main types of cohesive device, from Collerson (1994) and Gibbons (2001)	63
Table 2.7 Characteristic features of spoken and written language (Eggins, 2004, p. 93)	65
Table 2.8 Speech roles and commodities in interaction (Eggins, 2004, p. 146)	68
Table 2.9 Five steps in designing a genre-based syllabus, developed from Joyce and Burns (1999)	70
Table 4.1 Deconstruction of a recount text to inform syllabus design.	116
Table 4.2 Data collection instruments.	120
Table 4.3 Description of criteria for assessment.	124
Table 4.4 Description of additional criteria for assessment based on the Common European Framework.	127
Table 5.1 Hypothesized latent variable map for the generic structure and textual features of a recount genre.	158
Table 5.2 Person correlation and fit.	162
Table 6.1 Summary of the integration of assessment for learning into the syllabus.	177
Table 6.2 Profiles of persons used as model texts and to inform the 5-Point Likert scale for pronunciation and fluency	193
Table 7.1 Person fit statistics.	201
Table 7.2 Rater fit statistics.	216

List of Figures

Figure 1.1 Portion of an advertisement on the JREC-IN website for a teaching position in EFL (14 January 2017)	13
Figure 2.1 Generic structure of recount genre, adapted from Butt et al. (2000)	53
Figure 2.2 The continua of power, contact and effect (Eggins, 2004, p. 100)	58
Figure 2.3 Representation of interpersonal distance (Eggins, 2004, p. 91)	61
Figure 2.4 Representation of experiential distance (Eggins, 2004, p. 91)	62
Figure 2.5 Stages of the teaching/learning cycle from Callaghan, Knapp and Noble (2014)	74
Figure 4.1 Generic structure of the recount genre (adapted from Butt et al., 2000)	115
Figure 4.2 Summary of qualitative data analysis in the study based on suggestions by Bazeley (2009) and Burns (2010)	129
Figure 4.3 Person / Item / Rater map showing persons on common scale with items and raters	139
Figure 5.1 Matrix display to examine patterns in student emotions	145
Figure 5.2 Matrix display to examine patterns in response to the model texts	151
Figure 5.3 Person-Item map showing persons on common scale with items.	160
Figure 6.1 Student based model text with subtitles	183
Figure 7.1 Matrix display to examine patterns in student emotions	197
Figure 7.2 Matrix display showing reactions to the model texts	206
Figure 7.3 Progression of text authenticity	206
Figure 7.4 Person / Item / Rater map showing persons on common scale with items and raters	211
Figure 7.5 Person-Rater-Item map showing persons on common scale with items and raters	215
Figure 7.6 Chart showing bias/interaction between raters and students	217
Figure 7.7 Matrix display to examine themes of assessment	220
Figure 7.8 Matrix display to examine the impact on teaching	223
Figure 7.9 Student emotions and formality of talk in the classroom context	235
Figure 8.1 Model summarizing some of the main findings of the research	245

Chapter 1: Introduction

1.1 Rationale

The impetus for this study stemmed from the demands of my immediate teaching context in Japan, and perceived inadequacies in the teaching and learning of students' speaking abilities in my English as a foreign language (EFL) classroom. The institution in which the study takes place has a compulsory English language component for all students who enter the university, regardless of their course of study. Over time, the incoming cohort of students has had varying levels of English proficiency. With the demands of a falling Japanese population (Coulmas, 2007), and a desire for universities to maintain enrolment numbers, the number of students with a low English proficiency has increased considerably. A growing number of students taking compulsory English language courses are unable to speak in English beyond one or two-word utterances and memorized formulaic expressions. This situation places great demands on the teachers who are responsible for identifying and sequencing classroom content appropriate to students' needs, and who must also ensure that students' speaking abilities progress in a principled manner; and moreover, that feedback is provided that promotes continued and independent learning.

Within this context, there is no identifiable curriculum to assist these goals, beyond the title of the compulsory English class, "English Conversation". Individual teachers are entirely responsible for syllabus design, and the only form of assessment is the *TOEIC Bridge™* test provided at the start and end of each semester. TOEIC Bridge is presented as a preparatory test for low-proficiency students who might not yet have the required level of English proficiency that is measurable under the scoring system of the full TOEIC test, due to a flooring effect (Fryer et al., 2014). The TOEIC Bridge test does not include any form of spoken criteria and is entirely receptive in nature. Research undertaken (Fryer et al., 2014) within my own university also found that students who were undertaking compulsory English

classes had very low motivation to study English and low proficiency, with a mean TOEIC Bridge score of 117 out of a maximum possible score of 180.

Within this context, I had great concerns about whether or not I was assisting my students' speaking abilities in my class, and if my schemes of work were meeting their specific needs and goals. I received feedback from students in the form of generic institution-wide questionnaires at the end of each semester, and scores from the TOEIC Bridge test; but due to the tests' receptive nature and non-specific feedback from students, I felt this feedback was not adequate. I felt my classes lacked a principled and coherent approach to developing students' speaking abilities, with a reliance on textbooks; this made the justification for the selection and sequencing of work hard to identify. I also felt that assessment procedures were not helping me gain appropriate feedback on my students' speaking abilities, nor assisting my ability to provide appropriate feedback for their continued learning, with identifiable and relatable goals for that learning.

Whilst studying for my master's degree, it was clear to me that the literature presents varying criticisms of the way in which the communicative approach has been adopted in the English language classroom; it also indicates the neglect or abandonment of alternative teaching practices altogether. These arguments will be discussed in greater detail in Chapter 2. Of particular interest to me during my study was the methodology behind the genre-based syllabi. This methodology was identified as part of the communicative approach, stemming from a Systemic Functional Linguistics viewpoint. The methodology addressed aspects of both fluency and accuracy in a principled way, and it included assessment as an integral part of syllabus design. During my master's study I had adopted this approach in my writing classes, and felt that I had achieved considerable success (Wilkins, 2006). I began to explore ways in which I might adapt these principles to a speaking class. In reading the literature, I found promising examples of how the genre-based approach had been applied to casual

conversation. However, it was not an approach that seemed directly applicable to my immediate context, namely that of EFL, low-proficiency Japanese students in compulsory English education, with low motivation and confidence. If a genre-based approach is to play a greater and more accepted role in English-language classrooms, then further research is necessary with regard to its pedagogical applications and efficacy in a variety of different contexts. The current study aims to expand on the research in this way, and to hold resonance for teachers in similar teaching contexts and beyond.

Section 1.1 builds the rationale behind this study and introduces briefly the context in which the study takes place. The participants in this study will be introduced in greater detail in Chapter 4; however, as part of the discussion of the rationale behind this study, a further exploration of the research context will follow, in order to highlight the reasons for some of the decisions that led to the specific research design that is undertaken in this thesis.

1.2 The University Context

This study was conducted at Kyushu Sangyo University, a four-year technology-focused university in Fukuoka Prefecture, southern Japan. Admission to public universities in Japan is primarily based on the national entrance exam, and universities are ranked by “Hensachi”, depending upon the scores of the student in the test. The average score in the test is 50 points. Kyushu Sangyo University has an average score of 42, which gives it one of the lowest rankings of the 53 universities in Fukuoka. With an average TOEIC Bridge score of 117, students would be considered of low English proficiency. Whilst the university has a large student body of around 12,000, there are no English majors; however, every student is required to attend a compulsory English language course for at least two of their four years of study. This compulsory English course is conducted via the Language Education and Research Centre (LERC), where I work and where this study is conducted. Students undergo two 90-minute classes of instruction per week.

Perceived gaps in the current curriculum at the institution have led to the English department brainstorming ideas for new speaking syllabi that would assist students to develop greater fluency in spoken performance; and in the future, for a potential university-wide curriculum to address perceived inadequacies in English language instruction and learning. Although English is a compulsory part of the university curriculum, it is integrated into vocational and technological subjects that take precedence in overall curriculum demands. These factors must be considered when approaching the research design in this study. There is also an over-arching Japanese context to this study that must also be considered when approaching the research design and its rationale.

1.2.1 The concept of “Native speaker”

During this thesis, repeated reference is made to the concept of the “native speaker”. However, this description is not used as a measurement of speaking proficiency, but it is an officially designated title within the institution in which the study takes place. “Native speakers” are required to teach specific types of classes, namely “English conversation” classes, which are nominally intended as classes for improving a student’s fluency in English conversation, and are conducted entirely through the medium of English. While these designations are not explicitly stated in any form of unified curriculum, or in syllabi, the notion is widely agreed upon. Meanwhile, “Reading and Writing” classes are taught by Japanese teachers: it is assumed most of these classes will take place in Japanese, thus immediately creating a dichotomy between fluency and accuracy, and between “native” and “Japanese” teachers. The prevalence of this notion of “native speaker” is observable on the JREC-IN website, which advertises tertiary teaching jobs in Japan. At the time of writing, all of the top jobs on the EFL section of the site make mention of, or list as a direct qualification, the title of “native speaker”, as can be observed in the screenshot below:

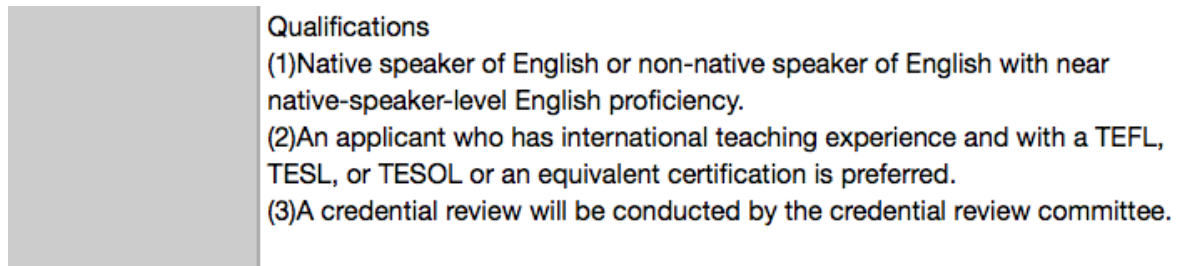


Figure 1.1 Portion of an advertisement on the JREC-IN website for a teaching position in EFL (14 January 2017)

The concept of the “native speaker” is hotly contested in the field of language education. Some researchers argue that the involvement of a “native speaker” is a criterion for “authentic English” (Harmer, 2007). Clark (2013), on the other hand, argues that ideas of a superior “native speaker” or standard English are weakening, and also that ideas of standard English are not linguistically or scientifically tenable. She argues that notions of a “native speaker” might inhibit success in communication. This designation of “native speaker” has profound implications for language teaching in Japan, which will be discussed in greater detail later in the results chapters of this thesis.

1.3 The Japanese Context

The desire for an increased focus on spoken output amongst the Japanese has been increasingly emphasized in Japan. Recent reports from the Japanese Ministry of Education, Science and Technology and Education (MEXT) on the Reform of English Teaching Methodology (1947, 1998, 1999, 2003, 2005, 2007, 2008, 2009, 2010 and 2011) have highlighted the need for improved speaking skills rather than a continued focus on written output. As early as 1947, MEXT’s first course of study guidelines for foreign language teaching stated that listening and speaking were the primary skills (Tahira, 2012). In April 2011, for the first time, instruction in English language communication became compulsory for fifth and sixth-grade elementary-school students. Also noteworthy were the requirements that for senior high schools, reading and writing courses were to be removed and replaced with “English Communication (I, II)” and “English Expression (I, II)”, with the position that

“grammar instruction should be given as a means to support communication” (MEXT, 2009, p. 7). The policy seems to reflect the inclusion of compulsory “English conversation” classes within the institution in which the current study takes place. Changes in policy can be seen as an attempt to move towards more communicative approaches, in order to reduce the dominance of the more traditionally used grammar translation method.

Goh and Burns (2012) argue that teaching speaking is a component of language learning that is neglected in many language curricula and assessments, and also that there is a difference between teaching speaking and actually speaking, as is typically taught in many language classrooms. A lack of focus on oral communication has led to what McDonough and Shaw (1993, citing Johnson, 1981) describe as the production of students who are structurally competent but who are often communicatively incompetent.

One reason for this lack of oral focus is proposed by Rapley (2008), who explains that although English-speaking skills are considered important in Japan, entrance examinations at the senior high school and university levels exert the greatest pressure on Japanese teachers of English, and as such produce teaching models that are not in accordance with the intentions of the MEXT reports. Rapley describes these “traditional Japanese methods” (p. 1) as focusing on elements of grammar and translation that are not conducive to productive language use.

There have been various moves towards addressing these problems in Japan: for instance, speaking assessment has been included in tests that are widely used in the Japanese context, such as TOEIC, the Standard Speaking Test, or the STEP-Eiken. Perhaps these modifications are meant to address the current lack of spoken English in these types of examinations. Japan has also seen an increasing emphasis on the communicative approach, which aims to emphasize the importance of spoken language (McDonough and Shaw, 1993). McDonough and Shaw (p. 22) define the “communicative approach” as being “for students

interested in using language rather than learning more about structure”. However, misinterpretations of communicative pedagogy (Mitchell and Lee, 2003) and cultural resistance to these communicative practices (Hu, 2002; Gorsuch, 2000) have meant that the approach has not always proved successful in Asian contexts. In environments of resistance and misunderstanding, the communicative approach may not lead to fluency, but rather to “desultory silence” (Scrivener 1994, p. 1), as teachers and students who attempt to develop speaking fluency in their classes fail to grasp the perceived need to speak, and thus lack the motivation or necessity to produce talk.

Recent evidence regarding the adoption of communicative approaches in Japan shows varying levels of success in their impact on learning and teaching. MEXT (2010) conducted a survey in 3,598 Japanese high schools, and concluded that most teachers who had claimed to be adopting Communicative Language Teaching approaches were primarily using Japanese in classrooms during oral communication courses together with students; this suggests they might not have been using materials appropriately. Moreover, a study in Japan by Sakui (2004) observed that the majority of class time in “communicative classes” was dedicated to teacher-fronted grammar explanations, chorus reading and vocabulary explanations. She reported that what might be categorized as more communicative activities, such as information gaps, played a much smaller role. This view was mirrored by Nishino (2011), whose research in 139 Japanese high schools revealed that teachers routinely failed to employ communicative activities, even though they held positive beliefs about Communicative Language Teaching. A major reason for these failings may be a misunderstanding of communicative approaches. According to Brown (2007), a major difficulty that has prevented communicative approaches from taking root in Japanese public schools is that it is a methodology with many interpretations and manifestations, and that teachers’ perceptions of communicative activities are varied and ambiguous. Tahira (2012) also points to a misunderstanding of MEXT guidelines. She claims that the MEXT guidelines are “obscure”

(p. 6), and that teachers receive little or no support and training for operationalizing the stated guidelines. Fennely and Luxton (2011) add that the newly introduced activities for elementary schools are not well understood, and this could lead to very serious problems in the future, regarding what MEXT expects in schools and what is actually being delivered.

As previously stated, part of this misunderstanding or resistance to communicative approaches may be cultural. McDonough and Shaw (1993) describe the communicative approach as polarizing “function versus grammar”, which they say is “unbalanced and incorrect” (p. 22). Joyce and Burns (1999, p. 1) echo these sentiments by stating that, “over the decades of the twentieth century grammar has moved at various times from being central in syllabus design to being eliminated altogether”. A traditional focus on accuracy and grammar may not be something that Japanese language teachers are confident about omitting entirely from their lesson planning.

A persistent and repeated desire for increased spoken proficiency among Japanese students is evident from MEXT, but the reality does not seem to reflect these goals. What can be concluded from the discussion so far is that there is a clear desire to improve the speaking abilities of Japanese students, and that a perceived gap in the current educational context prevents this from happening. However, closing this gap is hindered by factors such as cultural resistance, established teaching practices, the lack of appropriate training, and a misunderstanding of the communicative approach. Furthermore, language assessment procedures throughout the Japanese education system promote traditional teaching models based upon grammar and translation (Matsuura, Chiba and Hildebrandt, 2001). Many teachers and administrators may fear omitting traditional approaches to language teaching, due to the perceived receptive nature of university entrance exams and tests, such as TOEIC. These various problems highlight the need for an alternative or modified approach to English language teaching in Japan. Such an approach would need to develop oral fluency, while also

addressing cultural and pedagogical objectives that aim to improve accuracy and grammar. Could a genre-based approach that integrates both grammar and fluency provide a possible response to the issues of teaching methodology and spoken language assessment? A communicative approach that includes explicit grammar instruction could face less cultural resistance. This questions will be explored in more detail in Chapter 2. A further contextual aspect for the study to consider is the role of assessment in Japan, as this will influence the design of the syllabus to be integrated into the research design of this study.

1.4 Assessment

At present, the only formal assessment and measurement of students' communicative language at the current institution is conducted through compulsory TOEIC Bridge tests, which are integrated into individual syllabus designs by necessity, as they form 20% of students' final grades. These tests are conducted before enrolment at the institution, and once yearly thereafter. TOEIC Bridge and the corresponding TOEIC test have a very high status in Japan, with businesses and industry often requiring specific TOEIC scores as conditions of employment (Kubota, 2011). The institution involved in the present study uses TOEIC scores to measure student achievement, and even to conduct teacher evaluations.

TOEIC Bridge is an entirely receptive test, however, as it contains only listening and reading sections. There is no assessment of students' writing or speaking abilities, and no other means within the university's current curriculum of assessing the productive ability of students. Because of these limitations in the TOEIC and TOEIC Bridge test, it was necessary to create alternative forms of language assessment in this study, in order to inform cycles of action research, and to provide targeted feedback to students for their learning.

The current assessment approach at the university as a whole favours traditional approaches of grammar translation, which do not require spoken output. Brown (2003, p. 19) highlights the need for the adequacy of construct definition in assessing second-language

communicative competence. She defines communicative competence as “an abstraction that is rarely defined with any precision in terms of actual test performance”. Bachman (1990, p. 50) states that in order to:

... maximize the reliability of test scores and the validity of test use, we should . . . provide clear and unambiguous theoretical definitions of the abilities we want to measure and specify precisely the conditions, or operations that we will follow in eliciting and observing performance.

Cumming et al. (2002) remark that in their research, teachers used 27 different types of decision-making processes while evaluating a single written composition. As previously stated, the communicative approach is often ambiguous or misunderstood by teachers, which appears contrary to the demands of Bachman. It may be the abstract nature or interpretation of the communicative approach itself that has caused these issues. An approach that defines communicative competence more succinctly is required in any study that hopes to develop its findings beyond the immediate context, and to find resonance and acceptability amongst other teachers in similar situations. The demands of the LERC, and the limited number of teaching hours that students experience, require an integrated assessment procedure that encourages independent learning, assessment that is unambiguous, and assessment that is based upon sound theoretical definitions. Each of these criteria has informed the research design in this study.

1.5 Focus of the Study

So far, the introduction has highlighted that there is a pan-Japanese demand for increasing students' speaking abilities. However, assessment tests, established teaching practices and cultural norms that proliferate throughout Japan, and are still highly regarded by employers, have created an environment that that could counteract the achievement of these aims. Therefore, this study required a syllabus and research design that considered each of these factors.

As stated as a concern in 1.1, materials available to teachers may not always be appropriate for the teaching of speaking. As Slade and Widin (2004, p. 2) note, “nearly all the textbooks and materials available are based on written English and are not adequate for teaching speaking”. They highlight the necessity to find new and innovative ways of addressing the teaching of spoken language. In the context of this research study, namely Kyushu Sangyo University, the demand for increased student spoken output has become increasingly pressing, along with the need to find new approaches to the teaching of speaking. It is for these reasons that the study focusses on the development of an alternative syllabus by adopting a genre-based approach.

There are various arguments for a focus on genre-based syllabus design as an alternative approach. Feez and Joyce (1998), for example, argue that what has been missing in syllabus design is an approach that provides a framework within which present knowledge about language and language teaching can be organized to allow teachers to survey and analyse available syllabus elements, and to select, sequence and integrate elements into coherent, cohesive and comprehensive course design. Feez and Joyce propose that the genre-based approach provides a framework for integrated syllabus design because it focuses on the following: vocabulary and grammar; formulaic elements of simple exchanges in certain settings; whole texts and genres; topics and notions as a framework for planning; and knowledge of context that can be recycled for use in the ensuing contexts that are studied. By providing opportunities for the development of a syllabus with elements of fluency and accuracy, and with a focus on grammatical and structural elements of language use, traditional Japanese teaching methods could be used in conjunction with more communicative approaches, and could thus perhaps reduce some of the cultural resistance to these approaches. Teachers can also produce materials that address their concerns about entrance examinations, by basing syllabus design on whole texts and genres that mirror those

in the tests, or which contain grammatical and structural elements that are also found within the entrance examinations.

Feez and Joyce (1998) also explain that this approach allows opportunities for negotiation with learners about the design of the syllabus, together with opportunities for detailed explanation of the content and strategies to be used for monitoring achievement and providing effective intervention.

Specifically, the study will address the research question:

In what ways can a genre-based approach assist the teaching and development of Japanese students' speaking abilities?

1.6 Research Design

The previous sections outlined the impetus for this study, as well as the main issues embedded in the research problem, and how this problem fits into the wider scheme of second-language teaching in the wider context. The adoption of a syllabus adopting a genre-based approach in this context allows an opportunity to explore the ways in which these approaches may be able to benefit students' learning of speaking. This enquiry will be based upon data collection, the setting up of a database, and subsequent analysis of these data: these are the three key components of research, as identified by Wallace (1998). This section will briefly summarize the approach to research design that will be discussed in greater detail in Chapters 3 and 4.

1.6.1 Action research

This study locates itself within an action research paradigm. Burns (2011) describes action research in the English language classroom as “problematising” (p. 2) teaching. The teacher then becomes an investigator within their own personal teaching context, intervening in a deliberate way based upon systematically collected data. Action research was chosen as a

methodological approach in order to generate resonance with other teachers seeking practical knowledge of pedagogical value, building upon the work that has already been conducted, but primarily in an ESL context. It has also allowed me to address the concerns I highlighted earlier in the chapter in a systematic way, based upon data collection and analysis, rather than the existing intuitive feelings I currently had about my classrooms.

1.6.2 Data collection

Data were collected and analysed using both qualitative and quantitative techniques. Information was collected via classroom documents, student feedback, classroom observations, assessment tasks and a reflective journal.

Classroom documents included all lesson plans and worksheets used in class. Students were also asked to note reflectively what they thought about the new activities. Assessment procedures also asked pupils to provide self and peer evaluations. In this way, it is hoped that a balance between data collection and teaching was maintained. Video and audio recordings of students' use of English in the classroom were also utilized. Lesson plans for each stage of the syllabus included space to make detailed observations during class time, which addressed the research questions. A reflective journal also recorded observations and feelings after the event. Student feedback included their own reflective journal based on what was studied, as well as anonymous feedback provided at the end of the semester.

The range of data collection and analysis is designed to increase the objectivity of observations through triangulation. In this way, data can be compared and crosschecked to reach valid conclusions based upon numerous sources of information.

1.7 Thesis Outline

This section outlines the subsequent chapters of the thesis. Chapter 2 provides a review of literature relevant to the research questions, and to the syllabus design that incorporates both

the genre-based approach and assessment for learning. The main studies relating to these approaches and their theoretical underpinnings will be discussed. Gaps or shortcomings in the current body of literature will also be highlighted.

Chapters 3 and 4 will discuss in more detail the research methodology underpinning the study. The ideological choices regarding research methods and types of data collection and analysis will be discussed, with their appropriateness to the research question established. Data collection tools and database techniques will be described, with a summary of how themes were identified.

Chapters 5, 6 and 7 will present the analysis and synthesis of the data throughout the action research process. Through the use of student feedback, teacher observation and assessment procedures, the ways in which the genre-based approach was experienced by both the teacher and the students will be presented. Qualitative data will be presented from an emic, or insider, viewpoint through a process of inductive coding. Quantitative data drawn from assessment questions will be analysed to produce numeric conclusions, in order to compare, contrast or develop qualitative findings further. Rasch analysis will provide information on elements of the syllabus or assessment procedures that students found difficult or confusing. Correlations of assessment tasks will help determine the gap between the aims of the syllabus and what students understood in class. The use of averages and percentages in data collected via assessment tasks and questionnaires will also provide further opportunities for triangulation with other data sets, with regard to how the genre-based approach has assisted students to speak and to assess that speaking.

Chapter 8 will collate the findings of the thesis and address the research questions identified in this chapter. The implications of the study will be discussed. Based upon the information collected, suggestions for ways of implementing change at a range of levels in Japan's education system will be proposed, in order to assist the development of Japanese

students' spoken output, and effective ways of assessing that output. Chapter 8 will also discuss the study's limitations and possible future research enquiry. The thesis will also be summarized in order to illustrate its contribution to the field of applied linguistics.

1.8 Summary

This chapter has introduced the thesis by first describing the rationale behind the research, formed by the immediate teaching situation. Concerns about current teaching practice, and whether current teaching practice was meeting students' specific needs in developing speaking abilities, were also discussed. These concerns were reflected in the wider educational context, in which there was a growing insistence on the development of students' speaking skills, but a lack of evidence showing that such a desire was being fulfilled. In many cases, research showed that communicative language approaches were either being incorrectly applied or resisted, due to assessment practices or cultural traditions. A genre-based approach was mentioned as a possible remedy to the cultural resistance to communicative approaches, as this would provide a framework for explicit grammar instruction, as well as allowing activities that focused on fluency. The overall contexts, both in Japan as a whole and within the institution in which the study is undertaken, raised issues of research design and syllabus implementation. Assessment in particular was a component of syllabus design that would need to be addressed in the thesis. Action research was mentioned briefly as the approach underpinning the research design, which will be further developed in Chapter 3. Finally, an outline of the final thesis was provided. In the next chapter, a literature review is presented which overviews the major theoretical and conceptual frameworks that underpin the choice of the genre-based approach in this study.

Chapter 2: Literature Review

2.1 Introduction

This chapter overviews the major theoretical and conceptual ideas that underpin the choice of the genre-based approach in this study. Firstly, the historical context of language teaching in Japan will be discussed, in order to situate the current study within the Japanese educational system. Theoretical or ideological viewpoints, including different theoretical assumptions within this education system, will also be explored.

Secondly, the main theoretical background to the genre-based approach will be discussed in order to conceptualize the study and provide definitions in use. Thirdly, the challenges posed by the conceptualization of this theory as a practical pedagogical approach will be examined.

Following this, the major differences between spoken and written language will be analysed, in order to explain how an approach that was originally concerned with written output can be adapted to inform syllabus design and criteria generation for spoken output. Current theories and practice in utilizing the genre-based approach to speaking will be examined, with a focus on calls for follow-up studies, or gaps in the literature that are evident.

2.2 The Implementation of New Teaching Approaches

Before defining the genre-based approach, it is important to consider what came before in the Japanese education system. Examining this historical context will help highlight challenges, or lessons from previous approaches, that could inform the implementation of a genre-based approach in a wider context beyond this research. Some historical context will also help situate the current study and identify approaches that could be adopted into a genre-based

methodology in order to limit the effects of cultural resistance, or to build on the teaching skills, experiences and expertise that already exist.

Despite the efforts of MEXT to introduce new communicative methodologies into language teaching practices, many challenges are still evident. Feez and Joyce (1998) describe how new approaches to language teaching are often presented as discrete packages which challenge the legitimacy of what came before, and that this in turn has three main negative consequences: people assume that language teaching is based on fads and fashions rather than an evolving body of knowledge; teachers are divided between those who support the latest approach and those who reject it; and finally, that teachers lose access to valuable aspects of approaches which have gone before. Coupled with this issue in Japan is the prevalence and primacy given to traditional testing, which Rapley (2008) describes as exerting pressure to rely on traditional teaching practices such as grammar translation. Both of these factors mean that it is extremely difficult to initiate change in traditional approaches. However, this study aims to initiate change in the current teaching practice: firstly in the immediate context of the research as part of an action research process; and secondly, as research that resonates with teachers on a wider scale. At this stage, it would be useful to outline the different teaching approaches that have led to the current Japanese English-language teaching situation, before exploring the implementation and definitions of a genre-based approach within the literature, and how it will fit into this historical context.

2.3 Trends in Language Teaching

Tahira (2012) provides a useful summary of MEXT study guidelines for foreign languages since their formation in 1947. Following the Second World War, rapid economic expansion generated educational goals of attaining knowledge from Western culture through interaction with English speakers. The MEXT guidelines of 1947 stated four main objectives in language learning. Tahira (p. 1) describes these as: (1) “habit formation”, a term commonly associated

with “language drills”; (2) listening and speaking as primary skills; (3) imitations of English utterances; and (4) a focus on sounds and rhythm. These goals seem to mirror closely the audiolingual method of language teaching.

It is important to discuss the difference between the words “method” (i.e. the audiolingual method) and “methodology”. Richard et al. (1985) describe the difference by stating that “methodology” is the study of the principles and practice of language teaching, of what happens in the classroom, and the various teaching methods. Methodology therefore refers to both the principles of teaching and to what the teacher does in class, regardless of the principles that inform the teacher’s actions.

On the other hand, Richard et al. (1985) describe “method” as the attempt to find a single way of learning that suits all students, arising from the belief that there is one answer to the question of how students learn a new language. Based upon these definitions, the aim of a teaching method appears to be to prescribe what happens in the classroom, to specify how a teacher and students should behave. Thus, teaching methods attempt to influence classroom practice, rather than emerge from it. This is in contrast to “methodology”, which tends to suggest that what happens in the classroom should inform how learning takes place. The MEXT guidelines as outlined by Tahira appear to mirror trends and “methods” for language teaching, thus creating the problem outlined by Feez and Joyce (1998), of “discrete packages” of learning methods. This definition is important because in this environment teachers may be reluctant to adopt new methods; firstly because their current method is what they know well, and secondly due to a fear that a new method may soon come to contradict again everything that has happened before. This discussion has implications for this study, as previous methods must be considered, as well as ways of reducing resistance to new forms of learning.

2.3.1 The audiolingual method

The audiolingual method is based on the underlying theory of Behaviourism. Brown (2007) provides a summary of this method. The method became prominent due to changes in the field of psychology in the 1930s and 1940s, as well as political events such as the Second World War. Behaviourist theory, through the work of theorists such as Skinner (in Brown, 2007), led to stimulus-response theory, which is the notion that there is a direct link between action and reaction which played a part in the formation of the audiolingual method. The Second World War instigated a demand for soldiers, spies and diplomats who were competent in foreign languages. This audiolingual method can be characterized as follows:

- Substitution drills, lots of repetition
- Teacher does not explain grammar or language points
- Teacher corrects all mistakes
- Teacher controls what students say
- Only L2 used
- Emphasis on dialogues
- Orally-based, no writing at first, and reading done mainly in order to practise speaking
- Use of tape recorders meant students could study alone or without a teacher

The MEXT guidelines of 1947 appear to be paying attention to global trends of language teaching by adopting guidelines that assist the audiolingual method. This following of global trends is repeated again in the guidelines of the MEXT reports in the 1960s. Tahira (2012) attributes the prevalence of the grammar translation method, *yakudoku*, to the MEXT guidelines written in the 1960s that emphasized grammar rules and language structures. I argue that this adaption happened due to the Chomskyian notion of linguistic competence,

and that grammar translation had its roots in older approaches, but that these older approaches were adopted in response to Chomsky.

2.3.2 Chomsky and linguistic competence

Noam Chomsky (1967) stated that there were two aspects of human language: competence and performance. Competence related to the underlying rules of structures, while performance was what people did with language. Chomsky felt that linguistic competence was the more important of these. Emphasis was therefore placed on the mastery of language structure and the manipulation of grammatical forms. Chomsky believed that a native speaker has a set of grammar rules and that there are a finite number of these rules, within which it is possible to create an infinite number of sentences. Brown claims that although Chomsky's work did not give rise to any specific language teaching method, it did influence the development of language materials which presented explanations of grammatical rules. As previously discussed, contemporary MEXT guidelines are somewhat abstract or ambiguous, and it seems that a grammar translation method appeared to be the answer to the global trend associated with Chomsky.

2.3.3 Grammar translation

According to Richards et al. (1986), grammar translation focuses on learning the grammar of a language and using this knowledge to translate from one language to another. Richards and Rodgers (1986) and Brown (2007) characterize this approach as:

- Language is presented through sentences which show grammatical aspects of the language. Units of work are based on the introduction of a new grammatical point, a list of new vocabulary and a number of sentences for translation
- Written language is seen as the superior form of the language
- Learning is judged in terms of accuracy in L1

- The goal is for students to be able to read literature in L2
- Students do not talk
- L1 is used as the language of instruction
- There is no focus on the context or situation
- There is no concern with use or real communication

Here we see a method that in many ways utterly contradicts what had gone previously. Many of the tenets of the audiolingual method are completely contradicted by the grammar translation approach. This belief in adopting methods characterized by global trends can be seen again in contemporary MEXT guidelines that call for the communicative approach, and again we can see the adoption of methods that lead to a rejection of what has happened previously.

Preston (2007) refers to “folk linguistics”, which are the views that are commonly held about language, and which appear regularly in everyday conversation and in letters to newspapers. Folk views, however, can be extremely influential in Japan (Watanabe, 2004). Many politicians and public commentators advocate particular models of teaching over others, without feeling the need to develop any specialized expertise in the field. Takayama (2008) explains how MEXT guidelines that were aimed at modifying Japan’s education system to a “child-centered pedagogical approach” (p. 388) caused great controversy, fuelled by PISA rankings published in 2001 and 2004. Claims were made in the media and by politicians that any change from traditional teaching methods would undermine Japan’s competitiveness and turn it into a “nation at risk”. This caused MEXT to issue an unprecedented statement in 2005 that any educational reform proposed had been misguided. These folk linguistics are often based upon politicians’ or commentators’ own school experience (Law, 1995). Tahira (2012) has identified the lack of adequate teacher training in Japan, in terms of fulfilling MEXT guidelines; and part of this training should develop the

study of grammar beyond the grammar translation method, to include alternative approaches. Currently, many teachers may feel their only recourse is “folk linguistics”.

2.3.4 Communicative language teaching

Following global trends, Tahira’s (2012) summary of MEXT guidelines shows that from the 1970s onwards, a communicative approach was desirable. In 1989 the guidelines stated that developing students’ communicative ability in English was the central purpose of English education in Japan. The development of the communicative approach was born from the work of sociolinguists in the 1970s, and the notion of “communicative competence”.

The concept of “communicative competence” was developed by the sociolinguist Dell Hymes (1967, 1972). In contrast to Chomsky, Hymes claimed that language was not just concerned with usage, but also how to use the language appropriately in a variety of situations and circumstances. He believed that social interaction was much more important than mastery of language structure. Hymes’ theory posits that a competent speaker needs to understand not only grammatical rules, but also:

- The social and cultural rules which apply to the context or situation in which language is used
- The relationships between the interactants
- The purpose of the communication
- The topic
- How to speak or write strategically to achieve a purpose

The main aim of teaching in this approach therefore is seen as enabling students to use language appropriately in social contexts. McDonough and Shaw (1993) categorize the approach as having the following main features:

- Spoken and written language is important
- Language is viewed within its social context
- Language learners need to develop knowledge of the social context in order to develop appropriate language use
- Course content is based on student needs
- The teacher acts as a facilitator to arrange language learning opportunities
- Teaching is learner-centred
- Fluency is important as well as accuracy

From these definitions, we again see a tension between what has happened previously and what is now expected of teachers. Many Japanese teachers in the 1980s and 1990s would have undergone study and assessment within a system of grammar translation, or the audiolingual method, when they themselves were students; and thus they are highly resistant to new ideas. Richards and Lockhart (1996) explain that teachers' beliefs about learning are often based on their own experience as language learners. Coupled with the assessment procedures that are still the target of much learning, it is almost inevitable that the communicative approach would face problems in Japan.

2.3.4.1 Problems with communicative language teaching

MEXT guidelines have followed global trends, from an audiolingual approach whereby grammar was ignored and tasks were orally based, to a grammar translation approach where written language was seen as vital. McDonough and Shaw (1993) claim that in the early days of the communicative approach, teachers tended to abandon grammar in favour of using language in real-life settings. McDonough and Shaw (1993) explain that communicative language teaching has led to:

- An overemphasis on oral skills

- An avoidance of grammar teaching
- Inadequately preparing students for tasks
- A sharp decline in accuracy
- A devaluing of the teacher as a source of knowledge in relation to language teaching
- Shy or introverted students being placed at a disadvantage

McDonough and Shaw (1993) describe the design of communicative materials as polarizing function versus grammar, as if they were somehow in opposition. Yamaoka (2010) claims that the increased emphasis on fluency over accuracy in English teaching in Japanese schools has led to a general decline in the level of English proficiency in students, particularly in the level of accuracy of students' written work. These problems, however, do not appear to fully explain the lack of success of the communicative approach in Japan, as many of the above arguments continue to contradict the reality of the situation in Japanese classrooms. The school exam system tends to measure accuracy rather than fluency. Therefore, a decline in accuracy coupled with a corresponding improvement in students' fluency would be reflected in exam results only as a decline in accuracy. Any improvement in fluency would not be detected. If the exam system measures a decrease in accuracy but not an increase in fluency, this does not necessarily mean the standard of English is declining.

In contrast, MEXT guidelines as recent as 2012 call for more emphasis on oral skills. Tahira (2012) claims that the implementation of communicative language teaching has happened at a "sluggish pace" (p. 5). MEXT (2010) in its own studies has found that the majority of English classes are conducted in Japanese, with the majority of student output being also in Japanese.

One reason for this is that the communicative language approach is not a "method", but rather a "methodology". Brown (2007) describes this methodology as having varied

interpretations and manifestations, with the ambiguity regarding definitions leading to varied teacher perceptions of what constitutes communicative activities. One reason why this situation may be exacerbated is that the MEXT guidelines are often vague and difficult to understand. According to Fennely and Luxton (2011), the MEXT guidelines are routinely misunderstood, and the researchers highlight an urgent necessity for clear and unambiguous definitions of teaching activities and procedures that are to be followed. They call for teacher training, arguing that current provisions are inadequate for addressing the aims and objectives of the MEXT guidelines. It is small wonder that teachers who have themselves experienced a grammar translation system of education may feel more comfortable within this framework.

Added to this is the continued presence of an examination system that appears contrary to the MEXT stated guidelines for language use. The primary form of assessment that has the most profound effect on English language teaching in Japan is the high school and university entrance exams. Section 2.3.5 will present the integration of assessment into these different approaches to English language education in Japan, as in many cases assessment dictates entirely the motivations, aims and objectives behind different approaches to language education.

2.3.5 Assessment and entrance examinations

Watanabe (2004) describes the Japanese university entrance examination as “an emotionally charged issue” (p. 126). Indeed, entrance exams are generally seen as the defining measurement of student success. Watanabe also claims that a grammar translation method of teaching is prevalent due to factors such as the university entrance exams. Gorsuch (2000) identifies the entrance exams as an “institution in Japanese education” (p. 7). At the same time, she outlines that apart from the “Centre Exam”, these exams are not created by or under the scope of influence of MEXT, but are instead the creation of public and private universities. The MEXT guidelines make no mention of any kind of entrance exams in the

guidelines for studying foreign languages; nevertheless, these exams have a huge impact on Japanese media (Nishino and Watanabe, 2008) and the public perception of educational goals. According to Gorsuch (2000), teachers therefore feel they need to prepare students for exams. Referencing Cohen and Spillane (1992), who claim that university entrance exams are an essential element of instructional guidance, she feels that entrance exams create the main target of school language instruction. This sentiment is echoed by Mulvey (2010), who claims that entrance exams have long served as an important source of objective evaluation in Japan, “ensuring a level of quality control (however inadequately)” (p. 18). The *Chuuō Kyouiku Shingikai* (Central Education Council, 1999) states that despite changes in the teaching of language instruction by MEXT, a number of high-school educators continue to hold the opinion that unless entrance exams to universities are changed, the curriculum itself cannot be changed.

In my own context, this habitual perception of the importance of summative testing also prevails, albeit with the entrance exam now switching to a focus on receptive tests such as TOEIC, which rely heavily on grammatical knowledge and contain no productive speaking elements. Law (1995) claims that teachers are focusing instruction on students’ linguistic knowledge rather than their linguistic skills. The high value placed on university entrance exams therefore preserves an approach that favours grammar translation. Although MEXT guidelines appear to call for alternative methodologies, the strong influence of assessment procedures on public, teacher and student perceptions means that in reality (in the classroom), very little change has taken place in teaching practices since the guidelines of the 1960s. A key element in the prevalence of grammar translation methods in teaching practice in Japan therefore seems attributable to the assessment system. However, a change in the assessment system is not entirely impossible: the Central Education Council states that changes are being made in the entrance examination system, and have been for some time. Mulvey (2010) also claims that as the number of test applicants in Japan, with its rapidly falling birth rate,

continues to decrease, the entire system of entrance examinations will have to change.

Mulvey explains that the number of applicants applying for university is beginning to equal the break-even point for financial stability, thus rendering the entrance examinations futile: in other words, all applicants are accepted in university courses.

2.3.5.1 Alternative forms of assessment

According to Davison and Leung (2009), teacher-based assessment has become institutionally adopted in a number of education systems across the globe; including Asian contexts such as Hong Kong, China and Singapore, where assessment for learning procedures has become supported by policy. Hill and Sabet (2009) conducted a study in Japan involving speaking assessments that utilized “Dynamic Assessment”: the results suggested significant cumulative improvement in learners’ speaking performance. Ishihara (2009), also in Japan, considered that there was potential for using teacher-based assessment to develop learners’ productive skills in communication. However, the widespread adoption of assessment for learning practice has not enjoyed the same institutional recognition in Japan as in other Asian countries.

In Japan, summative assessment procedures such as university entrance exams, or TOEIC, remain the primary recognized measurement of student achievement (Cohen and Spillane, 1999; Mulvey, 2010; Watanabe, 2004). As previously stated, attempts by MEXT to address this situation in 2004 (Takayama, 2008), by introducing learner-centred methodologies, were abandoned as “misguided”. This does not, however, mean that there is no future for these assessment procedures in Japan beyond individual teachers. Takayama (2008) points out the homogenizing effect of the PISA rankings and the strong regional competitiveness Japan holds with high-ranking PISA nations such as Singapore, Shanghai and Hong Kong. Since these regions have adopted policy-supported assessment for learning procedures, and have performed well in PISA rankings, factors such as homogeneity and

regional competitiveness, which Takayama (2008) describes as being highly influential on educational policy, may compel Japan to follow suit. The Central Education Council (1999) claims that changes to the entrance examination system have been under way for some time. Mulvey (2010) predicts that changes are inevitable, due to rapidly falling admission rates for universities, which make the entrance examinations redundant.

Adoption of assessment for learning practices does not necessarily entail abandoning traditional Japanese assessment procedures, or creating a dichotomy between summative and formative assessment, but rather requires integrating them with new methods. Black (2009) argues that assessment for learning practice can provide a formative use of summative assessment tasks by treating them as an occasion for formative feedback. This can be done via peer or self-assessment activities that require students to think about the purposes of the work to be tested, or mark each other's test responses, in order to focus attention on criteria of quality. Kennedy et al. (2006) argue that the polarization of formative and summative assessment is not useful, and that we should look at summative assessment methods as productive learning opportunities. Davison (2008) prescribes summative assessment as an integral part of assessment for learning in the classroom, providing that results are used formatively to guide future learning and syllabus design. He promotes summative tests at different stages of a syllabus, from a level focused on criteria that help students decide what to do next, conducted by students and peers themselves, to system-wide published scales and standards, and formal tests. Therefore, if implemented correctly, assessment for learning could provide opportunities to achieve an assessment methodology that complements existing practices in Japan, rather than offering an alternative or contrasting view.

2.3.6 Current issues in Japanese language teaching

The historical context of Japanese education shows that MEXT appears to have closely followed global trends and has adjusted its guidelines accordingly, even though these

methods are often contradictory. Assessment, particularly the demands of the university entrance exams, appears to be the largest obstacle to changing the status quo. At each stage, guidelines appear vague and aspirational, with inadequate teacher training practices. This is particularly true of current ideology and the proposed adoption of the communicative approach, whereby a methodology of teaching is often misunderstood and misinterpreted. The notion that the communicative approach rejects accuracy (such as in Yamaoka, 2010), ensures an environment of cultural resistance from teachers whose beliefs will have been influenced by their own learning experiences. Meanwhile, assessment procedures, as the primary focus of measuring the success of Japanese students and the education system as a whole, have remained intact since the 1960s. This has compelled teachers to resist teaching approaches that have developed more recently, and which are often inadequately defined by MEXT.

2.3.6.1 The role of MEXT

At the tertiary level in Japan, MEXT has no official power or role in deciding the content of English language programmes, or of entrance examinations at private universities, which are the sole domain of the universities themselves. The accreditation associations for Japanese universities vary in their goals for language education. The Japanese University Accreditation Association (JUAA), which accredits university status to the university used in this present study, provides no English language teaching or learning guidelines as a criterion for accreditation. MEXT (2008b) itself declares that:

... it should be emphasized that the Fundamental Law of Education stipulates that the independence, autonomy and the merits of education and research by higher education institutions shall be respected. This principle of self-governance has been assured by Japanese Supreme Court decision. (p. 3)

Therefore, no centralized or institutionalized curriculum of language instruction exists at the tertiary level. Language curriculums are at the sole discretion of the individual university.

MEXT (2008b), however, has issued a pamphlet entitled “Higher Education in Japan”, which contains language course guidelines that it hopes are “useful, especially for those in charge of higher education policies” (p. 3). Section 5 of that pamphlet describes a goal of the “internationalization of universities” (p. 17), where MEXT argues that amid ongoing globalization, it is essential that universities develop an educational environment where students can acquire English skills. It further stipulates that it is very important for Japanese universities to conduct lessons in English “for a certain extent” (p. 17), or develop courses entirely in English. MEXT highlights that many universities in Japan already have classes taught in English, and that “there are 50 or more graduate schools where students can graduate by taking only lessons conducted in English” (p.17). Many universities in Japan have compulsory English components, and students who are non-English majors can be prevented from graduating if they do not also attain credits in English classes. This is true of the university in the context of this study, which is discussed in more detail in Chapter 3.

MEXT has gained considerable influence over EFL teaching practices in universities through the awarding of KAKENHI grants (Grants-in-Aid for Scientific Research) to individual teachers or teaching departments in universities. For example, in 2011 alone, 206 million yen was made available to universities under the category of “language teaching”, for research ideas considered to “promote creative and pioneering research in critical fields attuned to advanced research results” (Japan Society for the Promotion of Science, 2012). A primary source for discovering what is required for such an award is pamphlets such as the MEXT (2008b) “Higher Education in Japan”. In the university in which this study is based, there is no central curriculum, despite the existence of a Language and Research department.

In my own experience, this is also true of four other universities in the area, one of which is a dedicated language university. MEXT guidelines are, therefore, an important consideration for tertiary-level teachers.

2.3.6.2 Other considerations in the contemporary context

Yamaoka (2010) claims that there is a dichotomy between grammar translation and communication, which has brought about a situation whereby “students’ basic ability has declined” (p. 62). He defines this basic ability as the means for students to perform correct sentences, not just in writing but also in speaking: “The students have become able to say easy things, but other than that, their communication ability has not reached a higher level” (p. 62). He claims that there are a great number of teachers in Japanese schools who do not think communicative classes help students to pass the entrance exams. This viewpoint appears to summarize succinctly the major issues concerning the development of English education in Japan. There is a belief in a dichotomy between accuracy and fluency, coupled with the belief that entrance examinations require the teaching of grammar at the expense of communication.

McDonough and Shaw (1993), however, label this dichotomy as “false” (p. 21). They observe that although the movement towards communicative approaches began in the 1970s, this shift did not take place everywhere; and in some areas of the world the debate is “current, reflecting the differing and changing perceptions of the needs of the education systems” (p. 22). This appears to be the case in Japan. Despite MEXT guidelines, the communicative approach is very much a current trend, and as a methodology it is open to misinterpretation. McDonough and Shaw claim that there is much to the main principles of the communicative movement that should not be underestimated, though the approach requires varying degrees of change and modification.

As Feez and Joyce (1998) stated, new approaches are often categorized as discrete packages, challenging the legitimacy of previous methods. This does appear to be the case in Japan, where MEXT guidelines have often proved contradictory, depending on global trends. Feez and Joyce describe that what is missing is a “framework within which the sum of present knowledge about language and learning can be organized” (p. 13).

Feez and Joyce (1998) discuss how different types of classroom mirror five different syllabus types, as outlined in Table 2.1:

Table 2.1 Classroom types and their corresponding syllabus type. Adapted from Feez and Joyce (1998).

	Classroom situation	Syllabus type
1	In this classroom, the learners are studying grammar. They learn the rules and then translate the sentences. They start with fewer complex aspects of grammar and build up to more complex ones.	Structural
2	In this classroom, students are learning simple spoken exchanges focusing on functions of language such as apologizing, expressing opinions and making requests.	Functional – notional
3	Here the learners are doing a project on sport. Although they are doing a lot of reading, writing, listening and speaking, they do not really focus specifically on language.	Topic-based
4	In this class, the learners explore a range of elements of language: structures and functions. They also work on topics, tasks and different types of texts. Their work is based on a thorough analysis of the language needs that they have.	Mixed
5	In this classroom, there are no set aims for the learners. Their teacher negotiates with the students the work that they will do each week and the language that will be their focus.	Process

The structural syllabus focuses on lexical items and grammatical structures sequenced according to their perceived complexity, based on the idea that the learner accumulates the building blocks of language one by one in a process that eventually constructs the entire

language. Feez and Joyce (1998) explain that this syllabus type ignores context and meaning, but that sentences are illustrated in contrived sentences, isolated from real-life language. This syllabus type best fits the definitions of the grammar translation method outlined in 2.2.1.3. It also mirrors the “Course of Study for Senior High Schools in Japan” (MEXT, 1999), which calls for the study of sentence patterns, such as in “Section II.a”: “(a) Subject + Verb + Complement, in which the verb is other than be and the complement is a present participle or past participle, or the verb is be and the complement is a clause beginning with what, etc. that, if, or whether”. This is also seen in “Section II.b”, which lists grammar objectives such as “Adverbial use of infinitives” and “Basic use of the subjunctive”.

Feez and Joyce (1998), however, do not claim that this syllabus type is incorrect or incompatible with other syllabus types. Instead they claim that there is a type of syllabus that allows teachers to select the best aspects of all other syllabus models and integrate them into a coherent, cohesive and comprehensive course design. They label this syllabus type a “text-based” syllabus, which in this study I refer to as a genre-based approach. They identify relationships between a text-based syllabus and the structural syllabus by explaining that structural syllabi and related materials are useful to teachers designing a text-based syllabus because they provide a repertoire of vocabulary and grammar activities to draw on when preparing that part of the text-based syllabus which focuses on grammar and vocabulary.

Feez and Joyce (1998) also identify a “task-based syllabus” that perhaps best describes a form of communicative approach whereby the syllabus is recorded in terms of methodology: tasks emphasize communication, with learners learning by interacting communicatively and purposefully while engaged in activities and tasks. These kinds of syllabus goals can also be seen in the same MEXT (1999) guidelines as mentioned above, section III.2.(1): “teaching materials focusing on everyday life, manners and customs, stories . . .”. Feez and Joyce (1998) identify the limitations of this syllabus type, regarding it

as confusing due to the lack of content specification. These limitations appear to mirror the problems that the MEXT guidelines have encountered in Japan: namely, the misinterpretation of methodological approaches, and a cultural resistance to alternative models of education. Feez and Joyce (1998) introduce the text-based syllabus by explaining that when viewed from the perspective of methodology, it is possible to think of a text-based course as being task-based. The next section will consider the theoretical definitions of a genre-based approach. These theoretical definitions will inform the research design in this study, as a syllabus is introduced as part of an action research process, based upon the fundamental principles as defined in the literature.

2.3.7 Summary of section 2.3

Before considering the adoption of a genre-based approach, it is important to consider the historical context in which this approach would be implemented. A number of different approaches have been implemented in Japan that have echoed the wider understanding of English language learning and pedagogical approaches. These different approaches have often been presented as trends to supersede what has come previously, and as a better way of doing things. This presentation of different approaches in discrete packages has one major implication for this study: specifically, that the communicative approach has been seen as an avoidance of grammar teaching; and some research has suggested this has coincided with a general decline in accuracy. This is a dire indicator for teachers concerned with university entrance exams and summative tests such as TOEIC. The challenge for a new approach will be to present itself as a methodology that complements what went before, rather than dismissing it entirely; a methodology that draws on the best of all teaching practices and individual teacher expertise, while considering the historical context of English language education in Japan. It must consider issues of assessment in order to complement current aims and objectives, rather than appearing to contradict them. The next section will outline the theoretical underpinnings of the genre-based approach. These theoretical underpinnings

will inform the design of a teaching intervention in the form of a genre-based syllabus that constitutes the basis of this study's research design.

2.4 The Genre-Based Approach

As noted in the previous section, Feez and Joyce (1998) describe a “text-based syllabus”, while this study describes a “genre-based syllabus”. These terms are used synonymously in this study. The theoretical considerations behind the choice of synonymy in this vocabulary are explained in section 2.4.1. However, the word “genre” was also given preference over the word “text” due to the Japanese word for “genre” being the same as that in English: in fact, “genre” is an English loan word in Japanese, with the same meaning. Describing the notion of genre and a genre-based syllabus, therefore, would hold more resonance with students, based on their current linguistic knowledge in L1.

2.4.1 Text and genre

A brief description of these terms as synonyms will therefore follow. “Text” derives from “context”. The word “context” literally means “with-text”, from the Latin. In everyday use we use it to refer to the wording that comes before or after a particular linguistic item. For example, if I read the sentence “His behaviour was churlish” and I do not know what “churlish” means, I will need more of the text surrounding the word to tell me what it means. I can understand that churlish is referring to his behaviour, but I cannot tell anything more than that. If I read: “John sulked for two hours before lunch and when I asked him to help with the dishes he was rude. His behaviour was churlish”, I now have a clear picture of the sort of behaviour that constitutes being “churlish”. Firth (1950) proposed that the term “context” should be broadened to include the relevant social environment of any act of communication.

Halliday and Hasan (1976) refer specifically to a “text” as a coherent piece of written, spoken, non-verbal, visual or auditory language produced in an interaction with the intention

of communicating some kind of meaning. A text can be any length. For example, a stop sign at a junction is a text, as is a novel. Thus, we have Feez and Joyce's (1998) "text-based" approach, i.e. syllabus planning through the use of whole texts which are embedded in the social contexts within which they are used. The word "genre" simply refers to the text type, or any purposeful activity that is characteristic of a cultural group or community. It has a characteristic staged generic structure. For example, Butt et al. (2000) identify seven text types, or genres, that children engage with in primary school education: narrative, recount, information report, discussion, explanation, exposition, and procedure. Martin and Rothery (1993) define genre as "the different types of text used in our culture to get things done" (p. 147).

Egins (2004, pp. 54–55) illustrates an easily understandable definition of genre by quoting an extract from a horoscope: "You are on the threshold of a magnificent chapter in your life, with substantial opportunities emerging after the new moon on the 5th ...". She explains that most readers can quickly identify this extract as a horoscope, as the text is doing something with language that we recognize from previous experiences. When we identify a text as a certain genre we are stating the "purpose" of the text. By being able to identify the genre of a text, we understand its "generic identity", and the purpose of this text in the culture within which it is written.

A further reason why the word "genre" was chosen in this study, rather than "text", was a linguistic decision. In Japanese the word for "genre" is "*janru*", an approximation of the English word "genre", as it is a loanword from English with the same meaning. This allows the concept of genre to be more accessible to students during the explanation of a genre-based syllabus in English.

2.4.2 Context of situation and context of culture

Malinowski (1923), in his early anthropological work, developed the concepts of “context of situation” and “context of culture” in language use. Malinowski (1923) used the term “context of situation” (p. 300) to describe the immediate environment of communication between Trobriand Islanders in the South Pacific region, surmising that in order to explain to his European readership the islanders’ use of language, he also needed to describe their physical and social environment. Without knowledge of the context, it would be challenging to interpret or encode the text accurately.

Malinowski’s ideas, however, were not fully developed into a theory of linguistics. Firth built on Malinowski’s notions of context (see Halliday and Hasan, 1985, p. 8) in order to apply a conceptual model for different uses of language. He therefore adapted Malinowski’s term “context of situation” to refer to the immediate instantiation of any act of communication, which included “participants”, their “actions”, the “surrounding objects and events”, and the “effects” of the verbal action on the situation. The importance of this contextual model is shown in its mirroring in 1967 by Dell Hymes (in Halliday and Hasan, 1985, p. 9), who described the context of situation as “the form and content of the message; the setting; the participants; the intent and effect of the communication; the key; the medium; the genre and the norms of interaction”. Malinowski’s ideas and the concepts of “context of culture” and “context of situation” have become very important foundations for social and functional theories of linguistics that have emerged over time.

This development is furthered in the works of Halliday (1985, p. 5), who provides a fuller description of context and the further notion of “text”:

The terms, CONTEXT, and TEXT, put together like this serve as a reminder that these are aspects of the same process. There is text and there is other text that accompanies it: text that is “with”, namely the con-text.

He defines “text” by stating: “it is language that is functional. By functional we simply mean language that is doing some job in some context, as opposed to isolated words or sentences that I might put on the blackboard” (p. 5). He therefore explains that any instance of living language that plays a part in a context of situation is a text.

The non-linguistic factors in the social environment are also important. The term “context of situation” covers the social and material factors which are relevant to understanding the language that people use. This goes far beyond the superficial “setting” of communication such as “at the hospital” or “at the restaurant”.

Halliday (1985) identifies three features within the context of a situation that provide a conceptual framework to interpret the social context of a text and the environment in which the meanings are exchanged: namely, Field, Tenor and Mode:

Field – the social activity going on at the time (e.g. a “doctor-patient consultation” or “making a restaurant reservation”)

Tenor – the social roles people take up when they communicate (e.g. “doctor/patient” or “restaurant manager/guest”)

Mode – the medium (or channel) of communication adopted (e.g. speech or writing)

Halliday’s “functional” model of language helps us to see the relationship between context and text, and to focus on the linguistic consequences of this relationship. There still remains a gap, however, between this theoretical basis and planning a syllabus to help students understand this too, and thus develop their speaking abilities.

2.4.2.1 Context of culture

At another level, Halliday (1985) introduces the “context of culture”, i.e. the broad sociocultural environment operating behind social situations. This includes factors such as

ideology and ongoing social conventions and institutions, and the ways in which these affect the language used by people in any text.

This context of culture is an essential consideration for this study. The Japanese education system, and factors such as MEXT guidelines, will have influenced students and teachers in a number of ways before they enter tertiary education. The context of culture will influence:

- The ways in which the teacher interacts with students
- The kinds of textbooks that are used
- The expectations teachers have of students
- The kinds of English that students can learn
- The access students have to these kinds of English

This context of culture is therefore an essential component in the planning of any syllabus, such as the one involved in this study. These cultural influences will have to be considered, particularly the issue of grammar translation as dictated by assessment and teacher beliefs. Joyce and Burns (1999, p. 3) argue that the pedagogical approach associated with genre theory provides students with guidance “according to students’ needs and the social contexts which students need to be able to access . . . in which learners are taught aspects of grammar as they relate to spoken and written texts”. EFL students in the Japanese setting, who are unable to experience English-speaking social contexts directly, are therefore likely to benefit from a whole-text approach, through an introduction to cultural and situational texts as they are understood by “cultural insiders”, together with the grammatical and structural demands expected of certain text types (Kashima and Kahima, 1998). The notion of grammar is an important one to consider when we think about the kinds of English

that students can learn, and teachers' expectations of students in an educational context which is based upon a focus on grammar translation.

2.4.3 The role of grammar in a genre-based approach

As previously stated, one of the problems with the communicative approach is the perceived lack of grammar instruction associated with that approach, or even a view that the learning of fluency impedes the learning of accuracy. This role of grammar in Japanese language education would need to be addressed in a study such as this, which aims to introduce new ways of looking at language learning. It seems a necessity, therefore, that grammar is included in any new approach. However, a new way to look at grammar might also be necessary.

Martin and Rothery (1993) identify three models of grammar, focusing particularly on their impact on language teaching in Australian schools. Nevertheless, despite this particular focus, their analysis provides a good overview of how new approaches can be adapted into educational culture.

As previously mentioned, "folk linguistics" (Preston, 2007) often has a role to play in deciding which models of grammar are used in different educational contexts. This view is often based upon politicians' or public commentators' own school experience, which ignores developments in grammar teaching since they left school. Folk linguistics may have been the reason why the MEXT guidelines in the 1960s were widely interpreted in Japan as the grammar translation method.

Martin and Rothery's (1993) three methods of grammar study are those of "traditional", "formal" and "functional" grammar. Traditional grammar is based upon the study of Latin, and was applied to formal written English. It was adapted for use in schooling in the nineteenth century, and knowledge of traditional grammar was considered to be

essential for young people. The rules of traditional grammar were applied in a very prescriptive way. As Martin and Rothery (1993) describe it, traditional grammar focused on what people are supposed to say and do when they think about it, rather than what they actually say when they do not.

Formal grammar was developed to take account of universal regularities across English and other languages. Formal grammars are tools that have developed out of a different set of interests and tasks: ones that focus on the universal and context-free. It is here we see the influence of Chomsky and transformational-generative grammar, which Chomsky himself proclaimed had no practical relevance whatsoever to language teaching and learning in schools (1967). This kind of grammar focuses on the limits of human grammars, and is not intended to be used for practical purposes in education. However, it appears that the emergence of Chomsky's views coincided with the MEXT guidelines calling for the focus of language education to change from the audiolingual to grammars and structures. This was not without good reason, as Martin and Rothery (1993, p. 141) explain: "the formal grammarians are the most powerful group of linguists in the world today; even linguists who disagree with them often do so in a cringing way that defers to their work". For this reason, they explain that such grammarians' research and dominant institutional position must be taken very seriously. This dominant institutional position, coupled with folk linguistics, may explain a great deal of the thought behind Japan's approaches to language teaching.

Functional grammar as developed by Halliday (from Martin and Rothery, 1993) is intended as an "applied grammar" (p. 144), and has been utilized in numerous EFL contexts (Moore, 2009). It covers both spoken and written grammars. Unlike traditional and formal grammars, functional grammar is oriented to language as a resource for making and sharing meanings with other people, rather than being a set of rules for producing sentences. This version of grammar is referred to as Systemic Functional Grammar, or Systemic Functional

Linguistics. Table 2.2 summarizes the major differences between traditional and functional grammar, as outlined by Martin and Rothery (1993):

Table 2.2 The major differences between traditional and functional grammar, developed from Martin and Rothery (1993).

TRADITIONAL	FUNCTIONAL
Describes linguistic “good manners” (p. 140) rather than language.	Describes language in use
Focuses more on written language than on spoken language.	Focuses on both written and spoken language.
Uses “class” labels (parts of speech, e.g. noun, adjective, verb).	Uses “functional” labels (e.g. process, participant, etc.).
Identifies parts of speech and a set of rules describing relationships between words, groups of words, and clauses.	Describes relationships between words and groups of words across sentence boundaries throughout text.
Looks at sentence level only.	Shows how whole texts function.
Is interested in the grammatical structures in texts rather than the meaning and content of texts.	Functional analysis views language as a communicative resource and is primarily interested in how linguistic structures express meaning.
Looks at lexical cohesion at sentence level.	Looks at cohesion in a whole text.
Teaches about grammatical “correctness”, no room for exploration.	Allows exploration of language and how the language system operates on different levels.
Enables sentences to be parsed.	Provides students with a metalanguage that can be used to analyse texts and reflect on them.
Accessible to students who have been taught these terms.	Involves learning new concepts, and understanding and using a new metalanguage.
Static – a set system of rules that can be applied to all texts.	Dynamic – explores the variables in texts and shows how they work together in making a text successful / unsuccessful in achieving its aims.

The contrasts in Table 2.2 identify some of the major issues in language teaching in Japan that were discussed earlier. The traditional methods are often exemplified in MEXT guidelines that focus on written language. These sets of rules and “class” labels appear to be far more conducive to writing than to speaking. This may explain why students associated with the current study often transcribe extended utterances before they feel confident enough to speak. Table 2.2 also raises some issues for concern, however: the idea of understanding

and using new metalanguage may work to confuse students, as the focus of new vocabulary or instruction switches to metalanguage rather than the language contained in the goals of the syllabus. Nonetheless, much like Feez and Joyce (1998), who call for an eclectic approach to teaching, Lock (1996) points out that there is no best way or “method” (p. 270) for teaching grammar. What should be done is to consider a number of methodological options. One such option would be to continue to use labels that students are familiar with, in conjunction with new functional labels where appropriate.

Lock, however, points out that many of the rules of traditional grammar, as identified by Martin and Rothery, are “at best half-truths and did little to help learners see how the structures could be used meaningfully in context” (p. 265). Lock refers here to grammar practice associated with the manipulation of structures at sentence level, with little or no context provided. These sentence-level, non-contextualized language examples are identified in the 1999 MEXT language teaching guidelines. Feez and Joyce (1998) describe this as the “structural” approach.

As discussed previously, the 1970s saw a call for communicative competence and the introduction of communicative language teaching. However, as late as 1999, we still see in Japan elements of the structural syllabus. Whereas educational establishments elsewhere might have seen this as a case of grammar or no grammar, Japan appears to have adopted the stance of traditional grammar being better than no grammar at all. Canale and Swain (1980) argued the case for grammar teaching by including “grammatical competence” in their list of four proficiency areas for language learners. Grammar should therefore be an integral part of any new syllabus design. This does not just mean reverting to old ways of teaching, as folk linguistics may compel us to do; rather, as Lock (p. 270) points out, we should examine the advantages and disadvantages of various approaches to teaching grammar using the knowledge gained in recent years. A closer look at how genre-based syllabus design interacts

with the teaching of grammar will be discussed later. Before that, it is important to discuss the literature in relation to general pedagogical approaches to the genre-based model, and how this relates to the teaching of speaking.

2.5 From Theory to Practice

Butt et al. (2000, p. 141) describe some of the pedagogical applications of the linguistic theory behind Halliday's model. They discuss the "dynamic" relationship between language and context: "the relation between context and meanings is dynamic and reversible"; noting that if students have knowledge of the field, tenor and mode of the context of situation, they can foresee with greater certainty the grammatical structures involved. By selecting "contexts of situation" that are appropriate to students' learning needs, teachers can provide themselves with tools for presenting and describing texts that students can later recreate for themselves:

By investigating the grammatical patterns in an objective way, we can see how users of language create meanings and achieve intended effects . . . the patterns discerned in our analysis make the definition of style almost as simple as the description of a context of situation because both are aspects of the functional diversity of language. (p. 141)

Butt et al. show how grammar reacts to contextual and generic demands by exploring and deconstructing major text types that primary-school children are engaged in academically. They identify seven such text types: narrative, recount, information report, discussion, explanation, exposition, and procedure. Each of these has a predictable "generic structure" and patterns of language use associated with that structure. As an example, a "recount" genre is illustrated in Figure 2.1:

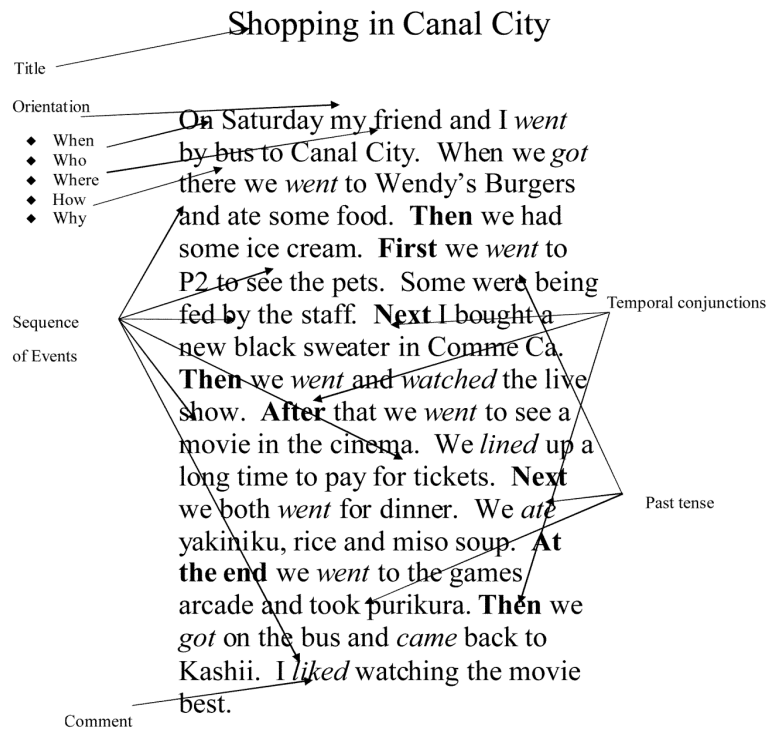


Figure 2.1 Generic structure of recount genre, adapted from Butt et al. (2000)

The example shows that each genre that develops in a particular culture has a particular purpose; particular stages with a beginning, middle and end, and particular linguistic features.

A “Recount Genre” (p. 143) such as this can be characterized by:

- Purpose: to reconstruct past experiences by retelling events and incidents in the order in which they occurred
- Structural features: Orientation (who, where, when what), events in chronological order
- Grammatical features: Past tense, human and non-human participants, temporal conjunctions or clauses, material processes

Butt et al. introduce further genres, shown in Table 2.3:

Table 2.3 Characteristics features of three genre types. Adapted from Butt et al. (2000)

Genre		
Narrative	Information Report	Discussion
Purpose:		
To tell a story as a means of making sense of the events and happenings in the world. It can be both entertaining and informative.	To present factual information about a class of things, usually by first classifying them and then describing their characteristics.	To present information about more than one side of an issue. It may end with a recommendation based on the evidence presented.
Structural features:		
<ul style="list-style-type: none"> • Orientation • Complication • Sequence of events • Resolution • Comment 	<ul style="list-style-type: none"> • Opening statement • Sequence of related arguments • Concluding statement 	<ul style="list-style-type: none"> • Statement • Arguments for and against • Recommendation
Grammatical features:		
<ul style="list-style-type: none"> • Past tense • Process types • Temporal sequence of events 	<ul style="list-style-type: none"> • Present tense • Relational processes • Tight thematic progression • Non-human participants 	<ul style="list-style-type: none"> • Human and non-human participants • Simple present tense • Tight thematic structure • Modality in opinion • Material, mental and relational processes

In Table 2.3, we see reference to different kinds of processes. Butt et al. (2000) identify seven different kinds of processes, with the main four being Material, Mental, Verbal and Existential (p. 51). Table 2.4 provides more information on these process types:

Table 2.4 Process types identified by Butt et al. (2000)

Type of process	What the process tells us:	Examples of processes:
Material	<ul style="list-style-type: none"> • What people do 	<ul style="list-style-type: none"> • Run • Eat • Sit • Swim
Mental	<ul style="list-style-type: none"> • About mental processes 	<ul style="list-style-type: none"> • Think • Remember • Assess
Verbal	<ul style="list-style-type: none"> • How people say things 	<ul style="list-style-type: none"> • Suggest • Exclaim • Whisper
Existential	<ul style="list-style-type: none"> • How things are • What things have 	<ul style="list-style-type: none"> • be / are • have / has

In relation to the elemental genres, Butt et al. note that “The purpose of a text influences grammatical as well as structural choices. Of course, this does not mean that we cannot be creative with language and break away from accepted patterns. What it does mean is that we can learn how to produce an acceptable text of any type simply by following the recommended formula” (2000, p. 20). Thus, these “elemental” genres provide a useful starting point for planning an EFL syllabus and the assessment criteria associated with it. Butt et al.’s point about creativity is an important one, as during syllabus design, generating

criteria that are too prescriptive and rigid may contradict the notion of assisting speaking ability.

The genres presented by Butt et al. relate to written texts, and would need to be adapted considerably in order to be used in a classroom designed to improve students' spoken ability. The next section examines the literature in relation to spoken discourse.

2.6 Register Theory

Drawing on the works of Halliday (1978, 1985), Martin (1984, 1992), Wignell et al. (1989) and Poynton (1985), Eggins (2004) develops the functional model in the context of speaking as "Register Theory", which clearly describes the relationship between spoken language and situation. A discussion of this theory of register will help to introduce some of the key concepts and terms behind the genre-based approach adopted in this study. Eggins (2004) examines more closely Halliday's conceptual framework of a text by asking the questions:

1. What is meant by context of situation, and what are the register variables?
2. How is register realized in language?

She does this by describing exactly what field, tenor and mode refer to, in a clear and accessible way that allows teachers to consider how to effectively plan and implement a genre-based syllabus.

2.6.1 Field

Eggins (2004) claims that the most useful way to analyse a text's field is to look at its pattern of "transitivity" (p. 213). Transitivity is the system for constructing experiential meaning in the clause. By examining transitivity patterns in a text, we can describe its field: the topic and the kind of treatment it is being given. Essentially this tells us "what is being talked about" (p. 249). This includes looking at three components of a clause:

- The participants: The people, places and things that can be related experientially to a process. Using functional grammar, this would be realized through the nominal group of the clause.
- The process type: Process tell us what is going on in terms of the physical world; the world of consciousness and of being. This is realized through the verbal part of the clause.
- The circumstances: The context in which processes take place and specific location. This is realized through adverbs or prepositions.

Egins describes that participants and processes are central to our representation of experience, while circumstances are less central “attendant” processes.

2.6.2 Tenor

The tenor of discourse describes the social roles we assign ourselves and others when we are speaking. Our role in an interaction will influence the language we use. Egins (2004) compares “informal” and “formal” situation types. Informal interactions are those between close friends who see each other often and freely express their feelings. Meanwhile, formal interactions are between strangers from different social levels who do not meet often and are not free to express feelings openly. Egins divides the tenor of discourse into three different continua:

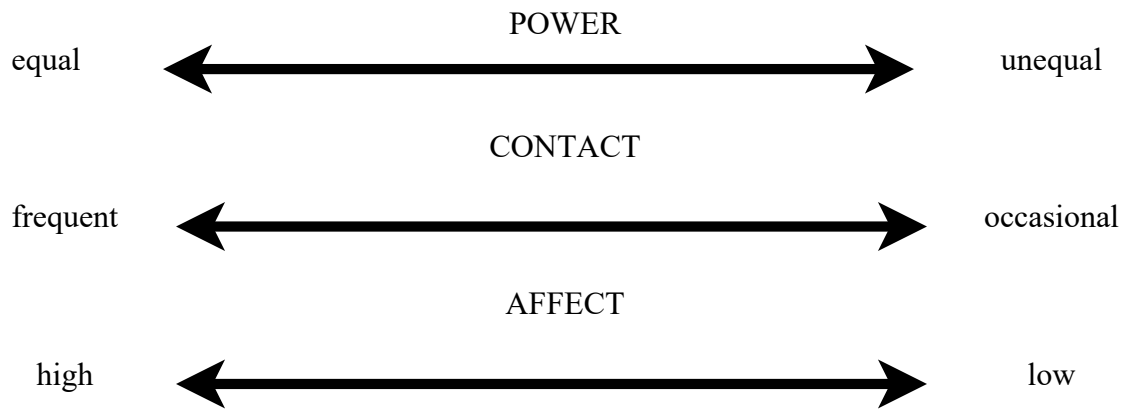


Figure 2.2 The continua of power, contact and effect (Eggins, 2004, p. 100)

Power refers to the extent to which the relationship between the interactants is equal (such as between two friends) or unequal (an employer or employee). Contact means the frequency of interactions, with friends, family members and neighbours at one end, and strangers at the other. Affective involvement refers to the extent to which emotions or attitudes are expressed freely between interlocutors. This may be high between friends, spouses and children, and low between passengers on a train.

Tenor also influences language differences in other ways. Eggins (2004) describes the act of communicating as involving a speaker or writer and a listener or reader. The speaker or writer selects a role for themselves and in turn allocates a role to the listener. Halliday and Hasan (1985) state that all communication can be categorized into one of the following speech acts:

- Offer (to give goods or services)
- Statement (to give information)
- Command (to demand goods or services)
- Question (to demand information)

In each of these instances, the listener has some choice in how they respond: an offer can be accepted or refused; a statement can be acknowledged or contradicted; a command can be undertaken or refused; and a question can be answered or disclaimed. Who takes the role of speaker and who is the listener is an indicator of which of the interlocutors is in the more powerful position.

Egins explains that in looking at how interpersonal meanings are expressed, the main grammatical features we need to examine are the subject and the finite, which combine to create the mood of the clause. The subject is realized by the nominal group, and the finite element is part of the verbal group. The remainder of the clause is called the “residue” (p. 150). The finite element has the function of locating an exchange with reference to the speaker making a statement/command that can be argued about. This is enacted through:

- Primary tense: past, present or future at the time of talking.

Through the primary tense we can argue whether an event did/will/should occur.

e.g. does, did, will

- Modality: this indicates the speaker’s judgement of the probabilities or the obligations involved in what is being communicated.

e.g. can, will, must

- Polarity: This indicates whether the clause is positive or negative.

e.g. was, wasn’t

Finally, another indicator of power is identified by Butt et al. (2000, p. 115) and Halliday (1994, pp. 82–83) as “Mood Adjuncts”. Examples of these are summarized in Table 2.5:

Table 2.5 Mood adjuncts in functional grammar

Mood adjuncts	Lexical examples
Polarity	not, yes, no
Probability	probably, possibly, certainly, perhaps
Usuality	usually, sometimes, always, never
Readiness	willingly, gladly, easily, certainly
Obligation	definitely, absolutely, by all means
Time	yet, still, already, once, soon, just
Typically	occasionally, generally, regularly, mainly
Obviousness	of course, surely, obviously, clearly
Intensity	just, simply, merely, only
Degree	quite, nearly, almost, totally, completely

In terms of real-life communication, the language choices we make on an interpersonal level are extremely important to how we get along in the world. For example, in the Japanese classroom, the teacher is supposed to talk the most, and students are expected to listen carefully. Outside the classroom, however, different situation types involve different expectations about speech roles, such as politeness conventions, turn-taking, attitudinal expressions, and so on. The challenge in designing a genre-based syllabus is how to open up the demands of different tenors and to provide learning contexts which enable students to meet these in ways satisfactory to them.

2.6.3 Mode

By discussing mode, we can examine the difference between spoken and written language more fully.

Eggins identifies two central concepts (p. 91):

- Interpersonal distance
- Experiential distance

Interpersonal distance relates to the spatial distance between interactants. Eggins places situations along a continuum based on the possibilities of immediate feedback between interactants.

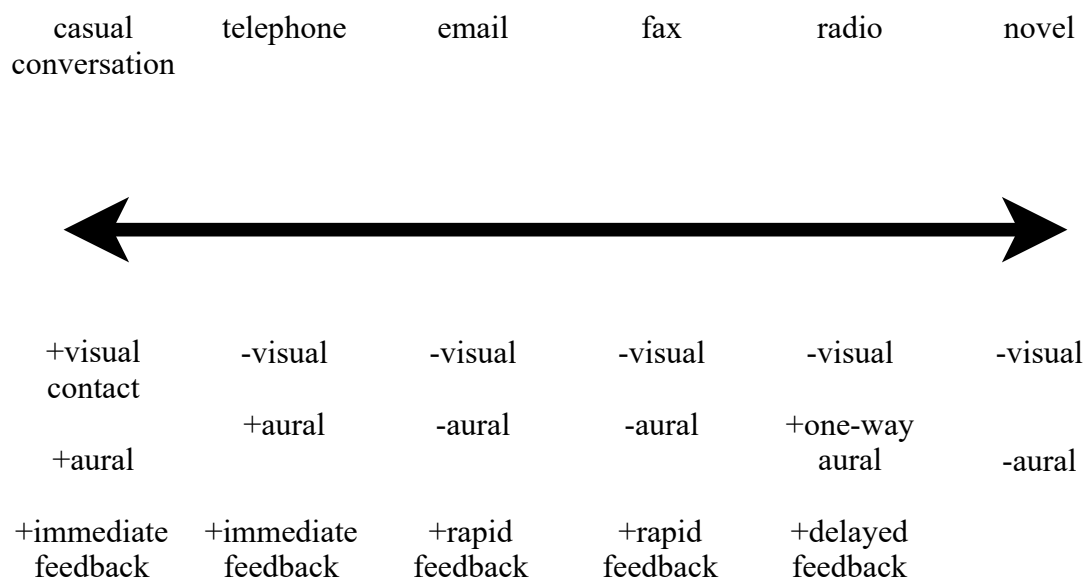


Figure 2.3 Representation of interpersonal distance (Eggins, 2004, p. 91)

Casual conversations typically allow both visual and aural contact between interactants, and hence provide immediate feedback. Novels, on the other hand, do not allow visual or aural contact between interactants. The reader can only be imagined by the author, and not considered in a real way, as in casual conversations. There is no opportunity to give feedback to the author. The continuum could also be updated with video-conferencing, which would provide visual and aural contact, plus immediate feedback; but it may still constitute a spatial distance equal to the telephone.

Experiential distance relates to the distance between language and the social processes occurring; but this tends to be measured in time rather than space. Eggins represents this continuum in the following way (p. 91):

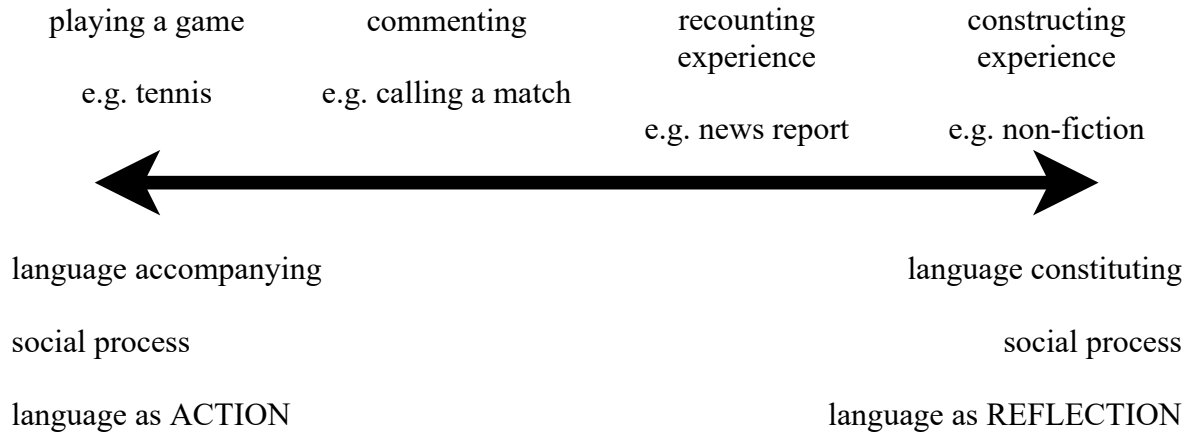


Figure 2.4 Representation of experiential distance (Eggins, 2004, p. 91)

In situations such as playing a game of tennis, language accompanies what is going on as it happens. In situations such as a non-fiction book about tennis, language constitutes “what is going on”. These two ends of the continuum represent language as action and language as reflection. When designing a syllabus, it is therefore essential to consider how the mode of discourse influences language choices, in order to support learners as they work along this continuum and select appropriate language choices for spoken situations.

This section has discussed the concept of register and Eggin’s notion of the continuum, to bring out differences of register in different texts. This enables us to identify which language features are sensitive to which aspects of the context of situation. Changes in field will influence experiential meanings in a text; changes in tenor will influence interpersonal meanings in a text; and changes in mode will influence textual meanings in a text.

2.7 Cohesion, Theme and Rheme

As stated by Eggins (2004), mode refers to how language is being used, whether the channel of communication is spoken or written, and whether the mode is that of action or reflection.

Collerson (1994) and Gibbons (1991) help to develop these ideas further by examining textual meanings. These express the relationship between language and its environment, including the non-verbal, situational environment (context) and what has been said before (co-text). Their discussions focus on patterns of “Cohesion”, “Theme” and “Rheme”.

Cohesion (Collerson, 1994, pp. 126–137; Gibbons, 2001, pp. 83–86) refers to the resources within language that provide continuity within a text, and cohesive devices which help bind parts of a text together. These ideas are summarized in Table 2.6:

Table 2.6 The main types of cohesive device, from Collerson (1994) and Gibbons (2001)

Type of Cohesive Device	Short Explanation	Example (in bold)
Reference	A way of referring to something that has already been mentioned. This is usually done with shorter words such as pronouns or demonstratives.	<i>Last week, there was a visitor. She left a note on the table.</i>
Ellipsis and substitution	“Ellipsis” is where something is missed out because we already understand the meaning. “Substitution” is where one word or phrase has been replaced by another word or phrase.	<i>Person A: What’s your name? Person B: Simon. “My name is” has been missed out. (ellipsis) <i>He said he would get me a present and he did just that.</i> (substitution)</i>
Lexical Cohesion	a) repetition b) synonymy c) antonymy d) hyponymy e) meronymy f) collocation	<i>b) poor and destitute c) left and right d) apple, pear, banana are hyponyms of fruit e) steering wheel, clutch, tyre, gearstick are meronyms of car</i>
Conjunction	This refers to the way that clauses are linked together in one sentence.	<i>and, then, so, but, because</i>

In terms of this study, cohesive devices will therefore help to create a whole, unified text. Each genre will have a typical range of grammatical structures and cohesive devices; these cohesive devices thus assist in forming a genre. If students are to create their own spoken genres, then they must be made aware of these very important cohesive devices relevant to the genre that is being studied.

The second major aspect of the mode of discourse involves “theme and rheme”: Cummings (2005) explains that these are necessary for the speaker or writer to keep listeners or readers well informed about where they are and where they are going in their talk. In writing paragraphs, this can be achieved through the use of the topic sentence which summarizes the paragraph. The co-text expands on this topic sentence. As we saw in Butt et al.’s (2000) generic description of a recount text, the same is achieved in speaking through the “orientation”, where we are told the “when, who, where, how, why” of a situation. Cummings (2005) describes that the same situation also applies at the sentence level itself. Speakers and writers use the first position in the clause to signal to their audience what the message is about. Cummings (2005) says that this first position in the clause contains textual meanings because it signposts the development of the text. The first position is known as “theme”, and the rest of the message is known as “rheme”. Theme can be defined as: “what or who it is going to be about”.

At this point, it is important to discuss more specifically the differences between spoken and written languages. This is particularly relevant in the context of this study, where the students often feel more comfortable writing than speaking, and often produce spoken text by first transcribing a speech to be memorized.

2.8 Spoken and Written Language

The importance of developing innovative ways of teaching spoken text has been proposed by Slade and Widin (2004). They echo a call by Crystal and Davy (1975), who claimed that ESL

pedagogy cannot be improved without an understanding of the realities of English conversation. Eggins (2004) provides a useful starting point for the discussion of the linguistic differences that typify speech and writing. Table 2.7 highlights these differences in relation to the variable of mode, which relates to the channel selected for communication:

Table 2.7 Characteristic features of spoken and written language (Eggins, 2004, p. 93)

Spoken and written language: the linguistic implications of MODE	
Spoken Language	Written Language
turn-taking organization	monologic organization
context dependent	context independent
dynamic structure	synoptic structure
-interactive staging	-rhetorical staging
-open-ended	-closed, finite
spontaneity phenomena (false starts, hesitations, interruptions, overlap, incomplete clauses)	“final draft” (polished) indications of earlier drafts removed
everyday lexis	“prestige” lexis
non-standard grammar	standard grammar
grammatical complexity	grammatical simplicity
lexically sparse	lexically dense

From Table 2.7 we see that written language is more compressed and frequently shows signs of editing. Eggins summarizes the difference between spoken and written language by explaining that spoken language is concerned with human actors carrying out action processes, with sequences of clauses dynamically linked together. Written language, on the other hand, is concerned with abstract ideas and reasons, linked by relational processes in condensed sentences.

Eggins (2004) also describes the process of “nominalization” (p. 94). This is the means by which we increase the information-bearing function of language by turning “things that are not normally nouns into nouns, with consequences for other parts of the sentence”. This increases the lexical density of a text. The main parts of clauses that can be nominalized are verbs and conjunctions. The following sentence is adapted from an example provided by Eggins (p. 94):

i) “I was sick for two weeks and I couldn’t hand my essay in. I’ve got a doctor’s letter.”

ii) “A fortnight’s illness prevented the submission of my essay. I enclose a medical certificate.”

The nominalized parts of the clauses are:

- Verbs: “hand ... in” becomes “submission”, “was sick” becomes “illness”
- Conjunctions: because becomes the reason

This highlights a distinctive difference between spoken and written texts. Spoken texts are organized in terms of ourselves, whereas written texts are organized in terms of ideas, reasons and causes. Finally, spoken text is often full of false starts, hesitations and interruptions, whereas written language can be drafted and edited, with students using a

dictionary to identify vocabulary. A current problem in the context of this study is that when students are allowed time to prepare for longer utterances, their speech does not contain these elements and sounds too polished and unnatural.

As mentioned in section 1.3, the students selected to participate in this study have a great deal of difficulty differentiating between spoken and written texts. Time afforded to preparation for a speaking assignment usually entails the written preparation of a text to be memorized by students. Any effort by a teacher or peer to initiate turn-taking is thus rendered extremely difficult, and the synoptic structure of the talk can sound unnatural and abnormally polished. This is reminiscent of Martin and Rothery (1986) (see section 2.3.3), who commented that in a traditional grammar system, language is often focused on what people are supposed to say and do when they think about it, rather than what they actually say when they do not. If we examine again the way in which the education system in Japan compels the learning of a grammar translation approach, and compare that with Eggins' continuum in Figure 2.4., it is not surprising that Japanese students are more comfortable using language as reflection, rather than language as action. The main source of language input is very often derived from this end of the continuum, so it often seems apparent that students are more comfortable with writing than speaking: in other words, they are often more accurate than fluent. The different contexts of EFL and ESL settings may also influence this, as language in action in English is something that students may only experience during class time.

As highlighted in section 2.4.2, Table 2.8 summarizes Eggins' (2004) discussion of Halliday and Hasan's (1985) "grammar of interaction", pointing out that whenever we use language to interact, the interlocutors are establishing a relationship, through a turn-taking sequence whereby they take on different speech roles. These speech roles are outlined as "giving" and "demanding". Concurrently, a speaker also chooses the type of "commodity"

they are exchanging. This commodity can consist of an exchange of information or that of goods and services.

Table 2.8 Speech roles and commodities in interaction (Eggins, 2004, p. 146)

COMMODITY EXCHANGED		
SPEECH ROLE	Information	Goods and Services
Giving	statement	offer
Demanding	question	command

Eggins (2004) proposes that this functional categorization relates to the types of clauses and grammatical patterns produced when we speak: whereas writing is generally monological in organization, speaking usually requires a form of turn-taking and a choice of roles. These are further dimensions of spoken communication that must be addressed in the classroom. Carter and McCarthy (2006) expand on the differences between spoken and written language. They focus specifically on differences not only in structure, but also in the grammars of writing and speaking. They point out, however, that there are clearly overlaps between grammar that is spoken and that which is written: in some situations, speech can resemble writing; as, for example, in formal speeches. They suggest that, when designing a speaking syllabus, it is preferable to balance it with a corresponding written syllabus, so that written differences are thus revealed to learners in combination, and the differences can be made clear. These considerations are important in the development of the intervention involved in this study, as opportunities for writing may also prove necessary, thus requiring a closer integration of speaking and writing instruction during syllabus design.

2.9 Designing a Genre-Based Syllabus for Speaking

In previous sections, the various language features that combine to make language function more effectively in a range of situations were discussed. It was pointed out how choices in field, tenor and mode can alter the meaning of texts, and how spoken and written language have different purposes and features. Different genres were also suggested, which students might need to know about and use. It is now important to discuss the literature associated with how this theory can be put into practice in the language classroom for speaking purposes. This section will outline syllabus design guidelines from the literature, in preparation for the speaking syllabus to be designed for this study.

The discussion so far implies that designing a language syllabus that meets students' needs means designing one which combines different features of language, such as syntax, lexis, functions and textual structures and purposes, as well as units of meaning that contain cultural, topical and social information. Table 2.1 outlines the five different types of syllabi identified by Feez and Joyce (1998) in language classrooms. They explain that teachers should be eclectic and select the best of these syllabus types during design. In a genre-based syllabus, teachers can organize their syllabus by focusing on aspects of language such as learning domains; language content; register; the language event; text progression; the macro-skills of listening, speaking, reading and writing; and appropriate starting points for each unit of work. This allows students to learn language; learn through language; and learn about language.

Joyce and Burns (1999) suggest five steps in designing a genre-based syllabus: these are summarized in Table 2.9:

Table 2.9 Five steps in designing a genre-based syllabus, developed from Joyce and Burns (1999).

Step	Points to remember
1. Decide on a starting point	<p>You can start with topics, contexts or texts:</p> <ul style="list-style-type: none"> • Topics: Look at topics which are relevant to the learners. Identify texts within each topic. • Contexts: Look at the contexts in which learners need to use language. Get students to identify texts within these contexts. • Texts: Start with texts that students need to learn.
2. Develop goals or aims	<ul style="list-style-type: none"> • Goals and aims should be specified to students. • Aims should fit with learners' needs and goals.
3. Sequence content	<ul style="list-style-type: none"> • Students should be able to understand why content has been sequenced in particular ways. • Courses can be sequenced in various ways: interests, needs, familiarity, complexity. • Sequencing should reflect how texts are used in their social contexts.
4. Analyse the features of spoken and written texts	<ul style="list-style-type: none"> • Analyse the grammatical elements of texts. • Look at the features in terms of discourse features, vocabulary, genre, socio-cultural knowledge.
5. Develop units of work	<ul style="list-style-type: none"> • Set short-term objectives. • Objectives can focus on features of texts and whole texts.

In this study, the above guidelines will be followed in order to develop a genre-based syllabus appropriate to the research questions. This will be discussed further in Chapter 3.

Slade and Widin (2004) claim that spoken interactions such as recounts are types of “storytelling texts” and have “identifiable generic structures(s)” (p. 9). Building on this concept, Slade and Widin outline a range of different spoken genres, including narrative, anecdote, recount, exemplum, observation, opinion, gossip, and joke-telling; all of which have identifiable generic structures.

Elaborating further the notion of spoken genres, Eggins and Slade (1997) identify highly interactive “chat” segments of talk, which are not amenable to generic analysis, and more monologically structured “chunk” segments of talk, which have distinctive beginning,

middle and end structures. This study refers to the “chat” segments as the micro-aspects of conversation, as opposed to more monologically structured “chunk” segments. As they unfold, chunk segments of talk are frequently supported by listeners through linguistic devices such as “expressing surprise or support”, “supplying helpful information” and “asking questions”. Eggins and Slade (1997) explain that while native speakers are already aware of the structures of spoken interactions and successfully fulfil a turn-taking role, in teaching second-language learners it is essential that this “native” knowledge is made an integral part of syllabus design.

The notions of identifiable spoken genres and the use of model texts for analysis with students in class (for example Nicholson and Butterworth, 2000) create a valuable starting point for utilizing the genre-based approach in a syllabus designed to enhance students’ spoken ability. The generic structures and grammatical patterns associated with various text types help form the basis for a set of criteria that are specific to particular spoken text types. As previously stated, some researchers and teachers believe that fluency and accuracy are in opposition. The genre-based approach offers a methodology that addresses this concern through the integration of grammar.

2.9.1 The role of grammar in a text-based syllabus

Section 2.3.3 discussed the different types of grammar teaching, how knowledge of grammar is an essential part of communicative competence (Canale and Swain, 1980), and the role of grammar teaching in a Japanese context. Some of the metalanguage and theoretical concepts behind a functional approach to grammar have also been discussed. In the Japanese context, the importance of having a role for grammar in teaching has also been highlighted. In this section, the role of grammar in a text-based syllabus will be discussed. As previously mentioned, the role of grammar teaching in parts of the Western world has been undervalued (Joyce and Burns, 1999). Although this has not necessarily been the case in Japan, traditional

methods of grammar teaching have remained; these approaches often seem in contrast to communicative approaches. Indeed, grammar study does not seem to have evolved beyond traditional guidelines. Joyce and Burns (1999) offer a way of “reintegrating grammar through text-based syllabus design”, by first asking such questions as:

1. Which aspects of grammar do students already know?
2. Which aspects will they need to learn about in the course?
3. Where should I start and how should I sequence the programme?

These questions provide an easily understandable framework for starting to plan a genre-based syllabus. Joyce and Burns identify a further seven areas of grammatical features worth identifying when asking the above questions. These comprise:

- The staging of the text: e.g. The “Orientation” stage might contain such grammar items as relational processes and linking verbs.
- Clause structure: Do texts utilize sentences with more than one clause? What elements are included in the theme position?
- Lexical choices: Technical or everyday vocabulary; vocabulary of judgement or attitude; descriptive vocabulary.
- Processes: Particular types of verbs used; different verbs used in different stages of the text; modal verbs.
- Noun Groups: Does the text utilize extended noun groups?
- Circumstances: e.g. Prepositional phrases of time, place and manner.
- Cohesion: Does the text utilize particular types of conjunctions?

Together with Table 2.9, Joyce and Burns (1999) thus offer a useful starting point for creating a genre-based syllabus that utilizes grammar and whole texts. Teachers are thus provided with a basis for presenting students with a syllabus that addresses both fluency and

accuracy in language use. These ideas will be adopted in the creation of the genre-based syllabus in this study, thus ensuring that teaching and learning are based on sound theoretical definitions.

2.9.2 Cycles of language teaching and learning

Butt et al. (2000) focus on speaking and writing in a cycle of activities, and provide useful information on planning students' learning by giving a clear idea of the knowledge and skills that need to be incorporated into a syllabus. This cycle includes presenting students with authentic spoken texts, in order to give them correct models and allow them to practise through interacting with other students in the class. Butt et al. (2000) provide further questions that will be considered when designing the genre-based syllabus to be utilized in this study; these questions mirror those of Joyce and Burns. When choosing a text for students to study, Butt et al. suggest asking:

- What is the purpose of the text and what is its context?
- What is the overall organization of the text?
- How does the text achieve its cohesion?
- What are the main grammatical features of the text?
- What are the main vocabulary choices?
- What comments can be made about how a spoken text is uttered?

These questions will be answered in Chapter 3, in which the main rationale for the syllabus to be used in this study is constructed.

Feez and Joyce (1998) describe the genre-based approach as the most effective methodology for use in text-based teaching and learning. Callaghan, Knapp and Noble (2014) suggest that within this approach is a teaching and learning cycle that enables teachers to

focus on genre, spoken texts, grammatical features and vocabulary choices in a range of contexts appropriate to the students. This methodology is pictured below:



Figure 2.5 Stages of the teaching/learning cycle, from Callaghan, Knapp and Noble (2014)

Chapter 4 will illustrate the syllabus design for this particular study in more detail, with reference to this teaching and learning cycle. In this section, ideas outlining design choices behind a genre-based approach will be discussed. Joyce and Burns (1999) and Butt et al. (2000) describe useful guidelines and questions to inform syllabus design. Eggins and Slade (1997) and Slade and Widin (2004) focus on spoken genres, and their research offers ideas regarding specific criteria for genre types to be studied in class. Finally, Feez and Joyce (1998) and Butt et al. (2000) offer a methodological cycle of teaching and learning that further assists syllabus planning. These concepts will provide the theoretical underpinnings of a genre-based syllabus that will form the basis of the action research design in this study. The

syllabus that will initiate the first cycle of the action research process will be illustrated in greater detail in Chapter 3.

2.10 Current Research

When examining instances in the literature of current research that applies the genre-based approach to teaching, the majority of cases stemmed from an ESL context in Australia. Joyce and Slade (2000) state that the development of teaching materials for low-proficiency students has been hampered by the belief that casual conversation is too fragmented and unsystematic for use in classrooms. This is despite the fact that in the ESL context, the ability to initiate social exchange and its interpersonal elements might mean the difference between integration and social exclusion. As such, it is vitally important that further research is conducted that focuses on this approach and its pedagogical efficacy. This is particularly true of the EFL context, where research is also difficult to identify, and the genre-based approach still struggles to gain widespread acknowledgement. In *Teachers' Voices*, Joyce and Slade (2000) present five different teacher-centred research projects that touch on a variety of issues. These include the simplification and introduction of authentic texts, which are argued to be essential components within a genre-based approach. The inclusion of simplified texts and authentic texts raises many issues when sharing this approach with low-proficiency students; these will be explored further in this study. Assessment procedures and the role of interpersonal skills are also introduced, and a number of issues are raised that suggest the difficulty of such an approach in a classroom with low-proficiency students. However, these studies do not necessarily aim to strengthen research in the wider field of English language teaching, or to stand up to peer-reviewed scrutiny; rather, they are primarily concerned with the immediate teaching contexts within which they are undertaken. Much of the academic literature associated with the genre-based approach appears to be concerned with the pedagogical principles of the genre-based approach to language teaching; but more systematic explorations of its practical applications are necessary.

In the EFL context, research in the literature appears to be even more scarce, particularly in regard to casual and pragmatic conversation. A number of PhD theses have attempted to focus on the teaching of the genre-based approach in conversation, such as Kunyot (2005) in Thailand, and Rivera (2012) in Columbia. Kunyot found that much research into the genre-based approach was undertaken in the ESL context, and as such, specific problems arose in the EFL context concerning issues of pronunciation and comprehensibility of authentic texts. The study was also limited in its data collection, due to the limited number of participants in the context of the study. Rivera points out that the genre-based approach in an EFL context needs to take into account cultural practices in the specific EFL context, and the need for carefully planned assessment practices; but he states that research in the EFL context has been neglected. His study is primarily concerned with a one-off lesson utilizing the genre-based approach, and does not fully explore assessment practices or the sequencing of content in a syllabus-level investigation. Although these two studies show an increasing interest in the application of a genre-based approach in the teaching of speaking, much more research is necessary in order to gather explicit conclusions regarding its efficacy.

2.11 Alternative Approaches

Hyon (1996) identifies three different schools of genre-based pedagogy: the first is labelled the “Australian” systemic functional linguistics methodology, which is adopted in this thesis. Hyon further identifies genre-based pedagogy in English for specific purposes (ESP) analysis, most particularly the work of Swales (1986, 1990); and finally, that of North American New Rhetoric studies. Swales (1990) summarizes the ESP approach as helping students “to develop their academic communicative competence”, with his book *Genre Analysis* being intended to offer “an approach to the teaching of academic and research English”. Swales’ pedagogical approach is exemplified by his “Creating a Research Space

Model” (1990), based upon his analysis of journal articles. The model attempts to explain the patterns and organization of scholarly research studies.

Swales’ approach appears in many ways to offer the same pedagogical opportunities as the genre-based approach adopted in the present study, with an emphasis on real-world texts and task-based classroom activities that emphasize a consideration of audience and the linguistical features of different genres. However, I concluded that the approach adopted in this study would allow a more systematic approach to classroom pedagogy, through the teaching and learning cycles emphasized in Section 2.9. The focus of the classroom activities was also removed from an ESP context and formal academic writing, being more focused on casual conversation. The New Rhetorical approach also presents a pedagogical approach with its basis in genre analysis, but which appears less concerned with a genre’s formal features and more on what makes communication effective in its social context. The goals and educational context of this particular study were best served by the “Australian” genre-based approach; in particular, within the context of low-proficiency EFL students. The systematic approach adopted in this thesis reduced the need for my students to engage with complicated meta-language that would prove more difficult to comprehend than the target language that was hoped to be learned in the classroom objectives. My own deeper understanding and past experience of the Australian approach also meant that I felt far more confident and capable of designing schemes of work for practical classroom use.

For similar reasons, I also decided not to pursue a scheme of work focused on the findings of Conversational Analysis (CA). I was concerned that the meta-language necessary to introduce such an approach to students was too complicated for my students’ proficiency level, and the large amount of language content needed in order to model conversational discourse to students was beyond their ability. However, Wong and Waring (2010) present ways in which CA can be used by EFL teachers aiming to improve their students’ oral

proficiency, particularly via an analysis of the structure of conversation. Wong and Waring argue that very few pedagogical materials address the turn-taking system, and this was of particular concern in my own teaching context. Despite choosing not to focus on a CA approach to lesson design, I was interested in adopting some of the turn-taking pedagogical approaches highlighted by Wong and Waring into my overall genre-based methodological approach. This also highlighted the notion of a genre-based approach as a model of teaching that is not adopted as a discrete and unique approach to pedagogy, but as an overall framework of teaching and learning that can utilize a number of different approaches depending on students' needs.

2.12 Summary

In this chapter, the historical context of English language education in Japan and the wider world was discussed. Language instruction has followed trends in discrete packages throughout the post-war period in Japan, with each new approach sometimes being adopted in complete opposition to what had come previously. Behind these changes in approaches, however, has remained one constant and over-riding factor: namely, the university entrance exams. The university entrance exams and highly valued standardized tests such as TOEIC have created tension between MEXT guidelines for instruction and what teachers feel is necessary in reality, particularly in relation to communicative approaches.

The theoretical underpinnings of a genre-based approach were examined in preparation for the research design of this particular study, including questions raised about the integration of assessment into such a syllabus, and potential conflicts that may arise. Alternative approaches to the genre-based approach were also introduced, with an explanation of why these approaches were not chosen as the primary focus of this research. What appears to be missing from the literature is a wide body of research that systematically and scientifically explores the practical aspects of the pedagogical approaches associated with

a genre-based approach. It is hoped that this study will add to the growing body of literature on the subject, particularly in the EFL context.

Chapter 3: Methodology: Research Perspectives

3.1 Introduction

This chapter outlines the main research perspectives that underpin the research design in this study. The first section discusses the idea of thinking about research at different levels of abstraction. This provides a framework that is first used to describe the main characteristics of different research perspectives, and then situates the present study in the paradigm of critical theory. The selection of an action research approach is discussed, together with a discussion of the limitations of such an approach and the design of the research conducted in this thesis itself. Finally, a more detailed description of the participants in the study is conducted, with an example dialogue aimed at illustrating their speaking proficiency in English.

3.2 A Framework for Thinking about Research

This thesis fits broadly within a perspective of social critical theory, adopting an action research approach to data collection and analysis. Although action research is a methodological approach, this discussion of perspective and approach is not deterministic; the research also draws from aspects of other perspectives and their methods and methodologies where appropriate. Johnson and Onwuegbuzie (2004) posit that researchers should put considerable effort into answering two questions as a starting point to research: “What methodologies and methods will we use?” and “How do we justify these choices?” The first part of this chapter responds to these questions, firstly by providing an overview of the different methods and methodologies employed by researchers; secondly, by positioning the current research within these different research perspectives; and finally, by describing why the research has been designed in this way.

McDonough and McDonough (1997) describe four levels of abstraction for thinking about research. These abstractions include the epistemology and ontology of research: that is,

the philosophical viewpoints regarding the nature of reality and knowledge. Part of this level of abstraction is the research perspectives: these are the broad traditions of research, each of which makes different assumptions about the nature of reality and the kind of knowledge that is produced. A third abstraction is the research approaches and methodologies: in other words, the conceptual framework and principles through which a piece of research proceeds. Crotty (2008) describes this abstraction as a theorization and a justification of the methods and procedures used to arrive at valid knowledge. The final abstraction is the research methodology, which means the specific techniques for collecting data: these include statistical analysis, surveys, interviews, observations, and so on.

Crotty (2008) argues that at the research perspective level of abstraction, the researcher is likely to decide the types of research questions they find relevant and interesting. These theoretical assumptions are likely to shape the next abstraction in the framework: research approaches. Having selected a research approach, the researcher then chooses specific methods of data collection. Johnson and Onwuegbuzie (2004) point out, however, that the relationships between the various levels of research are not deterministic. For example, an interest in a particular research topic may lead the researcher to choose an approach that is typically aligned with a different research paradigm from the one the researcher is comfortable with.

The research perspective under which a researcher works, and the research approaches and methods that are chosen, depend upon the views held by the researcher about the nature of reality and knowledge. The next section discusses these two theoretical perspectives of reality and knowledge.

3.2.1 Ontology

Cresswell (1998) describes ontology as the study of reality, and highlights the difficulty of exploring ontology because of the possible combining of ontological and epistemological issues. Cohen, Manion and Morrison (2013) define the two ends of the ontological spectrum:

1. Objectivity: The belief that there is an objective truth that everyone would agree upon.
2. Subjectivity: Reality is personal and socially constructed, and there are multiple realities. Our own beliefs prevent us from observing something objectively.

Cohen et al. (2013) claim that the above separation reflects the realist–nominalist debate. This debate is characterized by realists, who accept the notion of universals, and nominalists, who posit that ideas, which are represented by words, exist only in our minds and have no real existence. This dichotomy, offered by Cohen et al. (2013), is rejected by a phenomenological solution to this debate, presented by Systemic Functional Linguistics (Halliday and Matthiessen, 2006): this solution describes experience not as knowing but as meaning, and hence as something that is construed in language. In this belief, the concern is with the construal of human experience as a semantic system; and since language plays the central role not only in storing and exchanging experience but also in construing it, language is taken as the interpretative base.

3.2.2 Epistemology

Epistemology is the study of knowledge (Cresswell, 1994). Cohen et al. (2013) claim that beliefs about knowledge, such as what it is, where it comes from, and how it is acquired, are numerous. It is a researcher's epistemological assumptions that influence how they undertake and report research.

Similarly to the case of ontology, Cohen et al. describe polar opposites of understanding:

1. Knowledge is hard, measurable, tangible and objective.
2. Knowledge is personal, unique and subjective.

According to these definitions, therefore, there are two different views of reality and knowledge. One is objective: reality and knowledge are stable, neutral and free of human values. Alternatively, reality and knowledge are subjective; they are not neutral but influenced by human values and beliefs. It is because of these ontological and epistemological views that different research perspectives exist. Creswell (2013) outlines five different research perspectives and their assumptions: positivism, interpretivism, critical theory, feminism, and postmodernism. The next sections outline these different perspectives, and position the current study within the critical theory perspective.

3.3 Research Perspectives

The previous section discussed a possible framework for thinking about research at different levels of abstraction (in terms of perspectives, approaches and methods). Concepts of epistemology and ontology were discussed to highlight how a researcher's assumptions about the nature of reality and knowledge influence their research perspective, and subsequently the research approach they tend to apply. The next section discusses the five research perspectives described by Creswell (2013) and McDonough and McDonough (1997). This discussion follows the level of abstraction previously described, firstly by discussing the history of the paradigm, and then the predominant research approaches in each paradigm. The discussion then situates the current study within the framework of abstraction offered by Creswell (2013) and McDonough and McDonough (1997), among others.

3.3.1 Positivism

This section describes the main features of positivism and some of the key features implied. Chappell (2007) provides a brief history of positivism, from the discovery of the "New World" and the Renaissance to the present day. Chappell explains that the positivist

perspective began with Galileo and continued with Newton, who undermined the previously spiritually governed world by arguing that science is the primary investigative tool required to understand the world; and that through science, humans are capable of determining a set of absolute rules that interact with each other in predetermined and predictable ways.

The historical development of positivism is further presented in the modern context by Sorell (1991), who claims that the nineteenth-century philosopher Comte is generally associated with the emergence of positivism as a dominant discourse within the Western intellectual and cultural tradition, and its position as the dominant research paradigm of the social sciences. Sorell states that many researchers still insist that educational research must follow the methods and methodological framework of positivism. This positivist viewpoint has been resisted by some social researchers, particularly humanist sociologists, such as Mills (1963), Weber (1978) and Parsons (1971). Despite this resistance, it remains the dominant paradigm within the social sciences (Wagner, 1994).

Cohen et al. (2013) identify the underlying principles and assumptions of today's interpretation of positivism, including the following: its close links to empirical science; its adherence to a philosophy of science that sets few bounds on what science is capable of achieving; its contention that scientific knowledge is both accurate and certain; its view that positivism is entirely objective; its claim that a positivist's world is a "mathematised" (p. 27) world; and its view that the world is perceived as highly systematic and well organized (full of regularities, constancies, uniformities, iron-clad laws and absolute principles). Crotty also notes Husserl's (1970) view that the positivists' world is an abstraction from the "lived world" (p. 28).

Two predominant research approaches fall within the positivist paradigm: surveys and experiments (Cresswell, 1994; Wiersma, 2000). Cresswell (1994) notes that established procedures of statistical analysis are typically associated with this form of research. He

provides a checklist of questions for designing a survey or experiment. These questions make reference to control groups, measurement, treatment conditions, validity and reliability, tests and statistics. Wiersma introduces a number of key concepts that he sees as central to the notion of good research, including controlling variance, dependent and independent variables, randomization and bias. We see that positivism encompasses the realist ontology and objectivist epistemology. As previously described, Wagner (1994) describes positivism as the dominant paradigm in the social sciences. However, the definitions outlined in this section do not adequately express the ontological and epistemological beliefs that are apparent throughout the research design. Another research paradigm identified in the literature is that of interpretivism, which also contains some of the underlying assumptions apparent in this thesis.

3.3.2 Interpretivism

The term “interpretive” covers a diversity of research traditions and methods (Williams, 2000). Broadly, these traditions are called “hermeneutic” (Packer, 1989), after the kind of scholarship that attempts to explain the Bible. The term also includes historical research, which takes as its main object of inquiry the meanings of past events that are interpreted using primary and secondary accounts of those events. Interpretive research also covers major research traditions in sociology and anthropology that take as their main object of inquiry an understanding of the cultures of social and ethnic groups (Crotty, 2008; Spindler, 2014; Tobin, 2000).

Interpretivism emerged from the perceived failure of positivism in social and educational research to generate generalizable knowledge and predictive insights, combined with the increasing acceptance by the natural sciences that knowledge is always tentative, relational and conditional (Chappell, 2007). According to Crotty (2008), interpretivism questions notions of objectivity and external realities that are constituted independently of the

subject, and proposes that subjectivity is critical to the emergence of understanding and explanation in the social world. An interpretive epistemology emphasizes social practices as the focus of social research, and makes the assumption that all social practices are meaningful for those involved. Interpretivism also criticizes the positivist approach to investigation, which sets up artificial and simplistic experiments to prove connections between agents of cause and effect. According to Scott and Usher (1996), social phenomena are more usefully seen as open and intermediate, so that predictive generalizations are not possible unless the closure that is necessary is imposed from outside. Interpretivism therefore focuses on interpretation, meaning and illumination rather than generalization, prediction and control.

Ethnography is a major interpretive approach to research (Watson-Gegeo, 1998). Esland (in Young, 1970) and Watson-Gegeo (1998) describe ethnography as an approach that turns attention to understanding classrooms as social settings, where teaching and learning can be understood as constructed through inter-subjective meanings of teachers and students. This perspective opens the way to analysing teaching, learning and curriculum as institutionalized realities. Hammersley and Atkinson (1983) state that the central activity of ethnographic research is observation and interpretation. Observation is carried out “in context” in “naturally occurring settings” (such as classrooms). This is in contrast to “non-natural contexts” of interviews or set-up experiments. The aim of ethnographic research is therefore to investigate things as they are in real-life settings. Data gathering is done “in the field”, with naturalistic, non-experimental settings. Hammersley and Atkinson (1983), however, problematize the notion of the naturalistic setting, proposing that naturalism as well as positivism both presuppose a notion of the objective detached researcher. They argue that the intrusion of the researcher into the research setting shifts and alters that setting in unpredictable ways. This study also rejects the notion of the objective, detached researcher. This seems particularly true for classroom contexts where the teacher (as a researcher) plays such a commanding role. The research paradigm through which this study is conducted does

not, therefore, seem adequately defined by either positivism or interpretivism, despite the fact that some aspects of the assumptions evident in the research design constitute both approaches. The next section discusses the paradigm of critical theory, which seems best suited to describe the epistemological and ontological beliefs that underpin the research design.

3.3.3 Critical theory

The term “critical theory” was coined by Max Horkheimer, who became director of the Institute of Social Research at the University of Frankfurt in 1930. The Institute and its beliefs are known as the “Frankfurt School” (Geuss, 1981). According to Horkheimer (1972), Social Critical Theory is an idea of a future society whose members form a community of free people limited only by the technology of the time. Its focus was to analyse the social world under a variety of auspices, such as philosophy, literature, art, education, economics and politics, by asking whether thought and action were autonomous, and if theory had any practical significance.

Because of its beginnings as a means of opposing social injustice in modern society, critical theory seeks to expose and critique current practices, assumptions and theories (Crotty, 2008). It does not offer a neutral set of judgements, but makes the status quo problematic, in order to affect change. The social critical theory critique of positivism is that it relies on methods for the clarification of natural phenomena by removing uncertainties, yet at the same time leads to the acquisition of unproblematic beliefs. From the Social Critical Theory point of view, interpretive research methods have limitations, as they cannot account for the possibility that research might diverge from pre-conceived ideologies that may not be liberating for all people.

Social critical theory differs from positivist and interpretive perspectives as it problematizes the unproblematic nature of both research perspectives, by pointing out that

they do not contribute to freedom of thought for those who do not agree with dominant ideologies. In simple terms, social critical theory research questions and clarifies social values, through identifying social structures and ideologies that influence and dominate society; furthermore, it makes the knowledge gained from the identification of influential and dominant structures and ideologies problematic, by critically reflecting on their systems of meaning.

Kincheloe and McLaren (2011) provide a useful summary of the basic assumptions underpinning critical theory, as follows: that all thought is constituted by power relations that are social in nature and historically constituted; that facts can never be isolated from values or ideology; that the relationship between concept and object, and between signifier and signified, is never stable, and is often mediated by the social relations of capitalist production and consumption; that language is central to the formation of subjectivity, that is, both conscious and unconscious awareness; that certain groups in society are privileged over others, constituting an oppression that is strengthened when subordinates accept their social status as natural and inevitable; that oppression comes in many forms, and concern for only one form of oppression can be counterproductive because of the connections between them; and finally, that mainstream research practices are generally guilty of reproducing systems of class, race and gender oppression.

This description of social theory is mirrored by the beliefs that underpin an action research methodological approach, which Burns (2005, 2011) describes as being concerned with notions of democracy and social justice. Thus, action research is a major approach emerging from critical theory; it is also the main approach of this study. Specifically, this study aims to problematize the status quo, as well as historical and dominant educational ideologies and classroom practices, including the practices of the researcher himself. Action research methodologies used in relation to this study are outlined and critiqued in section 3.4

and onwards. Before this, the next section provides a brief survey of the final two alternative research perspectives, namely feminism and post-structuralism.

3.3.4 Feminism and post-structuralism

Feminism converges with post-structuralism in critiquing the epistemology bequeathed by the European Enlightenment (Malson, 2003). Luke and Gore (2014) describe feminism as a “re-vision” (p. 149), challenging the production of meanings and explanations in social activity from which women have been marginalized. They describe this motivation as ranging from an awareness that the playing field on which women compete is far from level, to a radical change in culture and society. According to Lerner (1986), feminism asks how epistemological categories are implicated in defining masculinity and femininity; how they function to define the “nature” of people; how they work to attach differential valuations to their skills and capacities; and how gender difference is a category of analysis around which every society is structured.

A post-structuralist approach is sceptical of dominant beliefs concerning truth, knowledge and power, which are seen as aspects of contemporary Western culture that shape modernist thought (Collinson, 2006). The argument is that knowledge is structured through hierarchically organized oppositions, and that modern epistemology privileges the masculine. Thus, empirical/analytic epistemology’s claim that rationality, objectivity and abstraction is the only guarantee of truth is actually a specifically masculine claim.

Positivist approaches are problematic in these paradigms because they rely on a pre-existing reality. If the assumption of a pre-existing reality is called into question and replaced by an assumption that reality is constructed through discourse, there are a number of consequences, such as issues of power and openness to change. Scientific method is also undermined by the notion of discourse, because science is seen as a social practice and can therefore never be objective and neutral. Scott and Usher (1996) posit that interpretivism

retains features of positivist enquiry in the idea that there are universal and generalizable truths concerning human subjects that exist outside the situated discourses that interpretivism uses in its construction of subjectivity. Feminist and post-structuralist beliefs appear to mirror closely those of critical theory. However, according to Crotty (2008), critical theory is also problematized by these perspectives because of its tendency to raise particular discourses to the status of universal truths, thus implying a pre-existing social reality. Nonetheless, Agger (1991) suggests that it is possible to forge links between critical theory, postmodernism and post-structuralism in order to democratize science; therefore, we should not to cast our arguments in positivist terms of pure representation, but those which recognize theories as being able to be subject to revisions and improvement. This study takes the view that current practice, when subjected to questioning and re-questioning, as is evident in an action research approach, leads to change that ultimately seeks to achieve improvement in the classroom.

This chapter commenced by developing a framework for thinking about research at different levels of abstraction. Concepts of epistemological and ontology were explored, together with the question of how a researcher's assumptions about the nature of reality and knowledge influence their research perspectives and approaches. The five predominant research perspectives in educational research were outlined in order to situate the current study within the critical theory perspective, but with links forged to post-structuralism. The next section describes the action research approach and the purpose of adopting this approach for this study.

3.4 Why an Action Research Approach?

This study defines action research as a methodology, and as such, does not take a deterministic or dichotomous approach to data collection and analysis. This section expands on the reasons for adopting an action research approach: firstly, by defining action research and how this definition aligns with my personal aims for conducting the research; secondly,

by outlining my background and beliefs; and finally, by highlighting the limitations of this approach in relation to the research questions.

Reflecting social critical theory and aspects of post-structuralism, action research is an approach to research that is focused upon the following aspects: “change” (Denscombe, 1998, p. 58; Cohen et al., 2013); “problematizing” (Burns, 2010, p. 2; Freire, 1976, in Crotty, 2008, p. 156); “cycles” (Denscombe, 1998, p. 58; Crotty, 2008, p. 157; Kemmis and McTaggart, 1998, in Burns, 2000, p. 8); and “participation” (Denscombe, 1998, p. 58). These elements of action research are the fundamental reasons why it was adopted as the main research methodology underpinning this study. As outlined in Chapter 1, the rationale for this thesis began with a desire for change in the status quo by problematizing what I had been doing previously. I wanted this change to be systematic and reflective, involving cycles of change based upon reflective observations and feedback from students. The researcher in this study is also the teacher, who is a critical component of the immediate context and therefore a participant in the research project itself; this contrasts with an objective and impartial outsider searching for universal truths, as might be evident in an ethnographic approach that calls for a detached and objective interpretive system of observation.

Change refers to a way of dealing with problems by discovering more about them, and forms an integral part of research (Denscombe, 1998). In this research, I wanted to discover more about my students’ problems with speaking in English, and ways of changing what currently happens, by using an alternative approach. I would then use action research to explore the impact of the changes I made, through data collection and analysis. Chapter 2 highlighted the problems faced by Japanese students in the current context, both local and national, when communicating in English. As the status quo appeared to be inadequate in addressing these problems, I felt that imposing change would allow me to explore these problems in more detail, and offer alternative routes for my own personal teaching context.

Problematizing refers to taking an area of your teaching you feel could be done better, and subjecting it to questioning by developing new ideas and approaches (Burns, 2010). However, problematizing does not mean a simplistic case of problem solving. Allwright and Hanks (2009) argue that current research models are often wanting in their ability to aid educators' understandings of the language learner. They state that divorcing researchers from practitioners and practitioners from learners creates an attitude of counterproductive mistrust, and incomplete findings. In this case, I felt that the way in which my students learned English could be improved, but I wanted to address this issue in a systematic way. By taking an action research approach, I felt I could raise awareness of the complexities my students face when trying to speak. By experimenting with a genre-based approach, I could examine my current practices and record scientifically whether any changes had an impact. I had attempted change before in an unofficial way; but rather than relying on assumptions and hunches, I wanted to use systematically collected data to explore my own context and changes to my own context.

Cycles in action research refer to the feedback loop in which initial findings generate possibilities for future change, which are then implemented and evaluated as a prelude to further investigation (Denscombe, 1998). Crotty (2008) refers to "spiralling" (p. 157) rather than cycling, to highlight the reciprocal relationship of reflection and action. Burns (2010) adapts the idea of a cycle by referencing Kemmis and McTaggart (1988) and their four stages of the cycle: planning, action, observation and reflection. I felt that this model provided a useful framework for planning, conducting and preparing my research for publication. The planning (Burns, 2010, p. 8) stage provided guidelines for narrowing the extent of my research: this included identifying a problem and narrowing down the focus into a research plan that was possible within the constraints of my workplace. These constraints included such factors as time; ethical considerations; what data collection procedures I could use that were appropriate for my students; and what benefit my students would gain from any

research I conducted. This allowed me to design a syllabus and series of lesson plans that were informed by the literature but were also appropriate to my particular context. I was also able to consider various avenues for data collection. This plan was then put into action (Burns, 2010, p. 8), and I was able to document my “Observations” (p. 8). At the “Reflection” (p. 8) stage, I could then evaluate the effects of the action, re-plan, and also begin to write my thesis. The subsequent chapters in this thesis detail these processes chronologically. The choice of action research gave me the confidence and tools I needed to address my research questions in a systematic way appropriate to my context. However, as previously stated, these plans could be changed at any moment, either through planned reflection, or in reaction to the immediate environment, specific participants, and other contexts that are explained in more detail in the results chapters.

“Participation” refers to the fact that practitioners are the crucial people in the action research process. Their participation is active, not passive, as might be noted in alternative perspectives such as positivism and interpretivism. Cohen et al. (2013) highlight a tension between teaching and research agendas, which action research explicitly challenges by re-working the theory–practice divide. Action research addresses this challenge by emphasizing issues of practice that are directly relevant to the concerns of teachers. Cohen et al. describe action research as a small-scale intervention in the functioning of the real world, and an examination of the effects of this intervention. Burns (2010) describes this process as a “focus on practical theory” (p. 133).

The idea of developing my own ideas of theories for practice (Burns, 2010) was very important to me. Tensions I had often felt in balancing the time I gave to theory and to practice were somewhat reconciled by adopting an action research approach. Ultimately, I wanted the results of any research to benefit my teaching and my students’ learning, and this provided my primary motivation for undertaking any project.

Cohen et al. (2013) develop this idea of participation further by identifying four further characteristics of action research that hold resonance with my own beliefs. These are that the research is situational: i.e. concerned with diagnosing a problem in a specific context, as well as attempting to solve it in that context. It is usually (though not always) collaborative: teams of researchers and practitioners work together on a project. In my own context, I was asked to be an editor of the university's journal. As I was a PhD candidate, I was therefore expected to share my experiences with other members of the department through this role. I invited other members to read my research proposals and participate in the syllabus I designed. Together we have formed an "Action Research Group" that carries out smaller-scale interventions corresponding with teachers' own research interests. Cohen et al. (2013) describe action research as participatory: team members themselves take part in implementing the research, and as such it is self-evaluative. Modifications are introduced into the ongoing situation and continuously evaluated within the ongoing situation, the ultimate objective being to improve practice in some way.

A key motivation underpinning my research was professional development and developing as a teacher. By being participatory in nature, action research therefore allows the teacher to become reflective, and instigates professional development (Burns, 2010). Burns argues that "doing AR (action research) can reinvigorate our teaching, lead to positive change, raise our awareness of the complexities of our work, and show us what drives our personal approaches to teaching" (p. 7). The personal approach to research appears in antipathy to a positivist perspective that knowledge is objective and certain. At this stage, I will provide a short recount of my own experiences with research, in order to situate the present study at the personal level of abstraction; this is in an attempt to illustrate how my own epistemological and ontological beliefs have been shaped by experience.

3.4.1 My research experiences

Since I began teaching after graduating from university, I had always considered that the main part of my job was to teach and develop as teacher by being eclectic and experimental in my teaching approaches, and thereby to learn by trial and error. This was an ideal that was emphasized during our teaching training, and one that has motivated me ever since. For similar reasons, I would also strive to select approaches that best matched specific groups of students, often through differentiation in individual classrooms. Differentiation involves providing students (often within the same class) with different avenues for learning and assessment, so that all students can learn effectively, regardless of ability. This viewpoint indicates an ideology that reflects the critical theory perspective as outlined previously; and in action research approaches, as Burns (2010) describes, it reflects “the democratic and social justice philosophies that underlie AR” (p. 131).

After attaining a job at a university as a lecturer, I was also expected to contribute to the research contributions of the Language Education and Research Centre (LERC) of my place of work. I felt the tension between theory and practice, as was highlighted between Cohen et al. (2000). As Burns (2010) and Cohen et al. (2013) describe, action research is a way of resolving this tension. It allowed me to conduct research that was immediately applicable to my context and students; it improved current practice; and it enabled me to develop professionally.

However, one aspect of the LERC that I encountered immediately was the strongly held belief in the positivist approach, and the necessity to uncover universal truths in published research. As a new member of staff, and also a new researcher, this prevailing perspective was heavily influential. I began to look at research in terms of measurement, controlling for variance, reliability and validity, generalizations and objectivity. As such, I also enrolled in statistics courses and learned various statistical procedures. Because of this, the tension I felt

between theory and practice increased. I often felt that my research, or research I collaborated in, was not positively contributing to my students' learning, or was taking up too much of the time I could have been devoting to my classroom procedures; especially given that a great deal of this research was based upon summative testing. However, since I instinctively felt the need to experience a range of approaches to my teaching, I persevered with many positivist approaches during my early career in the LERC. I hope that in this thesis I have been able to exploit the methodological nature of an action research approach to effectively combine these positivist influences.

As I became more confident in my own perspectives, became an integral part of many successful research projects at the LERC, and began my PhD study, I began to readjust my research perspectives. I became an editor of the LERC journal and introduced submission guidelines for qualitative research, whereas previously only quantitative research had been promoted. I was able to decrease my feelings of tension by re-adopting a critical theory perspective and adopting action research methods. However, my dalliance with positivist approaches was also a positive one. Many of the statistical procedures, which were primarily used as measurement tools, I found were also very useful diagnostic tools for identifying problems with individual students or elements of a syllabus: they therefore highlighted opportunities for qualitative exploration. This is reflected in this study by my application of quantitative data collection and analysis, primarily through the Rasch model.

I considered a positivist approach for this research, but did not choose it, for the following reasons: I was not comfortable with the notions of measurement that this approach often entails. In this regard, a key part of the syllabus design eventually integrated the assessment of learning strategies. Although this does not preclude summative assessments, I did not want to make summative assessments and measures my primary goals, as this would allow the research to become objective and mathematized. I had a preference for content that

focused on my specific students and teaching situation. Furthermore, I felt that controlling for variation was extremely difficult and also raised ethical concerns. During a collaborative research project that we conducted in the LERC concerning vocabulary acquisition (Anderson et al., 2012), each teacher or group of teachers consulted the literature for possible vocabulary learning interventions that could be introduced in class. The success of these different interventions was to be based on pre- and post-tests of knowledge of frequent vocabulary items. During the planning of this research we considered different areas of variance that we should attempt to control: time on task; the different teachers; age of students; sex of students, etc. Compiling a list of variance seemed arbitrary and endless. For example, would students be affected by studying after lunch? What if students had problems at home? In the conclusion of our research, after controlling for as much variance as possible, we discovered that it was the teacher, not the intervention itself, which appeared to have the greatest effect on students' vocabulary knowledge. Although this was an interesting conclusion, we also concluded that there might have been better ways of addressing the research questions. Controlling for variance also led to the creation of "control groups", and this raised ethical concerns. I was concerned about the fact that we were trying to improve students' vocabulary knowledge by introducing our own interventions; yet students in the control group were doing something that was not based on an intervention that was designed to improve current practice, but simply to continue current practice. As I entered the planning stage for the current research, the same questions arose, and I decided that an approach that was scientific and rigorous, but which also included epistemological and ontological notions of subjectivity, was vital.

This discussion of my beliefs and my background elaborates my reasons for choosing an action research approach; however, it also highlights certain limitations of my research, some of which are embedded in action research itself.

3.4.2 Limitations of the action research approach

Denscombe (1998) elaborates further advantages of action research by describing this approach as addressing practical problems in a positive way; an outcome, he claims, that has failed in social sciences research to the present. He also claims that it should benefit the organizations within which the research takes place, by improving practice and resolving problems. This may even be extrapolated to larger contexts, such as national educational policies. Denscombe also points out that participation in the research by practitioners democratizes the research process and enhances respect for practitioners' knowledge.

Denscombe (1998) also highlights the limitations of action research. The insider knowledge that action research entails can be a "genuine bonus" (p. 63), but it also poses some problems. The "insider" may often overlook something that is considered mundane or obvious, but which might register as an important factor to an outsider. Denscombe therefore suggests that an outsider "expert" read over data and data analysis at different stages of the action research process (p. 64), as such an expert can offer alternative perspectives. This insight is echoed by comments that can be found in Chapter 7.

Denscombe also highlights what Burns (2010) describes as the "thorny issue" (p. 191) of validity in action research. It might be argued that the local scope of action research means that findings will rarely contribute to broader insights. Thus, while a research project such as this might assist my teaching in the local context, it is far more difficult to make generalizations in a wider context. Features of setting and context are "givens" (p. 64) rather than factors that can be controlled or varied. Additionally, the integration of research with practice limits opportunities for exercising control over variables.

The ontology and epistemology of the action researcher is not objective; reality and knowledge are not neutral, but value-laden. Action research can include addressing perceived problems in everyday activity, and thus the practitioners have a vested interest in the

findings. Unlike positivist views of science, action researchers cannot be entirely detached or impartial. The study itself also contains other various limitations that should be clearly described in this chapter.

3.4.3 Paradigms for reporting action research

Stapleton and Taylor (2004) highlight the difficulties faced by interpretive epistemologies in deciding an appropriate thesis structure. They argue against structuring interpretive research according to a positivist paradigm, as this creates distorted portrayals of their inquiries as timeless and lacking an emergent nature. Stapleton and Taylor (2004) then present an electronic thesis in a non-linear hyperlinked format, which unfortunately does not provide a practicable model for structuring the present study. Publications such as *Teachers' Voices* (Joyce and Slade, 2000); *Classroom research in English language teaching in Oman* (Borg, 2006); and *Investigating English language teaching and learning Oman* (Borg, 2008) show excellent examples of action research written by teachers. Accounts are written in first-person chronological narratives, beginning with a description of the teaching problem and context; then the plan of action; an account of what happened as a result of the plan; and finally the conclusions. However, these examples, whilst written as part of formal study on a BA (TESOL) course, do not constitute a written approach appropriate for a PhD thesis. Zuber-Skerritt and Perry (2002) argue that little guidance is given in the literature on how to construct a PhD thesis that addresses the academic criteria of such an award. Zuber-Skerritt and Perry (2002) argue that there are two types of action research projects: one that involves practitioners within an organization, such as that exemplified in *Teachers' Voices* (Joyce and Slade, 2000), which are written in first-person narrative form; and “thesis action research” (p. 175), which focuses on theory and methodology. Zuber and Skerritt (2002) propose that the thesis action research should first define the research problem; then provide a rationale for the research; conduct a literature survey; describe the research process and procedure; analyse and evaluate the results of the action; and finally, present conclusions from the

research, together with its limitations, and suggestions for further research. Davis (2007) claims that the traditional “bulky tome” (p. 181) format of the thesis – following the traditional structure that she defines as literature review, methodology, research design, findings and conclusions – is a format under challenge. She creates an alternative structure for her PhD thesis based on action research cycles. She divides the reporting of each cycle into individual chapters, and within each chapter details the different phases of action research, such as planning and action. This is followed by an overview of the cycle, in which she describes lessons learned, and critiques of her plan.

This thesis is structured in a similar way to that suggested by Zuber and Skerritt (2002) and Davis (2007). The thesis begins in Chapter 1 with the focus of the research and the rationale behind it. In Chapter 2, a literature survey is conducted; but as suggested by Davis (2007), elements of the literature review are also included during the description of cycles of research, to illustrate the non-linear format of the action research approach. Chapter 3 describes the epistemological and ontological perspectives of my research, together with a justification for the action research approach. As with Davis (2007), my account of the action research then follows the two cycles of action research undertaken in this study. Each phase of each cycle – planning, action, observation and reflection – is then presented in a way that reflects the reality of the research itself. Where smaller cycles of action occur that do not fit the larger two-cycle format, I have included paragraphs illustrating “reflection-in-action” (see section 4.8), based on Schön’s (1983) concept of spontaneity in the action research process.

3.4.4 Limitations of this study

A number of limitations of this study became evident during the research design, the most significant of these being due to the research context. These issues included sample size, time constraints, language constraints and cultural constraints.

Firstly, the study itself uses a range of sample sizes, according to the cycles of piloting and designing the syllabus. Some of the sample sizes at various stages might be considered too small to make broad generalizations.

Next, time constraints provided a key consideration during the planning stage. Each semester, every student experiences only fifteen ninety-minute classes. This time is also devoted to broader TOEIC testing and preparation, as part of whole-school policy. Such time constraints limited considerably the opportunities for data collection. As each student is a non-English major, requesting time with students in addition to their fifteen allotted classes also proved difficult, as their main concerns lay outside their English education. The research design initially considered conducting structured or semi-structured qualitative interviews with students at various stages of using the genre-based syllabus, but time constraints (as well as language constraints) made this extremely difficult. It was concluded that written feedback in the form of student journals and other classroom documents, or non-structured talk recorded in teacher journals and logs throughout the course of the semester, was a more viable system of data collection, considering time restraints.

Language constraints presented another limitation of data collection. The gap between my Japanese language proficiency and the students' English proficiency made interviews difficult to conduct. Written responses in student journals and class documents allowed time for students to formulate more considered responses in English.

Finally, cultural constraints were also a limiting factor in research design. One cultural problem highlighted by MEXT (2003a) is that although it is vital for Japanese students to learn to communicate in English, they also struggle to communicate in Japanese, particularly in communicating personal opinions. Further cultural constraints can be categorized into two main areas: curricular and cultural (Barnlund, 1989; Neustupny, 1985; Ting-Toomey, 1985; Klopff, 1991; Wetzel, 1988). Curricular reasons include teaching

methods that promote the necessity for real-life communication, and lack lecture-centric classrooms where listening is the primary mode of communication between student and teacher. Providing feedback, and particularly critical feedback, is a new concept to many of the participants in this study. Cultural reasons outlined by Barland (1989), Neustupny (1985), Ting-Toomey (1985), Klopff (1991) and Wetzel (1988) include silence constituting an acceptable affirmative response to questioning; and implied meanings rather than explicit opinions (Okazaki, 1994). Other cultural contextual factors include issues such as formality, status, power, and social distance, which would dissuade participants from providing explicit feedback to questions (Kington, 2000). This is particularly true in a teacher-student scenario such as the one encountered in this research. Collecting data in summative questionnaires that sought students' opinions about the syllabus or classroom activities often produced responses that were sparse, as students were reluctant to question the teacher's methods or provide feedback that might be construed as negative, even when questionnaires were anonymous. Answers would often be short and vague. This discussion of the cultural context of the study leads to an ideal opportunity to present the participants in the study in greater detail.

3.5 The Participants

The majority of participants in this study were 18 to 20 years old, and had studied English since junior high school, which begins at the age of 13. There were both male and female participants, in roughly equal proportion. All participants were non-English majors, and English language classes in the setting were compulsory. Language proficiency amongst the participants was considered below the average proficiency for university students in Japan when measured by TOEIC Bridge scores (Fryer and Ozono, 2014), with an average TOEIC Bridge score of 117. Student confidence and motivation in the university as whole were also considered to be low (Fryer et al., 2014).

An important factor in this study is the difference between ESL and EFL classrooms. ESL involves instruction in the target language country where English is the main spoken language in the immediate environment. In the EFL context, the teacher is the sole provider of English input. Here, it is not usual for students to speak to each other in English, as they share the same L1. Thus, opportunities for the students to speak English outside the classroom are extremely limited at the institution, and there is very little need for them to do so in everyday situations.

English language classes at the institution are entitled “English Communication” and “English Reading and Writing”. Students must complete “English and English Communication I, II, III and IV” in order to graduate from the university. There are no syllabi or lesson plans dictated to teachers beyond these class descriptors. The compulsory taking of the TOEIC Bridge test is the only aspect of the teaching environment that might be considered as signifying that the institution had a unified curriculum, and often teachers felt compelled to include content in their classes that prepared students for this test.

Students’ English-language proficiency is below the average ability of Japanese university students, as measured by the TOEIC Bridge examination. TOEIC Bridge is a simplified version of the TOEIC test, and has scores ranging from 20 to 180. A score of 100 TOEIC Bridge points is equal to 260 TOEIC test points (ETS, 2012). The average TOEIC Bridge score of students entering the university is 103 (this increases to 117 after two years). However, the average TOEIC Bridge score for all people who sat the TOEIC Bridge test in 2012 in Japan was 130. If TOEIC Bridge is considered a legitimate measure of English proficiency, this means that students at the university are of comparatively low proficiency. It should also be noted that the TOEIC Bridge is intended for low-proficiency students who are not yet ready to take the TOEIC test, but want to prepare for it.

Having conducted a study of motivational factors at the university where this study also takes place, Fryer et al. (2014) explain that while students may be motivated to learn, they are unlikely to exhibit the same motivation to learn English as English majors. Fryer et al. found that students at the university were not motivated to study English, and held a belief that they were not good at English at the start of their study. “Amotivation” amongst students was defined by Fryer et al. (2014) as a lack of motivation by students to engage in English language learning, and a failure to appreciate the activity and its value.

3.5.1 Students’ spoken proficiency

My research questions were initially triggered by what I had observed about my students’ speaking abilities, and the concerns I highlighted in Chapter 1. Below is an example of two students’ spoken output, conducted before this study began.

Participants:

A: Teacher

B: Student (Japanese words are included in *italics*)

Student A:

1. A: How are you?
2. B: I’m fine thank you, and you?
3. A: I’m alright, I’m a bit hot. It’s hot in here.
4. B: (nods)
5. A: What have you been doing recently?
6. B: I get up at 6. I brush a tooth.
7. A: OK
8. B: Yes
9. A: Did you do anything else?
10. B: Err, come to school.
11. A: I see. Today?
12. B: Yes. Today.
13. A: What did you study today?
14. B: *Eigo*. *Eto*, err, English.
15. A: This class?
16. B: Yes. This class
17. A: OK, I see, thanks for speaking with me!
18. B: Yes.

Student B:

1. A: Hey! What have you been doing recently?
2. B: I went to beach.
3. A: Uhuh, when?
4. B: Last, *lasto*, last week. (Long pause)
5. A: Which beach?
6. B: (difficult to understand) by bike.
7. B: *Eto*, I went, went *McDonald*.
8. A: I see. Which beach did you go to? *Doko*?
9. B: Ahh! *Doko*? *Shikanoshima*.
10. A: Ah, it's nice there. Was it fun?
11. B: Fun!

These dialogues illustrate examples of the speaking proficiency of students before the intervention took place in this study. Student A starts with a fluent response to the greeting and reciprocates, thus immediately contributing to a conversational style. However, this reciprocation is a typical greeting taught to students in junior high school. When the conversation moves on from this generic format to more unstructured talk, fluency is inhibited. Rather than responding to information or reciprocating personal feelings, Student A simply nods, which may in natural conversation indicate an end to the interaction. The student in line 6 apparently misunderstands the question, as he begins with “I brush a tooth”. He could either have been drawing on a “daily activities” dialogue he may have studied previously, or he misunderstood the purpose of the activity. In line 6 he also uses the present rather than past verb tense to describe his activities, and there is a lack of conjunctions with which to achieve textual cohesion. In line 7 the teacher offers an “OK” in an attempt to encourage the student to contribute further; however, the student’s answer is a simple “yes”, either signifying that he misunderstood the “OK”, or perhaps signalling that he wanted the conversation to end by not sharing any further information. The teacher persists with the development of the text in lines 9, 11 and 13, in an attempt to draw more information from the student about his recent activities; but responses are primarily single-word answers, or in the case of line 14, in Japanese. The use of Japanese might signify that rather than the student wishing to end the conversation, he simply cannot continue, as he lacks the linguistic resources to reply. This is corroborated by the consistent use of “yes”, even when not

appropriate, such as in line 18, and relying on the vocabulary used by the teacher to answer questions in a verbatim manner. Overall, the initial enthusiasm of the opening exchange is not maintained beyond the generic greeting.

Student B shows a similar speaking proficiency as Student A. In line 1, the teacher uses “Hey” rather than the generic greeting; this may be unfamiliar, and is therefore not reciprocated by the student. Unlike Student A, however, Student B does orientate the listener to what and where and correctly, uses the past tense in line 2. However, the article “the” is missing from his answer, and he does not provide any further information voluntarily. The teacher attempts to develop the conversation by asking the student more orientating information: however, in line 6, the student confuses “which?” for “how?”, and later the teacher resorts to using Japanese to illustrate a “where?” question. In line 9, the enthusiastic emphasis in the student’s response suggests that rather than wishing the conversation to end, understanding is the main obstacle to a fluid conversation. Nevertheless, as with Student A, Student B relies heavily on answering questions using the same vocabulary that the teacher used, and providing one-word answers. These exchanges help to highlight current gaps in student knowledge and proficiency, which provide a useful starting point for exploring ways in which a change in teaching and learning could assist their speaking abilities.

3.6 Summary

This chapter has outlined the research perspectives that informed the research design for this study, in different levels of abstraction. The methodologies and methods that influence research were discussed, in order to reflect the choices that lay behind an action research approach to the present study. Limitations of an action research approach were highlighted, as well as certain contextual factors that limited the generalizability of this research. The next chapter will outline the action research process that aimed to address the research question

identified in Chapter 1, and will provide a description of the structure of the remaining chapters of the thesis.

Chapter 4: Cycle One: Planning and Action

4.1 Introduction

This chapter begins the two cycles of the action research process undertaken in this thesis (see section 3.4.3). As such, this chapter outlines the syllabus design, methodology and instruments employed for data collection, and the planned analysis of this data in Cycle One; these correspond to the “planning” and “action” phases of action research, as proposed by Kemmis and McTaggart (1998). An initial timeline of data collection and analysis is presented, together with a description of the participants in the data sample. A syllabus that adopts the major theoretical underpinnings of the genre-based approach explored in Chapter 2 is constructed and described in detail, with possible ethical issues during the data collection process taken into consideration. The procedures for the analysis of this data collection are then described. Subsequent chapters will then be further structured to reflect Kemmis and McTaggart’s “reflection” and “observation” phases of action research, and future cycles of action research.

Kemmis and McTaggart’s (1998) model is described by Burns (2010) as a “classic” (p. 8), and as a useful model for summarizing the essential phases of action research; it thus provides a convenient means of structuring the thesis. The four phases constitute one cycle of action research, and the research undertaken in this thesis will cover two full cycles of action research. Additionally, a third and final phase of “planning” was also conducted, and will be described in the final chapter of this thesis, as a means of developing a discussion about future directions and subsequent cycles of action research that could be undertaken beyond the scope of this study.

It is important to emphasize at this point that this chapter primarily illustrates initial assumptions and planning at the start of the data collection process, before iterations of the action research cycle took place. Some of these assumptions, as well as syllabus, assessment,

and consequently data collection plans, changed as the data collection and analysis progressed. Various changes and modifications were made throughout the action research process at different times and in different classroom settings, based on the inherent reflective and dynamic nature of such an action research approach. Reflecting these concerns, Schön (1983, p. 22) introduces the concept of “reflection-in-action” to illustrate the spontaneous ways of thinking and acting undertaken in the midst of the general cycles of action research, which are described by Kemmis and McTaggart (1998) and by Schön himself as “reflection-in-action”. Burns (2010, p. 14) describes this reflection as “on our feet”, which introduces changes to our plans in the midst of the ongoing research. To discuss these smaller cycles of planning and reflection within the larger cycles set out by Kemmis and McTaggart, the end of this chapter will contain a report on “reflection-in-action”, to highlight changes to the plan that were made spontaneously during the research.

4.2 Planning: The Research Question

Burns (2010) summarizes the “planning” stage of action research with the following checklist of criteria that should be fulfilled during this phase, which includes:

- Finding a focus area for your research
- Developing and refining your questions
- Referencing the literature
- Organizing equipment and materials
- Addressing ethical issues

Sections 4.2 to 4.4 will be structured with reference to these criteria as part of the narrative of the planning phase of Cycle One. The planning process of Cycle One began with an identification of a problem in my teaching context, which is described in detail in Chapter 1 and section 3.5.1: specifically, the perceived lack of speaking abilities of my students. The

primary goal of this research, therefore, was to explore how a change in syllabus design to follow a genre-based approach might assist the speaking abilities of Japanese students. This is exemplified as a question below:

In what ways can a genre-based approach assist the teaching and development of Japanese students' speaking abilities?

Such a research question would necessitate the construction of a deliberate intervention in the classroom, involving a genre-based approach. It was decided in the planning stage of Cycle One that the intervention would take the form of a semester-long syllabus. A potential issue of asking such a research question is raised by Ellis (1996), who claims that the EFL teacher could be disadvantaging students by focusing specifically on oral skills; particularly when the examination process is focused upon the grammar translation method. The genre-based approach is a methodology that encompasses elements of explicit grammar instruction, and for ethical reasons it was important to consider the concerns of Ellis (1996). The demands of the TOEIC Bridge test, as required by the institution in which the study takes place, also had to be taken into account when considering syllabus design.

4.3 Planning: Syllabus Design

The next step in the planning phase of Cycle One was the creation of a genre-based syllabus as a deliberate intervention within the current teaching context. Based upon the theoretical underpinnings of the genre-based approach outlined in Chapter 2, such a syllabus was constructed, and initial plans were made for the chronological implementation of the syllabus. The starting point of the planned intervention involved choices based on the types and number of genres that were capable of being taught within the time constraints of the participants' classes and the curriculum of the university. The issue of the TOEIC Bridge assessment dictated by the university also had to be taken into consideration, as per the

concerns highlighted by Ellis (1996) in section 4.2. Slade (1997) offers eight essential genres in spoken interaction that have recognizable generic structures, and which are detailed in section 2.9. Slade (1997) explains that participants in conversations weave in and out of different genres during conversation. Considering the proficiency levels of my students and the time constraints of the educational context, it was decided that only one genre would be studied in depth in the planned cycles of teaching and learning that constituted the intervention in this research. The justifications for this choice will be explored later in this section.

However, as it was essential to raise awareness of different genres, I decided that initial lesson plans would focus on highlighting a range of different genres, in a range of different contexts, and highlight the differences between them. Music is a useful tool for motivating students and gaining their attention; and in order to raise students' awareness of genre, at the start of the first class I asked them to identify different musical genres based on short audio clips. After identifying the musical genre, students were then asked to think about why and how they were able to identify each one. As part of this discussion, I was able to develop the idea that different genres have different key elements that identify them. In the case of music, this might include the type of instrument used, the type of lyrics, the rhythm, and so on. Appendix 1 shows an attempt to transfer this idea into the context of language learning. Students were asked to match short examples of written text with genre types, in order to illustrate how by analysing structure, vocabulary and grammar we are able to identify different genres of writing. It was explained that by deconstructing genres while focusing on key structural, lexical and grammatical characteristics, students can then also produce different genres independently. By saying to my students "Knock, Knock" and indicating the start of a joke, I was able to introduce the idea that there are also different identifiable genres in spoken English. I was also able to discuss with students the reason we might tell jokes and the different purposes for different types of writing or speaking: such as

sharing important information, exchanging goods and services, or establishing social bonds. By making it clear that jokes were a way of developing social bonds, for example, I was able to share with students the relevant cultural knowledge within which genres are based. From this basic introduction of a genre-based approach to students, I then began to introduce the semester-long syllabus, the content of which will now be discussed in greater detail.

Rothery (1996) describes key aspects of the Sydney School of genre pedagogy as part of a Teaching and Learning Cycle. The Teaching and Learning Cycle is based upon the modelling, deconstruction, joint construction and independent construction of texts. The genre-based approach in this cycle holds a view of language learning that is most successfully achieved through working with whole texts. These texts are selected in relation to a learner's needs and the social contexts which learners wish to access. According to Feez and Joyce (1998), the methodology which supports a text-based syllabus is based on a model of teaching and learning in which the learner gradually gains increasing control of text types. In this way, it is possible to develop sound principles for selecting and sequencing the content elements of the syllabus.

As previously discussed in this section, I had initial concerns about the number of text types I would be able to include within the contextual constraints of the teaching and research situation. I justified the choice of only one text type for the initial syllabus with the acknowledgement that future cycles of action research would introduce additional genres, and comparisons would be made in class based on the content of previous syllabi. Joyce and Burns (1999) explain that within the Sydney School of genre pedagogy, teachers start planning a syllabus by focusing directly on the spoken texts which students need to learn and use. Butt, Fahey, Spinks and Yallop (2000) develop this idea further by stating that the text type, and associated metalanguage necessary to deconstruct it, will depend on the students' developmental stage and existing knowledge, as well as the requirements of the teaching

context. I began my own syllabus design by first developing a pair of broad goals or aims that matched my students' needs and abilities. These initial aims included preparing students for the TOEIC Bridge test and developing speaking skills for casual conversation outside the classroom. With Slade's (1997) eight essential genres in mind, I focused primarily on the grammatical components of each spoken genre. The grammatical features of the recount genre commonly appeared in the TOEIC Bridge test, with a great number of questions in both the listening and reading portions of the test addressing verb forms and conjunctions. The recount is also a genre that is familiar to students; this might mitigate the possible confusion associated with introducing a new style of teaching and learning, by allowing students to reflect on past experiences in their language study. The recount is also a genre immediately relevant to students as a common form of casual conversation, and would rely on students drawing on their own experiences. For these reasons I decided to sequence the content of my syllabus based on the recount text, with the caveat that in initial classes I would raise awareness of differences between different genres, and that in the future, different genres would also be explored.

With the recount genre selected as the primary focus of the syllabus, I referred to the model of the teaching and learning cycle (Callaghan, Knapp and Noble, 2014) in section 2.9 and Figure 2.5, in order to sequence the content of my syllabus into units of work and individual lesson objectives. This section and the examples in the appendices attempt to illustrate the format of the semester-long intervention as it was originally planned. Many aspects of the syllabus described in this section were adapted and changed significantly at different times during the course of Cycle One, and these changes are discussed in further detail in section 4.8. At times, these changes would occur from one weekly class to the next, or from one semester to another. A tentative plan for data collection was also constructed before the intervention discussed in this section took place. The process of data collection constitutes the next section of this chapter. The first stage of the model of teaching and

learning (Callaghan, Knapp and Noble, 2014) requires students to interact collaboratively with the teacher, to build up a shared experience of the context of the text they are going to use. This was first approached through the activities described earlier in this section and exemplified in Appendix 1. With specific reference to a recount genre, I discussed with students the reason we ask and tell each other about what we did last weekend, or during our vacations, thereby establishing knowledge of the social activity and the subject matter, and developing the notion of Field discussed in section 2.6.1. We were also able to discuss what kinds of people initiate such recount texts and the types of relationships they might have, thus invoking the notion of Tenor described in section 2.6.2. Differences between speaking and writing were also highlighted, using Eggins' (2004) discussion of the differences between spoken and written language detailed in Table 2.7. These initial discussions then led to the modelling stage of the syllabus design. In the planning stage, an idealized written form of the recount genre identified in Chapter 2 was chosen to share with students, in order to exemplify the key structural and lexico-grammatical components of a typical recount genre.

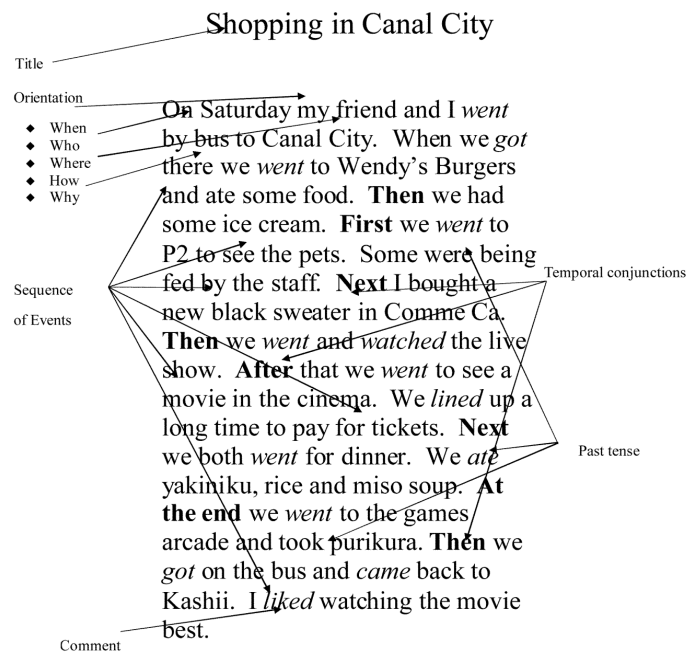


Figure 4.1 Generic structure of the recount genre (Adapted from Butt et al., 2000)

Figure 4.1 shows the recount genre as it was first presented to students during Cycle One of the research process. The idealized model is based upon the generic structure of a recount text as presented by Butt et al. (2000), described in section 2.5 and Table 2.3. The purpose of a recount genre is to retell past incidents in the order in which they occurred, and as such the structural features include an orientation detailing the who, what, when and what of the event to orientate the listener, followed by major events in chronological order. Key grammatical features include past tense verbs and temporal conjunctions. These grammatical features were consistent with grammar items in the TOEIC Bridge test. The structure and vocabulary were chosen to match the proficiency levels of the students, whilst still intended to be somewhat challenging. Table 4.1 shows how the generic structure and lexical-grammatical elements identified in the model recount text were collated into criteria for informing syllabus design and assessment. Assessment criteria are discussed in greater detail in section 4.7.1 and Table 4.3.

Table 4.1 Deconstruction of a recount text to inform syllabus design

Structural Criteria	Lexical- grammatical Criteria	Turn-taking Criteria
Orientation	Greeting	Asking questions
Sequence of events	Past tense verbs	Supplying helpful information
	Temporal conjunctions	Expressing surprise or support

Table 4.1 shows how additional criteria were generated based on the turn-taking nature of a casual conversation. During the planning phase of Cycle One, it was envisioned that these criteria would inform a weekly sequencing of classroom content based upon the teaching and learning cycle (Callaghan, Knapp and Noble, 2014). Appendix 2 shows a copy of the syllabus with each stage of the teaching cycle highlighted. After deconstructing the model text in Figure 4.1, joint construction of the text included a step-by-step framework of activities that allowed students to then create their own recount text via explicit instruction and feedback from the teacher. Independent construction of the text involved students constructing a new recount text during a conversation between teacher and student. This conversation was assessed by the teacher and by the students themselves. The assessment instrument for data collection and the rubric in which the criteria were described to students can be seen in Appendix 19 and 20, and it is referenced in more detail in section 6.5.2.

Appendices 3 to 7 contain scans of a portfolio of work completed by a student during the course of Cycle One of the intervention. Appendix 3 shows the “Study Plan” distributed to students during the initial classes of the syllabus intervention, following the context-building activities mentioned previously in this section. In the study plan, each of the key elements of the recount text were sequenced into lesson content, beginning with greetings and ending with turn-taking strategies. On the study plan, students had space to keep reflective notes on how they felt the class went and their understanding of the class content.

Students were asked to write their reflections in Japanese and English, and this formed part of a weekly student journal that is discussed in greater detail in section 4.6.3. Appendix 4 illustrates how tenor was discussed with students through the practice of different types of greeting. The content of the worksheets matched the speaking proficiency levels of the students, as exemplified in section 3.5.1. Appendix 5 shows a series of worksheets building on the criterion of “Orientation”, first by asking simple comprehension questions based on the model text, and then later by developing their own orientations based upon their own experiences. Appendix 6 shows an example of worksheets addressing the grammatical elements of a recount text. On the worksheet in Appendix 6, students are asked to identify the temporal conjunctions in a model text, and then think about how they could use these same conjunctions in their own recount texts during joint and independent construction. Appendix 7 exemplifies worksheets that attempted to develop turn-taking strategies that would help students ask and answer questions during a casual conversation. The final classes of the semester were organized in order to conduct a final conversation between student and teacher, during which students would independently construct their own recount genre based on their own recent experiences, and structured with reference to the content of the syllabus. These worksheets were collected as part of the data collection process that will be described from section 4.5 onwards. Before the data collection methodology is discussed, section 4.4 will discuss ethical considerations in regard to the data collection that took place.

4.4 Ethical Considerations

Before each phase of data collection, and according to the Aston University “Policy on Research Ethics”, all research involving human participants was subject to ethical review. This study involved observations, recording of speech, and the use of classroom documents such as worksheets and reflective journals; all of which required ethical approval.

The PhD Student Research Ethics Approval Form (REC1) was submitted twice. The first submission concerned the collection of student data described in Chapter 4, and the second dealt with the collection of additional data described in Chapter 5. This form included information that outlined my research aims; the methods I intended to use; and details of the informants, the method of access and sampling, and location.

I also detailed the ethical issues I considered would arise from the research, and the steps I took to address these issues. This included asking for consent from all participants, ensuring they understood exactly what they were expected to do, and that their data might be shared with PhD examiners or my supervisor. A consent form that explained these considerations was distributed in English and Japanese, and gave participants an option to withdraw from the process at any time, as well as to withhold any data collected. Participants were also told that they would remain anonymous, and that all data collected would be held either on password-protected computers or locked in a cupboard in my office. Teacher participants described in Chapter 5, section 6.3.4 were offered an opportunity to read any of my completed written reports, and the option to see the completed thesis before formal submission. Due to the intrusive nature of video recording during their final speaking interview with the teacher, students were given an option on the consent form to not be video-recorded. A number of students asked not to be video-recorded, and they were assured that this would have no bearing whatsoever on any form of assessment or classroom procedures that consenting students also experienced. This was also true of any form of data that students produced; but all students kindly consented to analysis and anonymous sharing of other data types. Section 4.5 onwards will describe the data collection procedures utilized during Cycle One of the action research process.

4.5 Action: Data Collection

Burns (2010, p. 54) summarizes the next phase of the action research cycle as “putting the plan into action” according to the following criteria, which offer a useful guideline for structuring Sections 4.5 to 4.10 of this thesis:

- Identifying the main ways used to collect data for action research
- Discussing the combination of classroom activities and data collection
- Describing the selection of appropriate methods for research questions
- Consideration of a range of observation methods
- Incorporating technology into data collection
- Triangulation

This section describes the process of data collection with these criteria in mind. As part of the action research cycle, data were collected and analysed as part of an ongoing process, with data being reflected on in combination (Burns, 2010) with the action. Table 4.2 provides a summary of the main data collection instruments utilized in Cycle One. Quantitative and qualitative data were generated by both teacher and students.

Table 4.2 Data collection instruments.

Quantitative data collection points	Qualitative data collection points		
		Teacher-generated	Student-generated
	During class and immediately after class	Feedback written on students' weekly class work Observational notes written on lesson plans	Student reflective journal Classroom documents Spoken class comments (noted on lesson plans by teacher)
	End of the week	Reflective journal	
Final teacher assessment	End of the semester	Written feedback to students on portfolio work	University-wide questionnaire Anonymous written feedback

Table 4.2 shows how classroom activities were combined with data collection throughout the semester, with quantitative data collected at the end of the semester via a final spoken assessment involving a student and the teacher. The sub-sections of 4.5 will now discuss these data collection procedures in more detail.

4.6 Qualitative Data Collection Points

Quantitative and qualitative data were collected at various points during the action phase.

These are summarized in Table 4.2, but will now be described in more detail in sections 4.6.1 to 4.6.5.

4.6.1 Classroom documents

Classroom documents included the following: syllabus guidelines; lesson plans; worksheets; student journals; recorded examples of students' spoken recount texts at various stages, including videos; and assessment events. Worksheets were collected weekly, together with student journals, to form a portfolio. Written feedback was provided on worksheets, as well as spoken feedback given during class. Where possible, spoken feedback was noted by the teacher on the lesson plans during class time. At the end of the semester, student portfolios were considered without any formal assessment awarded, but with written and spoken feedback given to students.

4.6.2 Lesson plans

For each of the fifteen weeks of the research, a lesson plan was constructed, detailing the following: class materials; the aims and objectives of the particular lesson; types of assessment planned; notes on classroom events as they happened; spoken feedback given to students; interesting events during class; and reflection on the class. Appendix 8 shows part of an example lesson plan, with notes written by the teacher during the class, based on classroom observation. The "Aims" of the lesson plan described the overall aims of the class, while the "Objectives" showed planned activities for that class. Space was provided on the lesson plan to record observations during class time. Space on the lesson plan was also provided for students to write comments on immediate reflection after class had finished. Student portfolios were collected at the end of the lesson and their reflective comments were summarized on the lesson plan sheet. This allowed a cross-referencing of teacher and student observations and reflections on that particular class. A space was provided for notes to be made concerning plans for the next week's class, and emerging themes in the data were also noted.

4.6.3 Teacher and student journals

Notes made on the lesson plans were factual, descriptive and spontaneous. The weekly reflective journal aimed to collate and summarize all of these types of data in a second period of more considered reflection. The research journals elicited written content by referencing a set of questions adapted from Richards and Lockhart (1996), shown in Appendix 9. The teacher and student were not expected to answer all questions in the list, but to use them as a framework for producing reflective content on the syllabus that was appropriate. Teaching questions attempted to summarize what principles and beliefs were informing the teaching; the roles of the learners in the classroom; and any improvements that could be made. The research generated data that were a reflection of the research process as well as the pedagogical considerations. The journal also provided space to identify broad patterns and trends in the data. An example of a teacher journal entry can be seen in Appendix 10, and a student journal in Appendices 3 and 11. The teacher journal shows a written reflective account of the previous lesson, followed by tables for entering problems that were observed in class and potential resolutions, either via reflection-in-action during the ongoing syllabus intervention, or as reflection-on-action in preparation for Cycle Two of the research cycles.

4.6.4 University-wide questionnaires

The university also has a compulsory formal questionnaire that is distributed to all students at the end of each semester. The teacher receives quantitative feedback shortly afterwards. Included are anonymous and unstructured comments from students, which typically reflect on the class. These data were also utilized where appropriate.

4.6.5 Anonymous feedback

At the end of the syllabus and after the final assessment procedure described in section 4.7.1, students were given a sheet of paper similar to that given at the start of the syllabus, for their

weekly reflections (shown in Appendix 3). Students were thus able to provide anonymous feedback regarding their overall thoughts on the syllabus and their own performance.

4.7 Quantitative Data Collection Points

As well as the qualitative data collection during the action phase of Cycle One, quantitative data were also collected using an assessment instrument at the end of the syllabus; both as a teacher summative assessment, and by the students as a form of self-assessment. The final classes of the syllabus involved a conversation between the teacher and a student, during which students would independently construct their own recount texts based on their own recent experiences.

4.7.1 The assessment instrument

Using the criteria illustrated in Table 4.1, an assessment instrument was created (see Appendix 12). During the action research cycles, this assessment instrument changed many times. In Cycle One, 26 students were asked to create their own spoken recount genre with the teacher, as a form of final assessment. During Cycle One, each criterion in the genre was graded on a 3-point Likert scale from 0 to 2, with 0 determining that the criteria had not been used or had not been used successfully, and 2 indicating that criteria had been used effectively and fully. Six initial criteria were identified in Cycle One to inform the assessment, and are described in Table 4.3:

Table 4.3 Description of criteria for assessment

Criteria	Description
Generic greeting	Students take part in a turn-taking greeting, reciprocating and answering generic greeting phrases. This is not necessarily a vital component of a recount text; however, it serves as a useful starting point for initiating a recount text with my students.
Orientation	Students develop key “WH” information at the start of their talk, to orientate the listener: e.g. “Last Saturday I went to the cinema”
Sequence of events	Students describe a number of events that occurred during a particular event, ordered in a logical time-ordered manner.
Past tense verbs	Students use past tense verbs appropriate to a recount: e.g. had, ate, went.
Temporal conjunctions	Students connect ideas and events through use of temporal conjunctions: e.g. next, then, after that.
Answering questions	The speaker would answer questions from the listener, as appropriate to the overall structure of the text.

Table 4.3 shows how key features of a recount genre were defined as criteria for assessment, including grammatical features and structural features. With the inclusion of self-assessment procedures in Cycle Two, assessment data also allowed a comparison of student self-assessment and teacher assessment that could highlight differences in student and teacher perceptions of their spoken performance. These observations are only possible with appropriate analytical methods. Section 4.9.1 will discuss these analytical methods further, after section 4.8 provides a short report on the “reflection-in-action” described in section 4.1.

4.8 Reflection-in-Action

Section 4.1 described Schön’s (1983) notion of reflection-in-action, and that during the action research process there are smaller iterations of the cycles of reflection, whereby change occurs spontaneously in response to students’ needs or events in the classroom. This section highlights two major changes to the syllabus as a response to reflection-in-action during the intervention.

4.8.1 The model text

During the action phase of the research, I noticed fundamental problems with the model recount text as presented in Figure 4.1. Specifically, based on this model, students were not preparing themselves for creating a spoken recount text, but for writing a recount text, which they would then memorize and produce as a monologue. During students' practice with their peers, I noted numerous times in my lesson plans and reflective journal that they relied heavily upon written models for their speaking. Students felt far more comfortable with written output as a means of producing spoken output. Another symptom of this was that during practice with their peers, there was no turn-taking. Although students were able to understand the basic concepts of the genre-based approach and could see value in its application, rather than having casual conversations and fulfilling the aims and objectives of the syllabus, the students were memorizing monologues and repeating them to each other. At first, I created a new model text that looked more like a transcript than the written monologue presented in Figure 4.1 (See Appendix 13). This new model helped to develop the notion that turn-taking was an important part of conversation. However, turn-taking was also proving to be the most difficult aspect of casual conversation for my students to construct independently; and while writing transcribed monologues, they also planned possible responses to any questions that might be asked in writing. I introduced some audio and video models to the syllabus in reaction to these events, but with varying degrees of success. My students' reliance on written language to inform their speaking was not something that could be eliminated easily, and this problem is discussed further in Cycle Two.

4.8.2 Turn-taking

Due to the reliance on written output highlighted in section 4.8.1, and the resultant lack of turn-taking during student conversation, I felt it necessary to include additional criteria for classroom teaching and assessment that would reflect the necessity for turn-taking in student conversations. In section 2.8, a discussion of the differences between spoken and

written language was highlighted. Eggins (2004) explains that written language is more compressed than spoken language and frequently shows signs of editing. When students rely on written language to inform their spoken language, the synoptic nature of the talk can sound unnatural and abnormally polished. This unnatural talk was evident in my students' practice conversations in class, as they memorized pre-written scripts for their conversations. Examples of this can be seen in section 5.5 and Appendix 23. In Table 4.3, my initial criteria had failed to fully address Halliday and Hasan's (1985) "grammar of interaction", which describes that when we use language to interact, the interlocutors are establishing a relationship, through a turn-taking sequence whereby they take on different speech roles. These "speech roles" can be seen in Table 2.8, where they are outlined as "giving" and "demanding". Concurrently, a speaker also chooses the type of "commodity" they are exchanging. This commodity can consist of an exchange of information or an exchange of goods and services. Eggins and Slade (1997) explains that while native speakers are already aware of the structures of spoken interactions and successfully fulfil a turn-taking role, in teaching-second language learners, it is essential that this "native" knowledge is made an integral part of syllabus design. As such, I consulted the Common European Framework of Reference for Languages: Learning, teaching assessment (Council of Europe, 2001, 2008), and developed additional criteria based on these level descriptors, as appropriate to the students' ability. These can be seen in Table 4.4:

Table 4.4 Description of additional criteria for assessment based on the Common European Framework

Criteria	Description
Pronunciation	The student uses English sound system with intonation reflecting thoughts and emotions.
Fluency	Speaks spontaneously and at length with a smooth flow of language, without noticeably long and unnatural pauses.
Illustrations	Conveys more detailed and finer shades of meaning to differentiate and eliminate ambiguity.
Repeats	Repeats and recasts information when prompted by the teacher.

In order to develop further Eggins and Slade's (1997) idea of native knowledge, I created new model texts to be introduced to students, based on audio recordings I made of native speakers holding authentic recount conversations. Students listened to the audio and answered comprehension questions based on the content; we later reflected on pronunciation and fluency, and how detailed knowledge was shared between the two speakers, as well as how information was repeated and recast during the conversations. However, these modifications, which were not as thoroughly pre-planned as other elements of the syllabus, raised further complications and difficulties, which are explored in greater detail during the observation and reflection phases of Cycle One, described in Chapter 5.

4.9 Preparing for Data Analysis

This section will detail the planned analysis of the data collected by the tools discussed in sections 4.6 and 4.7. Data analysis followed both qualitative and quantitative processes, in order to triangulate findings and increase generalizability. In this section, procedures for qualitative and quantitative analysis will be explained in detail.

4.9.1 Qualitative data analysis

Data analysis adopted the process of inductive coding, through an emic approach to data analysis that is consistent with the ontological and epistemological underpinnings of the

critical theory perspective and the action research approach described in Chapter 3. Inductive coding is a systematic procedure for analysing qualitative data using detailed readings of raw data to derive concepts, themes, or a model through those raw data (Thomas, 2006). This fits with the notions of Strauss and Corbin (1998), who describe this process as beginning with an area of study and allowing the theory to emerge from the data. Furthermore, Scriven (1991) describes the inductive process as “goal-free” (p. 56), whereby the researcher describes actual effects of an intervention and not just planned effects. Burns (2010) describes the process as an “insider” (p. 107) approach, because we look at the data from the perspective of someone inside the research, with the data providing the categories. This is in opposition to a deductive approach, where categories are pre-determined and gathered from the theory and literature. Thomas (2006) offers a slightly different account of the separation between inductive and deductive coding by describing a “general inductive approach” (p. 3) where analysis takes place through multiple readings and raw data, but where research objectives also determine the type of analysis that takes place. Cook and Crang (2007) describe the inductive process as the most common approach for identifying themes in qualitative data, which Charmaz (2006) describes as “grounded theory” (p. 40). “Constant comparison”, which lies at the heart of this method, refers to the breaking down of data into discrete units and coding them to categories (Glaser and Strauss, 1967; Lincoln and Guba, 1985). The categories created in the inductive process are derived from the participants’ own language, and are also those seen as significant to the research questions. The process allows the researcher to conceptualize the participants’ experiences and develop theoretical insights into the social process that is occurring (Lincoln and Guba, 1985). It is important, however, to view this approach to qualitative analysis critically. For instance, the insider may perceive themes and categories that match pre-conceived concepts. As such, at this point it is important that the discussion of my own research history and teaching context in sections 1.2, 1.3 and 3.4.1 is taken into account, to provide a sense of the types of preconceived ideas I

might possess based upon my own experience and the research context. Figure 4.2 is my own attempt to summarize the stages of analysis conducted in this study:

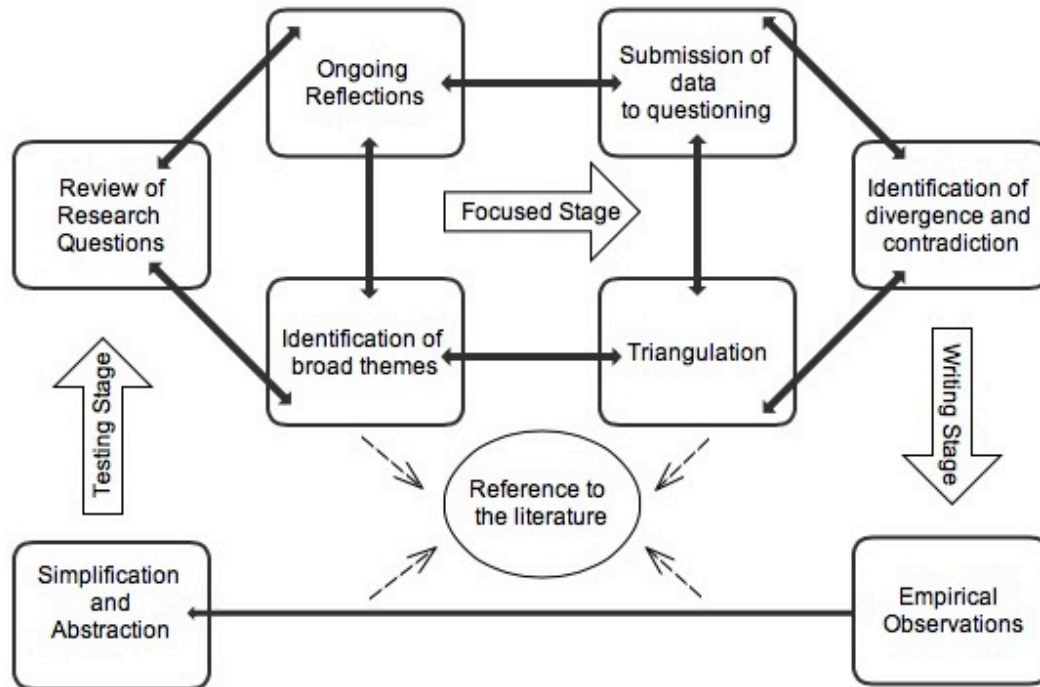


Figure 4.2 Summary of qualitative data analysis in the study, based on suggestions by Bazeley (2009) and Burns (2010)

Data analysis in this study, as represented by Figure 4.2, involved a number of strategies to explore deeper meanings in the data, as suggested by Bazeley (2009). These included: questioning the data to improve interpretation and categorization; triangulation to compare and find patterns in the data; using divergent or contradictory themes in the data to challenge generalizations; use of the literature at various stages as a source of explanation; modelling of theory; and writing to prompt deeper thinking. This section explains the ideas underpinning the analysis adopted in this study, as well as the main categories and concepts that arose from the data, and which are displayed in subsequent chapters.

Data analysis began “in combination” (Burns, 2010, p. 104) with data collection. As soon as I received data, I read them, reflected on them and made notes in my weekly journal (see Appendices 8 and 10). This included comparing data from my own journal with student journals (Appendix 11). Schiellerup (2007) echoes the notion of ongoing reflection on the data by stating that analysis is an interpretation of experiences that occur during the collection process, and is not a process to be conducted merely during dedicated moments of focused data interpretation.

Burns (2010) suggests that there is a “stopping point” (p. 104) of this ongoing process, when a more focused assembling of data and its analysis will occur, and the broad picture that has developed is refined by coding data into more specific patterns and categories. In Figure 4.2 this is emphasized by the “focused stage” arrow. This stage began with multiple repeated readings of the entire dataset, to make myself as familiar as possible with what had been collected. Ellis and Barkhuizen (2005) argue that in order to arrive at concepts in the data, researchers need to read and re-read texts very carefully, and that this reading process is continuous. Bazeley (2009) suggests that this stage transfers the focus from themes to “categories” and more abstract “concepts” (p. 2). Bazeley argues that identifying themes acts as a good starting point for qualitative research, but effective reporting requires using the general ideas generated from the data to build an argument that moves beyond descriptive reporting. He describes themes as “little more” (p. 3) than organizing areas discussed by participants, and without considerable explanation they do not communicate with the reader or construct meaningful abstractions of the research.

In order for analysis to move beyond description, which Bazeley argues is “not sufficient” (p. 4), data must be strengthened by ensuring they are “challenged, extended, supported and linked”. “Triangulation” (Burns, 2010, p. 95) is one way of linking and cross-checking data in order to strengthen findings by providing objectivity. Burns outlines five

forms of triangulation, comprising methods triangulation, time triangulation, space triangulation, researcher triangulation and theory triangulation. Each of these five types of triangulation was employed in this study. Data were drawn from a range of sources, such as journals; classroom documents; data collected throughout the course of the genre-based syllabus; numerous classes at different times of the week; feedback classes that were taught by other teachers; and the use of qualitative as well as quantitative data collection and analysis methods.

The idea of comparison and triangulation of datasets is further supported by Bazeley (2009), who suggests a three-step formula to work through for developing a coherent model of data analysis: “Describe; Compare; Relate” (p. 5). After we have described themes as a starting point in analysis, by explaining how participants related these themes, how many people talked about these themes, and what was not included, Bazeley explains we should next compare data in these themes across different sources, to record meaningful associations. Burns (2010), Miles and Huberman (1994) and Bazeley (2009) also suggest identifying contradictions in the data, in order to challenge generalizations and provide sources for further analytical thinking; this provides hints of what is happening for the larger sample, and thus refines categories and concepts. The process of challenging data begins to define the conditions under which certain categories arise; the interactions, actions and strategies involved; the consequences of certain interactions, actions and strategies; and how these vary depending upon circumstance.

Through cross-referencing, questioning and challenging data, themes are thus developed into categories and concepts; this serves as a means of increasing levels of abstraction in the interpretation of the data, in what Miles and Huberman (1994) refer to as a “ladder of abstraction”.

Schiellerup (2007) describes this process as the aggregation of codes into “super codes”, “code families” and “networks”. Super codes are formed through the combination of ordinary codes that have been generated previously; code families are based on certain codes sharing a similarity; and networks refer to networks of codes where relationships between different codes can be specified. Schiellerup was writing with particular reference to the computer software he used to assist his coding process. This study, however, did not employ any software for the analysis of qualitative data.

Welsh (2002) compared manual techniques for qualitative data analysis, with the use of Nvivo. She concluded that although the software yielded more reliable results for gaining an overall impression of the data due to the reduction in human error, nevertheless it was not as useful in terms of interrogating the text in more detail. The main reason for this was the existence of multiple synonyms, which led to only a partial retrieval of information. She also highlighted problems of the software’s usefulness in relation to the way in which thematic ideas emerge. She claims that using the software makes it more difficult to understand how different themes form a whole, due to the ease with which searching takes place. She concludes that researchers should be open to recognizing the value of both electronic and manual methods, and not rely too heavily on one over the other. Stroh (2000) and John and Johnson (2000) also highlight the advantages of software by explaining that it saves time and increases flexibility. However, echoing Welsh, they conclude that it places a focus on volume and breadth rather than depth and meaning, and thus distracts from the work of real analysis (John and Johnson, 2004). After experimenting with Nvivo for a short time, I found it to be very useful as a starting point for thinking about how to store and code my data; but I felt I could make more sense of the data and create a deeper analysis by using a manual approach. Eventually, I migrated most of my data into a manual analysis, as I found it more intuitive. However, due to the large volume of student journal data (n=260), I continued to use Nvivo simply as a database. I was concerned with the notion that depth, analysis and meaning might

be negatively influenced by the use of software, given the adoption of action research, and the call by Denscombe (1998) for rigour in this form of research. Software was utilized in the quantitative research, namely Facets (Linacre, 2007a) and Winsteps (Linacre, 2007b). This software and the reason for its choice will be discussed in section 4.7.2.

At each stage of the above processes discussed so far in this section, as part of the action research cycle, continual reference was made to the theoretical and methodological literature. Bazeley (2009) argues that the experiences of other researchers in the same field can lead to inspiration and motivation, and also provide ideas for categorization. Reading the methodological literature also provides additional ideas for refining and extending analysis. As such, while writing about my findings in subsequent chapters, I also make regular reference to the literature included in my original literature review, or any gaps that I found in that literature. The process of reference to other themes in the literature thus assists a critical examination of what the data say about the research questions, by allowing a linking of discoveries in the study to a larger theoretical framework.

Figure 4.2 illustrates entry into the writing process: the subsequent chapters of this thesis. Bazeley (2009) states that writing up research is itself a tool for analysis, as concepts are dissected and ideas explored, with a summary of arguments to support conclusions. Bazeley suggests a series of writing strategies that lead to superficial focus and repetitive organization. These include relying on quotes as evidence; organizing chapters by source; organization of chapters according to voice; and organization of chapters according to method. In response to this, Bazeley suggests organizing empirical chapters by theme or issue, so that they can then be compared, contrasted and developed. In this thesis, therefore, chapters will be organized according to the major concepts that were drawn from the data analysis. As previously indicated, this study also utilizes quantitative data analysis where appropriate. This analysis is detailed in section 4.9.2.

4.9.2 Quantitative data analysis

Quantitative data analysis is carried out through the software Facets (Linacre, 2007a) and Winsteps (Linacre, 2007b). This software is primarily constructed to perform Rasch analysis, but it also displays output data in graphs and charts, using mean scores for comparison and correlations, in Microsoft Excel. I have often used this form of measurement to assist the assessment of large numbers of students in my classes; I have found it particularly useful as a diagnostic tool for determining aspects of my syllabus that are not understood by students, or for identifying students who may require additional assistance. Rasch as a diagnostic tool is emphasized by Engelhard (2009), who used the analysis to take account of students with disabilities who are compelled to take state-wide assessment programmes in the US. Engelhard argues for a mixed-methods approach to analysing data, which stresses the importance of teacher judgements and qualitative analyses in the interpretation of quantitative data. In the case of this study, a mixed-methods approach allows a Rasch analysis of quantitative teacher and student assessments to be used as a diagnostic tool to highlight issues with the syllabus and with individual students. This can then lead to cross-referencing, or avenues for further qualitative investigation to be conducted (Bonk and Ockey, 2003).

This section explains the Rasch model to be used for the analysis of data generated by the assessment instruments. It is argued that this model can be used within an action research methodology, to analyse elements of the syllabus design, assessment procedures, and also individual students.

The Rasch model is a form of Item Response Analysis, and is particularly useful for examining issues in assessment and classroom planning, such as levels of easiness and difficulty of criteria in a test or intervention (Linacre 2007a). It also assists in pinpointing assessment criteria or students who are not performing as anticipated; or criteria that have been simply misunderstood by students in class and require better explanation. Rather than

relying on raw scores, therefore, it is possible to look at specific components of spoken text production and how students are reacting to them. The Rasch model has the potential to provide valuable information that directly addresses the research question by exploring multiple ways in which the criteria generated by the genre-based approach assist language learning.

Output graphs and tables also allow the comparison of student and teacher ratings, and can identify which parts of the syllabus the students found too difficult. This qualitative ordering of criteria enables the teacher to determine which criteria the students are struggling with, and which are appropriate for the participants. It also allows the assessment and comparison of students' speaking ability, in relation to their ability to address certain criteria according to difficulty in a hierarchical fashion. The Rasch model has become increasingly popular in the field of second-language testing (Beglar, 2010; Laufer et al., 2004). Engelhard (2009) argues that the Rasch measurement is fundamental to the trait of language performance, which is probabilistic in nature, as opposed to deterministic. The Rasch model is based on two basic aspects:

- Higher-ability learners have a higher probability than lower-ability learners to answer correctly
- Easier items will have a higher probability of being correct than more difficult items

Bond and Fox (2007) offer an accessible description of the Rasch model, by using the analogy of jumping. They propose that jumping is a trait, and that an individual with more of this trait can jump higher. If raters ever needed to measure this trait, they could ask the jumper to jump and measure the height, thus giving a measure of their ability. This ability could easily be labelled in metres and centimetres. It is then possible to compare the jumping abilities of different people. For example, somebody with the ability to jump four metres would be able to jump twice as high as someone with the ability to jump two metres. The

ability of the jumper and the difficulty of the jump can be compared because of the measurement in metres. Each extra metre of jumping ability correlates to an equal level of jumping ability.

Unfortunately, measuring students' speaking ability is not as simple as using metres for jumping. In the genre-based intervention designed for this study there are six criteria, in a 5-point Likert scale that gives a possible top score of 21 points during assessment. However, if Student A scored 20 points, we do not know how much more ability he/she has than Student B who has scored 10 points, unless the scores have been calibrated with the criteria on a scale of difficulty with intervals much like metres. The Rasch model therefore allows the user to construct a measuring-tape of speaking ability, using software such as Facets (Linacre, 2007a), which can convert test scores into measures of criteria difficulty or student language ability.

McNamara (1996) explains that the Rasch is a probabilistic model, so that measures of success or failure are not deterministic. A probabilistic model will produce results that vary, rather than a single result. This variety of responses is recorded by the Rasch model in the form of "fit statistics". The importance of these statistics is highlighted in the works of linguists such as Ellis (1985), who argues that language acquisition is varied in individuals and depends on individual learner factors: for example, age, motivation and personality, and contextual factors such as doing a test or using language in a classroom, as opposed to over the phone to a water company. Certain sections of an assessment or syllabus may prove uniquely difficult for different types of students. Thus, item response theory, such as Rasch analysis, allows identification of individual students or individual criteria that are causing problems, which can be addressed on either a whole-class or individual basis. For these reasons, the Rasch model is adopted as a key tool for data analysis when conducting this study.

Another benefit of the Rasch model is that it also allows analysis, through the use of the Facets software (Linacre, 2007a), of how different raters perform on the assessment instrument: specifically, which raters are too harsh or too lenient, or which raters are not behaving as expected, meaning that they do not understand how to grade the criteria. This provides a useful analysis of student self-assessment, and it could determine, with further qualitative investigation, whether elements of the syllabus have been appropriately understood by the students. This feeds very well into an action research process, whereby key problems can be quantitatively identified and qualitatively reflected upon.

Whilst it is difficult to describe models of statistical analysis without reference to actual data, it is hoped that this section provides some justification for use of the Rasch model, and its role as a diagnostic tool to provide valuable avenues for cross-referencing with qualitative data. The next section summarizes how the results of these different data analysis procedures will be reported in subsequent chapters of this thesis.

4.9.2.1 The inclusion of Rasch analysis in action research

Section 4.9.2 describes how Rasch analysis can be implemented for pedagogical purposes in a number of ways, as follows: for identifying students who might benefit from remedial intervention (Engelhard, 2009); highlighting issues in syllabus content by identifying items that are too difficult or too easy for students (Linacre, 2007a); highlighting students who are finding the syllabus too easy or too difficult (Linacre, 2007a); highlighting elements of the syllabus that are being misunderstood by students (Bonk and Ockey, 2003); and highlighting raters performing self and peer assessment, both teachers or students, who are too severe or lenient in their assessments, do not understand the criteria they are assessing, or are inconsistent in their assessment (Lunz and Stahl, 1993). If we think of action research as “problematizing” our teaching context, then Rasch analysis helps us to identify problems and

seek potential solutions by applying qualitative interpretation to quantitative data (Cheung, 1991).

Rasch analysis uses the Rasch equation to convert raw scores in tests into logits (Linacre, 2007a). According to Holster (2015), traditional statistical analysis, exemplified by Brown's (2005) *Testing in Language Programs*, provides statistics such as item facility values (IF) and item discrimination (ID). These will identify most of the same problematic areas in test responses as Rasch analysis, and it is unlikely that classroom grades would change to any substantive degree between the two, in the case of a thoughtfully developed test. Rasch analysis, however, provides benefits beyond traditional item analysis. Holster (2015) argues that two of these important practical benefits are the variable map, which provides a quick visual summary comparing students with features of the syllabus; and data-model fit statistics, which enables the diagnosis and identification of students requiring remedial instruction, as pioneered by Engelhard (2009). Holster (2008) uses the variable map and fit statistics to provide a practical guide to inform teachers about vocabulary that is likely to cause difficulty for students of different levels, and to highlight the need for remedial instruction for high-proficiency students with mis-fitting responses. Holster (2015) also argues that Rasch analysis has benefits for language programmes beyond the identification of misbehaving items, thus providing insights into individual students' behaviour that are simple enough for non-specialists to interpret.

Figure 4.3 illustrates a variable map associated with the data collected in this study; it has been replicated here to illustrate its benefits, as posited by Holster (2015).

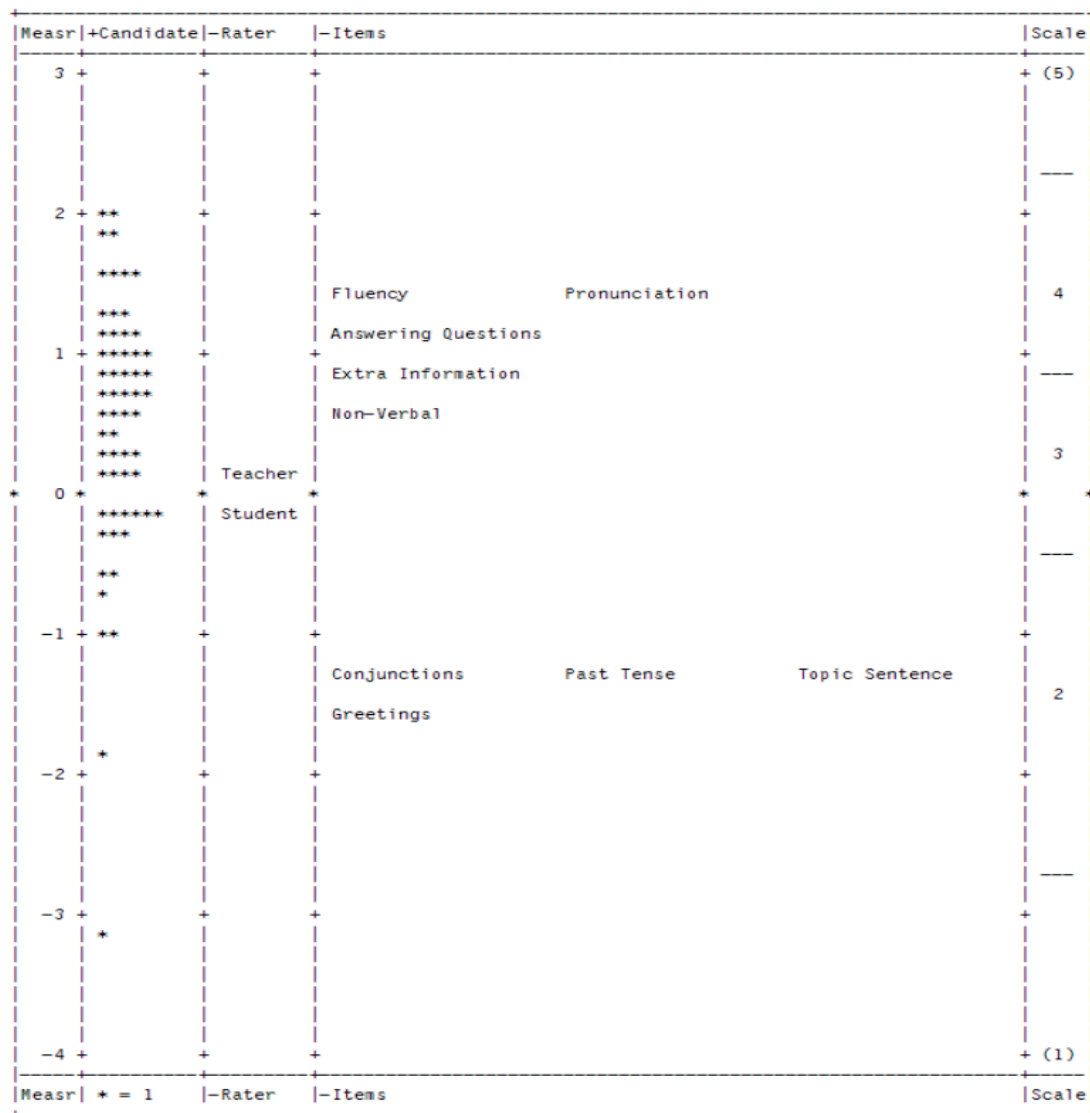


Figure 4.3 Person / Item / Rater map showing persons on a common scale with items and raters

By converting raw scores into logits, the Rasch model provides a psychometric measure of student performance (“Candidate”), illustrated by asterisks, and item difficulty (“Item”), illustrated here by assessment criteria generated from a recount text. In this instance we also have raters and their respective leniency and severity. In Figure 4.3, we see high-performing students, and the most difficult items on the test, at the top of the image. Where a student aligns with an item directly, the student has a 50% chance of completing this item successfully. We immediately get a visual idea of which criteria students found most

difficult, how many students were able to accomplish these criteria successfully, and we also see groups of criteria of similar difficulty. In this case, it is interesting to note that the easiest items are lexico-grammatical items. By triangulating this visual data with qualitative data, we can begin to examine which parts of the syllabus students are struggling to understand, and are problematizing the current teaching and learning cycle.

The second value of Rasch measurement in an action research approach is the use of fit statistics (Linacre, 2007a). Fit statistics are generated from the differences between observed responses and statistically expected responses, known as “score residuals” (Bond and Fox, 2007). For example, in a test where observed responses have values of 0 or 1 (wrong or correct), expected responses can take any value between 0.00 and 1.00, so observed values and expected values can never be exactly equal. When person ability and item difficulty are perfectly matched, the probability of success equals 50%, so the residual is 0.50 for a correct answer and -0.50 for an incorrect answer. Very low residuals occur when high-proficiency students succeed on easy items or low-proficiency students fail on difficult items, while large residuals will occur when low-proficiency students succeed on difficult items or high-proficiency students fail on easy items. The mean-square fit statistics provide an analysis to determine whether the observed data fit the expected distribution. The mean-square statistic has an expected value of 1.00, indicating observations that perfectly correlate with the Rasch model, with a lower limit of zero and no upper limit (Bond and Fox, 2007). Mean-square values below 1.00 indicate responses that are more predictable than expected, called “overfit”, while mean-square values greater than 1.00 indicate less predictable responses, called “misfit” (or “underfit”). Engelhard (2009) used similar fit data to illustrate its usefulness in the development of test modifications for students with disabilities. In his study, he used fit statistics to identify which individuals in large test cohorts were not performing as expected. By locating “mis-fitting” students in the quantitative data, and using a mixed-methods approach to triangulate findings with detailed qualitative interpretation,

Engelhard (2009) was able to identify disabled students who would benefit from the modification of test administration and teaching methods.

In the case of the current research, these fit statistics could be used to identify students in need of remedial attention, and could give an indication of how to deliver more detailed and targeted feedback on classroom performance to individual students. Fit statistics could also be used to locate mis-fitting items. Items that do not respond as expected in the test might indicate that they are being fundamentally misunderstood by students during the teaching and learning cycle, and emphasize a gap between student and teacher expectations. In conjunction with the visual assistance of the variable map, differences between student and teacher ratings of their performance on the test might also highlight a fundamental gap between student and teacher understandings of what has taken place in class, and a gap in student and teacher expectations. The triangulation of quantitative and qualitative data in these methods goes some way to addressing the criticisms of action research for being too subjective, as cross-checking will strengthen findings and objectivity.

4.9.2.2 Software: Winsteps and Facets

Winsteps (Linacre, 2007b) and Facets (Linacre, 2007a) are software tools that utilize the Rasch model. Winsteps (Linacre, 2007b) conducts Rasch measures from simple datasets of two facets, usually persons and items, and with a much simpler interface for entering and interpreting data output. Facets software (Linacre, 2007a) is designed to handle many-facet Rasch measurements by constructing measures from combinations of different facets, such as persons and items, as well as raters, tasks, and other structural facets. This many-facet capability means that data entry, conceptualization and interpretation are far more complex. In this study, during Cycle One, Winsteps (Linacre, 2007b) was the chosen method of data analysis, due to the simpler nature of conceptualization and interpretation, my own level of knowledge of Rasch measurement, and the two-facet data I was generating. As my

knowledge of the model and the software increased during the research process, and with the inclusion of further facets such as student and teacher raters, in Cycle Two the Facets (Linacre, 2007a) software was adopted.

4.10 Presenting the Results

During the writing process, Burns (2010), Bazeley (2009) and Miles and Huberman (1994) suggest numerous forms of visual aid to assist displaying data; they are particularly useful, firstly for data comparison and secondly for data conclusions. Bazeley (2009) describes matrix displays as an extremely effective way of detecting patterns in data. The matrix highlights frequency of responses and details of their content, allowing an analysis of how data varies under different circumstances and how often data events occur. This study will employ matrix displays to present data as they undergo various levels of abstraction during analysis. For conclusion purposes, Bazeley (2009) and Burns (2010) suggest flow charts and models to represent findings. Bazeley suggests that models are simplified versions of findings. This simplification is a process that polishes findings into theory. Chapter 5 begins the presentation of data analysis from Cycle One, with qualitative and quantitative analyses presented in broad themes generated from the raw data. Avenues are suggested for further investigation, comparison and abstraction, in preparation for Cycle Two of the action research process.

4.11 Summary

In this chapter, I have outlined the basic structure of the thesis, and the planning and action stages of Cycle One of the research. The construction of a genre-based syllabus was presented, and data collection procedures and planned analyses were outlined. Ethical considerations and procedures resulting from this data collection process were identified. The next chapter will begin the presentation of data analysis, and the “observation” and “reflection” phases of the action research structure adopted in this thesis.

Chapter 5: Cycle One: Observation and Reflection

5.1 Introduction

In this chapter, data analysis will be explained in detail, followed by a reflection on the data analysis that was undertaken in Cycle One. Section 4.9 illustrated the steps taken in preparing the data for analysis. In this chapter, there is a description of when, how and what reflection took place in the action research cycle, and how and why action research cycles will continue to be developed into Cycle Two.

5.2 Observation: Undertaking Data Analysis

Burns (2010, p. 104) suggests that one can begin to understand the meaning of the data collected during action research by asking:

- Do these data answer my questions? If so, how?
- What are the main messages so far?
- What are the gaps in the messages I still need to fill?
- Am I still asking the right questions?
- Do I need other kinds of data?
- Are some pieces of data more important than others?

At this stage in the action research process, I began assembling my data and looking for broad patterns and trends that addressed my research question. From these broad patterns I began to refine the data into two specific categories that will be discussed in greater detail in section 5.3.

5.3 Qualitative Data Analysis

Figure 4.2 was an attempt to summarize the analysis of qualitative data during the observation phase of Cycle One, including the simplification and abstraction of the raw data into broad themes with the research questions in mind, and triangulation to find patterns in

the data. During Cycle One, four broad themes emerged from the qualitative data; these are now discussed in greater detail.

5.3.1 Student emotions

A central theme that repeatedly and consistently appeared in the English language data is words and ideas that reflected the *emotions* of students. The word “nervous”, or a synonyms of “nervous”, were among the most frequent words to appear in the data, and in a range of sources. Notions of frustration were also expressed in the datasets. These included comments students made about not being able to complete a task, or finding a criterion or task too confusing or too difficult. However, not all instances of student emotion were negative: there were also examples of enjoyment and satisfaction. Figure 5.1 highlights and summarizes instances of student emotion in the data during Cycle One, as well as the data source, and the time period with which the instance is associated. Exempla are presented in the form of representative quotations. Sections 5.3.1 and 5.3.2 will discuss the data displayed in this matrix in further detail.

Time period	Student Emotions				
	Source	Nervousness	Frustration	Enjoyment	Satisfaction
Post-assessment	Student Feedback	<i>I was very nervous</i>	<i>I was confused by unexpected questions</i>	<i>It was a little bit fun</i>	<i>I could say what I wanted to say</i>
	Teacher Comments	<i>Some students were visibly shaking</i>	<i>Students dwelled painfully on mistakes or unknown answers rather than allowing the conversation to flow naturally</i>		<i>They were clearly pleased when finished</i>
Weekly	Teacher journal	<i>When talking to me (in English) they are more nervous than with their partner</i>	<i>One student commented that they were frustrated about understanding the questions, but not knowing how to formulate a response in English</i>	<i>Students seem to enjoy the opportunity to talk to each other</i>	<i>Audible sighs of relief</i>

Figure 5.1 Matrix display to examine patterns in student emotions

Nervousness, or synonymous emotions, were the most frequent emotions most easily identifiable in the data. Nervousness was reported by students themselves, and was often visible to teachers through students' body language and speech patterns:

Some students were visibly shaking

Figure 5.1 illustrates that nervousness was reported primarily when students were undertaking some form of teacher assessment, or when talking face-to-face with the teacher. Nervousness in a non-test situation and when talking to classmates was infrequently reported; emotions were generally positive at this time. During in-class practice, students and teachers also reported student enjoyment or satisfaction. The context in which nervousness was most

apparent was during the final assessment, which included a conversation between teacher and student (see section 5.5 for exemplar transcript). The main reason for anxiety might be explained by Poynton's (1990) description of tenor as being broken down into three continua: power, contact and affective involvement (see Figure 2.2). The teacher–student relationship promotes unequal power. Teacher–student contact is occasional, and affective involvement is low. Student nervousness arose in the data when students spoke with the teacher, even when formal assessment was not the key goal of the conversation. Another reason for nervousness might be that a casual conversation includes visual and aural contact, and feedback is immediate, as illustrated in Figure 2.3 (Poynton, 1990). The nature of a casual conversation means that students' use of the English language is accompanying the action. The notion of immediate feedback, which might include negative feedback from a person whom they consider a model of accurate language, would be extremely daunting to students.

Figure 5.1 also shows that frustration was a major theme of the data. Student frustration was primarily expressed by students through the use of negative verbs such as “I can't” or “I didn't” in relation to a particular criterion or activity in the classroom. Frustration was far more apparent after the final assessment than in data collected before the assessment. One reason for this might be the anonymous nature of feedback after the final test, allowing students to be more critical. However, this criticism was not usually aimed at the syllabus itself but towards students' own performance. Student judgments on their proficiency during class-time were sometimes negative, but often tempered with positive comments:

I didn't understand what I will do, but gradually I understood

Or, positive verbs such as “I could”:

I could speak it fluently

After the final assessments, students' feedback contained a larger number of negative verbs describing their performances:

I couldn't come up quickly with the words

A major contributor to the theme of frustration was reflection on practising the micro-aspects of discourse, such as question/response. Micro-aspects of talk are discussed in section 2.9, and identified as "turn-taking" in section 4.8.2. These are aspects of talk that do not contain easily identifiable generic and monologic structures, and are usually spontaneous. These micro-aspects of conversation might include recasts of language; expressing support or surprise; providing helpful information, or help with vocabulary choices; question and answering (see section 2.9). Teachers noted that some formulaic greetings were misunderstood, for example:

Teacher: *How's it going?*

Student: *By bus.*

The above interaction was noted in the teacher journal as occurring quite often, due to students attaching significance to the words "how" and "going" without recognizing it as a formulaic expression. Students expressed frustration and even embarrassment when this mistake was pointed out, or they realized during self-reflection afterwards. Frustration was also recorded during the final assessment and during in-class activities, when conversations were ended abruptly or contained long pauses. Pauses occurred as questions were either not understood, or when students were unable to formulate a suitable English response. In peer-assessment, students often had great trouble asking and answering questions to encourage extended dialogue.

Turn-taking during conversation proved to be a large source of frustration for many students, shown through such comments as:

I couldn't answer questions smoothly. I want to speak smoothly.

I noted in my journal that rather than using strategies to continue the conversation, some students would take extended periods of time to ensure they gave the correct answer to whatever question or idea they were confronted with, at the expense of fluency. In a natural conversation this would usually create feelings of discomfort, so I noted in my journal my own feelings of frustration that the syllabus was not addressing issues of fluency. Feelings of frustration, therefore, developed from the gap between receptive understanding and productive ability in the target language. This frustration was exacerbated in the final assessment, due to a gap between what students thought they could achieve during class time with their peers, and what they felt they could achieve with the teacher. When micro-aspects of discourse were introduced to the conversation by the teacher in the final assessment, it created a sense of frustration. During class practice, students were unable to replicate micro-aspects of discourse by themselves, so this interaction simply did not take place. As well as nervousness, therefore, frustration also increased when students engaged in casual conversation with the teacher. This was due to the teacher introducing micro-aspects of discourse to the conversation, which increased the difficulty level of the speaking interaction.

Figure 5.1 also illustrates that not all emotions experienced by students during the course of the syllabus were negative. There were also elements of enjoyment and satisfaction. Enjoyment was often closely associated in the data with terms like “opportunity” or “chance”, as students described how using the syllabus gave them opportunities to use English in class for a purpose. This sense of enjoyment appears to be closely linked to feelings of satisfaction, as students and teachers commented on being able to achieve

something students had not necessarily had the opportunity to do before: to speak for an extended period of time in English:

Students seem to enjoy the opportunity to talk to each other

Even after assessment, many students displayed either through body language or through feedback that it was a worthwhile achievement to be able to speak beyond one or two-word answers. Some students even showed surprise at what they were capable of achieving. In future cycles of action research and syllabus design, detailed feedback that promotes the sharing of learning goals and how students were meeting them might work to lessen feelings of nervousness and frustration. Students who had also shown frustration and nervousness also indicated enjoyment and satisfaction, both during the syllabus and after the completion of the final assessment:

I was confused by being asked an unexpected question, but I said what I wanted to say

One student mentioned that their nervousness was so great they thought they were going to die, but were finally pleased to announce that they did not, in fact, die, and could have a conversation in English. The data seems to suggest that enjoyment and satisfaction were closely related, and that a genre-based approach afforded increased opportunity to speak in English, which students found rewarding and even surprising.

During coding, I cross-referenced with the established literature the apparent themes of satisfaction, and there appeared to be a resonance with definitions of intrinsic motivation. Ryan and Deci (2000) define intrinsic motivation as motivation that is displayed via someone acting for the enjoyment entailed in an activity, rather than due to external pressures. This is in contrast to extrinsic motivation, which relies on external rewards to get a task done. In the case of this study, the extrinsic motivator, i.e. the final interview test, appeared to reduce motivation by increasing nervousness and frustration. Ryan and Deci (2000) explain that

extrinsic motivators such as a test can lead to a reduction of intrinsic motivation by leading to over-justification of outcomes. During the final interview, there were more instances in the data of nervousness and frustration. However, during class time, teachers observed, and students reported, enjoyment and satisfaction. Ryan and Deci (2000) outline three defining characteristics of intrinsic motivation that appear to reflect comments in student feedback, as follows: belief in a personal control over their learning; belief that they have the ability to attain their desired goals; and a desire to master a topic beyond short-term goals such as exam results. Through a process of deductive coding, it would appear that intrinsic motivation was promoted during the syllabus, but that the syllabus would benefit greatly from giving students some control over their learning; it should also match their ability levels to learning goals, and make these learning goals specific, in order to meet the definitions of intrinsic motivation offered by Ryan and Deci (2000).

5.3.2 The model text

As first illustrated in section 4.8, reflection-in-action highlighted that the modelling of the text was a persistent problem during the creation and implementation of the syllabus, and this issue was also a major theme during reflection-on-action (Schön, 1983). Figure 5.2 categorizes the raw data into themes associated with the modelling of the text, and also allows cross-referencing of data between different sources. The main themes that emerged from the data under the concept of modelling the text were *text authenticity*, i.e. the mode of the text, and *the model speaker*, i.e. who modelled the text. Within these themes, sub-categories included the differences between spoken and written language, issues concerning pronunciation and fluency, and providing model texts that were appropriate to the level of the students and desired learning outcomes.

Time	Data Source	Text Authenticity		The Model Speaker	
		Idealized Text	Transcribed Text	Pronunciation and Fluency	Learning Goals
Cycle One	Teacher Journal	<i>Students transcribed their "conversations" in preparation for the final assessment</i>	<i>Helped students to think more about how to speak quickly</i>	<i>The native-speaker model is too complex and rapid for them to understand</i>	<i>The [written] model text is too perfect. It's setting a standard of language that is beyond them and they're marking themselves negatively</i>
	Classroom Documents	<i>There is more writing than speaking taking place</i>	<i>They seemed to understand the differences between speaking and writing</i>	<i>They're looking shocked at how they think they're expected to speak when shown the "native speaker"</i>	<i>Students are relying heavily on writing and memorizing rather than having a conversation</i>

Figure 5.2 Matrix display to examine patterns in response to the model texts

Students did not comment explicitly on the model texts; however, I was able to observe the effect of the model text on students; and as the primary researcher during Cycle One, I was able to contrast student reactions to the model texts as I adapted them during reflection-in-action.

When the syllabus was first introduced to students during the pilot study, model texts used for analysis in the classroom were based upon idealistic written transcripts of speech, much like students were used to reading in their textbooks (see Figure 4.1). The use of written transcripts raised immediate problems that were a constant theme of the teacher journal and written notes on lesson plans during Cycle One. Chapter 4 outlined how the idealized recount text was replaced with an idealized transcription (Appendix 13), followed by the introduction of audio clips of native speakers (section 4.8.2). In Figure 5.2, "Text Authenticity" refers to the extent to which the model texts chosen for the syllabus in this study relied on spoken or written examples that were either designed or created by the teacher for the specific teaching situation, as opposed to genuine instances of conversations that were taking place for real-life purposes outside the classroom. It quickly became apparent that written models of spoken language were not appropriate for speaking activities, particularly for creating conversations. Numerous entries in the teacher journal show that before being

asked to practise conversations, students were writing dialogues and memorizing them; thus exacerbating the tensions between spoken and written English, as illustrated by Eggins (2004) outlined in Chapter 2. Their spoken conversations followed a monologic organization, with standard grammar and a polished aspect to their fluency that did not sound like natural speech. Their conversations lacked elements of turn-taking and spontaneity. Students themselves seemed pleased with their ability to construct longer utterances, and I remarked in my teacher journal:

Students are definitely using more English in the classroom, in amongst the noise I can hear much more English as opposed to Japanese, or silence

Students were using English in utterances that lasted for longer periods than was the case before the syllabus was introduced, but there was still a clear gap between student performance and teacher goals. The aim of the syllabus was to achieve authentic casual conversation, rather than reciting long memorized monologues to a partner, with little interaction. Although longer utterances and the accurate use of grammar and vocabulary were pleasing, their talk was not illustrating the typical characteristics of spoken language. The teaching journal reflected a concern regarding a long-term effect of the syllabus: that students might only understand and produce a simplified form of language, which would not allow them to cope with language in the real world:

although my students have a low proficiency and it is nice to hear them using English in longer stretches, is this really English?

The journal reflected my concerns that students might even undermine their language-learning skills:

by giving students a false model they may be missing out on gaining skills that help them to decode real language when they hear it outside the classroom

Although the model was assisting my students' understanding of grammar and how conjunctions could help them to sequence events and create longer utterances, it was not helping them to hold casual conversations, particularly with strangers. The model text complemented students' accuracy, but was inhibiting their spoken fluency.

The aim of the transcription-based model text shown in Appendix 13 was to introduce the idea of turn-taking in conversation, and the necessity of spontaneity when constructing longer utterances with a partner. The new model and transcript also allowed classroom discussion about the differences between spoken and written language as an explicit element of instruction. However, turn-taking and the micro-aspects of conversation continued to be difficult and frustrating elements of the syllabus, even with the introduction of the transcript shown in Appendix 13. Students maintained their reliance on creating monologic chunks of language that they could memorize and repeat to their partner. Comments in the teacher journal and in classroom documents show that aspects of turn-taking only took place at the end of extended monologic chunks that had been pre-prepared:

Students are waiting for their partners to finish their whole monological chunk before asking any questions.

Classroom documents collected from students showed hastily written transcripts scribbled on their worksheets in preparation for speaking with a partner, sometimes in Japanese and translated into English. It was clear from the data that the lesson plans were teaching students to write, memorize and recite, rather than engage in casual conversation. The aims and goals of my syllabus were not being fulfilled. It was clear that during Cycle Two, new strategies for introducing model texts were necessary, and that the aims and goals of the syllabus needed to be made more explicit, particularly in regard to turn-taking and emphasizing the differences between spoken and written language.

As well as notions of authenticity in model texts, the concept of the “native speaker” of English also arose in the data. This notion creates various issues, many of which might be political or ethical considerations that inform syllabus design. Tollefson (2000) warns that the spread of English promotes significant social, political and economic inequalities. Tsuda (1994) suggests that the teaching of English is driven by Anglo-American policies of monolingualism and cultural imperialism. Tsuda explains that in a Japanese context, English is often seen as a vital skill for students’ economic futures, with seemingly little attention given to these political considerations. In the context of this study, instructors such as myself are officially referred to as “native speakers” in curriculum documents. The data collected in this study showed that notions of a “native speaker” had quite profound implications for students’ language acquisition. After introducing the “native speaker” audio in response to reflection-in-action, new themes began to emerge in the data. Lesson plans show that the first model text I presented to students was between two “native speakers”, one from the US and the other from Wales. Themes of fluency and pronunciation arose in the data in connection with this model text, exemplified by a comment from a student journal:

Native speed is so quick!

And comments in my own teaching reflections:

The native-speaker model is too complex and rapid for them to understand

They’re looking shocked at how they think they’re expected to speak when shown the “native speaker”

In focusing on the authenticity of model texts, ethical considerations of model language use had not been addressed.

The examples from the student journals show that students believed the learning goal of the class was now to be able to speak as a “native speaker”, and that anything less than this was seen as a failure on their part:

Speaking skill is important but it's hard because of Japanese accent

A large gap was developing between teacher and student expectations initiated by the introduction of native speaker models. Notions of accent as a measure of spoken performance were troubling concepts to the participants in this study:

I need to practise more American accent

My teaching journal noted that:

Students seem to be under the impression that they need to speak like me by the end of the course, mostly they respond with amused resignation!

The abstract notion of what makes a “good English speaker” created a large gap between perceptions of students’ own proficiency in English and teacher expectations, as well as producing negative emotions. It was a clear necessity to promote in class the notion of mutual intelligibility. Studies have shown that in some cases, speech marked by non-native accents is understood by native listeners just as well as native-produced speech from a familiar dialect (Munro and Derwing, 1999). However, an accent can sometimes have adverse consequences for the L2 speaker (Flege, 1988), when listeners can have serious problems in understanding accented speakers. This was of particular concern to me, as further studies have shown that non-native speakers can often face prejudice and discrimination due to their speech patterns (Derwing, Rossiter and Munro, 2002; Dávila, Bohara and Saenz, 1993; Derwing and Munro, 2005; Munro, 2003).

Therefore, issues of pronunciation and fluency not only raised ethical considerations, but also issues of concrete and relevant learning goals for my students. Becoming a “native speaker” was clearly not a viable learning goal, either ethically or academically. A note in my research journal highlighted the need for mutual intelligibility:

It should be made clear to students that I don't expect them to speak like me at the end of the semester, but that they can have a casual conversation as Japanese speakers of English that can be understood in a global marketplace.

Reflective data showed that in order to empower students and demystify learning goals, a notion of mutual intelligibility was essential. Derwing and Munro (2003) and Jenkins (2006) argue that there is no universally agreed definition of what constitutes “mutual intelligibility”. Smith and Nelson (1985) attempt to define intelligibility in three parts: the ability of the listener to recognize individual words or utterances; the listener’s ability to understand the meaning of the word or utterance in its given context; and the ability of the listener to understand the speaker’s intentions behind the word or utterance. In Cycle Two, therefore, I decided that it would be advantageous for students to share Smith and Nelson’s definition of intelligibility; I would integrate it into lesson aims and goals, as well as assessment, which would then be made explicit to students. In Cycle Two I felt it would be necessary to quantify issues of pronunciation and fluency, and achieve concrete learning goals relevant to my students’ needs. This will be discussed in further detail during the “planning” phase of Cycle Two (section 6.3).

5.4 Quantitative Data Analysis

Section 5.3.1 highlighted the nervousness and frustration that students felt when facing spoken assessment. Assessment caused anxiety in students, particularly when conducted face-to-face with the teacher during speaking. During a conversation, assessment increased inequality in the power relationship between the interlocutors, thus making casual

conversation more difficult. Section 5.3.2 also highlighted how students were misunderstanding learning goals when the model of a native speaker was introduced. “Native-speaker” models implied assessment criteria for students that were vague and abstract, particularly in terms of pronunciation and fluency, and unobtainable within the time constraints of the learning programme.

During Cycle One, whilst assessment considerations were integrated into the original syllabus design at the start of the action research process, these were largely summative in nature; and although they provided useful quantitative data for analysis, they did not necessarily adequately address the aims of the syllabus. Criteria for assessment were based upon the deconstruction of a personal recount text, and reflected the concepts in the literature as outlined in Chapter 2. Additional criteria attempting to address concerns about fluency and pronunciation were introduced with reference to the Common European Framework (section 4.8.2). Section 5.4 discusses the quantitative analysis of students’ spoken assessment using the Rasch model, and triangulates these data with the two major themes of “student emotions” and the “model text” generated from the qualitative data.

As part of the syllabus design and criteria generation for assessment, it was useful to create a latent variable map that hypothesized the difficulty students might have with various elements of the syllabus. This hypothesis could inform syllabus design and lesson content based on the assumed needs of my students. These assumptions could then be tested, challenged and verified, or used to inform future syllabus design. For example, criteria that were hypothesized to be difficult might prove to require less attention in the classroom than was first assumed. Equally, criteria that the teacher thought were not difficult might actually prove more challenging to students than assumed. Table 5.1 shows the hypothesized difficulty level of criteria generated for assessment, and to inform syllabus design:

Table 5.1 Hypothesized latent variable map for the generic structure and textual features of a recount genre

Logit Scale	Student's use of structure and features	Structure and textual features
5.00		Answers questions
4.00	High use of features	Repeats
3.00		Illustrations
2.00		Pronunciation
1.00		Fluency
.00	Moderate use of features	Temporal conjunctions
-1.00		Past tense
-2.00		Orientation
-3.00		Sequence
-4.00	Low use of features	
-5.00		
Polytomous rating scale used (X=0,1,2)		

Table 5.1 shows that it was hypothesized that turn-taking, such as answering questions, as well as fluency and pronunciation, would prove to be the most difficult aspects of the syllabus for students to master; whereas lexical-grammatical criteria, such as using temporal conjunctions and past tenses, might prove to be easier. Table 5.1 shows that during assessment, scores were awarded on a rating scale of 0, 1 and 2, depending on the fulfilment of these criteria as judged by the teacher. A score of 0 was awarded to students if they did not use that particular item in their talk; 1 indicates that students sometimes used that item or used it fairly effectively; and a rating of 2 means that they used the item effectively and often. In previous sections, the necessity of making this scoring system explicit to students was

clear, and the necessity for a more robust rubric is, therefore, discussed in section 5.6, and in the planning stage of Cycle Two in Chapter 6.

During assessment, students used the content provided in the planned teaching and learning cycle to individually construct their own recount texts in a conversation with the teacher in the final classes. When participating in the conversation as an interlocutor, I awarded scores for the relevant criteria based on my own judgement. The criteria for assessment are presented in Table 5.1 and discussed in greater detail in Table 4.3 and Table 4.4. After the assessment, the scores were prepared to form a text file for entry into a Rasch analysis using Winsteps (Linacre, 2007b); a data control file is presented in Appendix 14 to illustrate how data were organized for this purpose (Linacre, 2007b). The analysis in this section is for the 27 students included in Cycle One of the action research. Figure 5.3 shows the assessment criteria ordered vertically according to their difficulty, as determined by the Rasch measure in response to student scores; these can be compared with the hypothesized difficulties presented in Table 5.1. At the bottom of the hierarchy of difficulty we see “Sequence”, thus indicating that students found sequencing events to be the easiest item on the test. The most challenging item on the test was “Answer Questions”, which correlates with the hypothesized order of difficulty, thus providing preliminary evidence of construct validity.

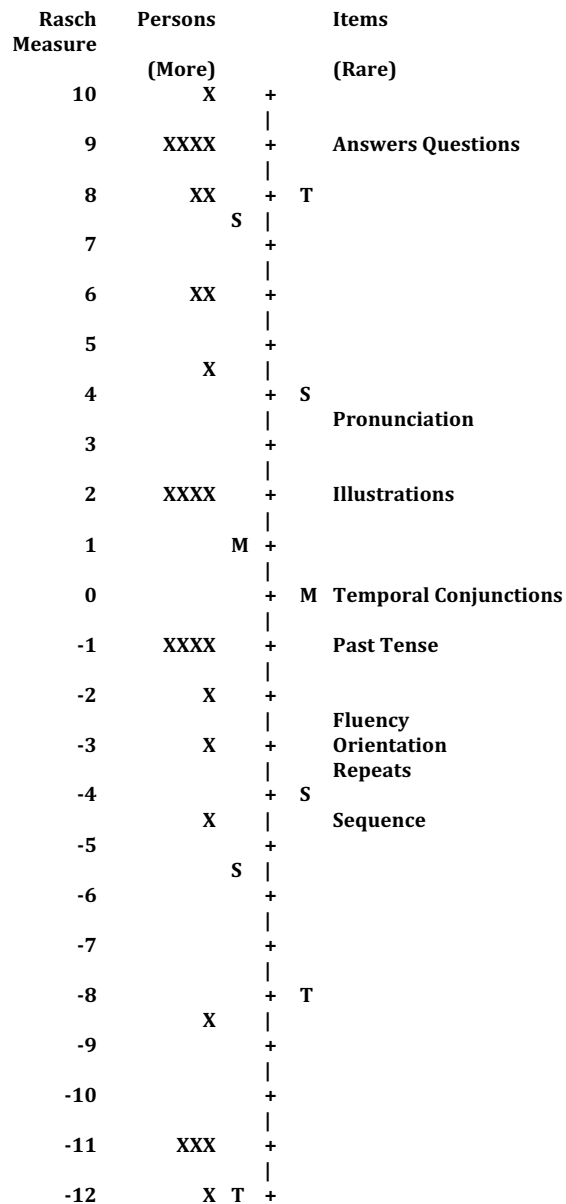


Figure 5.3 Person-Item map showing persons on a common scale with items

Each “X” in Figure 5.3 represents one student, and the higher the “X” appears on the figure, the more proficient the student’s language ability, based on this measurement. “M” shows the location of the mean of the persons or items, and “S” shows one standard deviation above or below the mean. The vertical spacing is the approximate placement of the items on the linear Rasch dimension, so that “pauses” to “past tense” has roughly the same increase in item difficulty as “illustrations” to “pronunciation”. In the test, five students are performing

at least one standard deviation below the person mean, and seven students are performing one standard deviation above the person mean. Most students taking the test performed within one standard deviation of the mean. We can see that four students have ability equal to “answers questions”; so Rasch calculates that they have a 50% expectation of success on this item, based on inferences drawn from response patterns in the data (Linacre, 2007b). Their probability of success in easier items increases above this 50%.

From Figure 5.3, it was possible to identify a number of problems with the assessment. Firstly, the test had been too easy for the highest-level students, who were able to tackle all items effectively; while conversely, four students could not perform effectively on any items in the test. The implications of this were either that the assessment needed more very easy and very difficult items in the test, or that some students were not learning these items effectively in the scheme of work in the classroom. This was useful as a diagnostic tool, as it informed my next cycle of research and indicated problematic points in my syllabus; it also allowed targeted support for individual students, perhaps by differentiating work for them in the classroom.

A potentially valuable, but underdeveloped, use of Rasch analysis of classroom tests is shown by Engelhard’s (2009) investigation of person fit statistics as a diagnostic tool in mixed-method research: this illustrates the duality of Rasch analysis, where the same analyses can be conducted for persons as well as items; thereby allowing mis-fitting persons to be identified and qualitative investigation to be conducted, in order to determine causes and possible remedial intervention. For illustrative purposes, the original dataset of 12 items and 27 persons is shown in Table 5.2, with persons arranged in order of fit.

Table 5.2 Person correlation and fit

		Infit	Outfit
Person			
No.	Score	MnSq	MnSq
12	7	1.20	5.43
14	22	1.32	5.40
10	9	1.87	2.64
26	16	1.77	1.36
6	8	1.68	1.25
16	13	1.40	1.05
7	11	1.34	1.14
25	23	1.22	.65
3	12	1.21	.91
11	22	1.02	.57
5	20	1.01	.61
24	21	.99	.58
2	17	.89	.53
18	22	.89	.43
9	18	.87	.50
19	15	.78	.51
1	23	.76	.26
4	23	.76	.26
23	6	.68	.34
20	15	.47	.30
21	15	.47	.30
8	12	.42	.31
13	16	.38	.25
15	13	.36	.27
22	13	.36	.27
MEAN	15.2	.96	1.04
S.D.	5.7	.43	1.39

In Table 5.2, Infit MnSq, Outfit MnSq and Standard error are all “fit” statistics that indicate how accurately or predictably data fit the Rasch model. “Outfit” is an outlier-sensitive fit statistic, and sensitive to unexpected observations by persons on items that are relatively very easy or very hard for them. “Infit” is an inlier-sensitive fit statistic that is more sensitive to unexpected patterns of observations by persons on items that are roughly targeted at them (Linacre, 2007b). Infit was an innovation of Ben Wright (Bond and Fox, 2007), who noticed that the standard statistical fit statistic (which we now call “outfit”) was highly influenced by a few outliers (very unexpected observations). He therefore devised the infit statistic, which was more sensitive to the overall pattern of responses. Infit weights the observations by their statistical information, which is higher in the centre of the test and

lower at the extremes. The effect is to make infit less influenced by outliers, and more sensitive to patterns of inlying observations.

Five persons show misfit large enough to warrant investigation. Person number 14 is extremely proficient, scoring 22 out of a possible 24, and has acceptable infit, but very large outfit, which is consistent with failing on a very easy item. This result is probably of no concern, but investigation of unexpected responses can clarify the reason for the misfit. Persons 12, 10 and 6 are of low proficiency, and have misfit that warrants further attention. Not only are they of limited proficiency, but they do not respond consistently with the other people's response patterns. Person 26 is of slightly higher than average ability, but has an infit mean-square figure of 1.77, so is also of concern. These mis-fitting students are deviating from the latent Rasch trait that defines the expected trajectory of this sample of persons through this curriculum; this identifies them as possible candidates for remediation.

Qualitative investigation of these students helped to identify the causes of this misfit. Person 6, for example, had very poor attendance, and was therefore unprepared for this form of summative assessment. Person 14 wrote a number of comments in their student journal that showed confusion about some of the criteria:

Without temporal conjunction the meaning of the sentence is hard, so I need to make it soft, but it is very difficult to understand

I couldn't understand about linking the content of my story, I'll try and practise every day to comprehend it

I can understand past tense verbs, but sometimes I forget the irregular verbs, I have a little bit unease about using in conversation. I want to practise more

In this particular assessment, we see that more consideration is needed in preparing students for assessment in order for their abilities to be measured more effectively; and in this

case, it appears that criteria need to be made more explicit to students, so that they know exactly what is expected of them, in order to reduce misunderstanding. During Cycle One, although intricacies of each criterion were discussed in class and exemplified via model texts, an explicit assessment rubric was not shared with students, and model texts proved to be confusing, or prompted students to rely on written output. Section 5.3.2 described how model texts were leading to a gap between student and teacher expectations, which led to frustration and nervousness. In future cycles of action research, it would be necessary to narrow this gap. Moreover, in further cycles of research, it would appear that an assessment rubric might need to be constructed beyond the simple assigning of scores on a polytomous rating scale, and that this should be made explicit to students. These findings correlate with themes presented in the qualitative data analysis in section 5.3.

In Figure 5.3, we see that criteria based on elements of generic structure and lexico-grammar in a genre are the easiest criteria for students. By contrast, criteria that are not necessarily genre-specific, but universal speaking abilities, were by far the most difficult for students. In Figure 5.3, the criteria are separated into two quite distinct groups according to level of difficulty. It proved problematic to measure the many different aspects of speaking ability on one unidimensional line, as Rasch analysis dictates. For example, one of my students who had performed very well in a classroom environment became extremely anxious in a spoken conversation, especially in the assessment situation. In such cases, it was therefore impossible to measure any of that student's language ability, since the student was unable to produce language. Measuring students on one summative assessment did not seem to be an accurate measurement of their overall speaking ability, and might equally have mirrored my own teaching ability and syllabus design. As discussed in previous chapters, micro-aspects of conversation were more difficult to assess than generic structural and grammatical elements of a recount genre. Such a speaking test, therefore, is likely to be measuring two different forms of spoken ability: mastery of structure and lexico-grammar;

and mastery of spoken output, including pronunciation, fluency and turn-taking. This means that one summative test for students' speaking may not be appropriate or an accurate reflection of their skills, and that other forms of assessment are necessary. In my teaching journal I noted:

Some of the students did not perform as well as I expected. How do we assess items that perhaps do not fit onto a unidimensional speaking assessment like the one in this study? Some students are clearly able to internally organize ideas and grammars, but personality issues may mean they are not able to express them effectively externally, so multiple assessments addressing different performance traits may be necessary.

My teacher journal also reflected other concerns:

Whilst the criteria provide a nice framework for informing syllabus design and helping students perform longer utterances, I feel that this approach is too prescriptive, especially with the final test, there does not seem to be much room for experimentation with language and the kinds of mistakes and spontaneity you might find in casual conversation

Based on a triangulation of both quantitative and qualitative data, it became clear that a genre-based approach must entail greater formative assessments being integrated into the syllabus at various stages; it must also provide opportunities for students to experiment with language that allows them to be more spontaneous, and also less anxious about making mistakes. Formative assessment would be able to address not only structural and lexico-grammatical elements of spoken output, but also universal speaking abilities such as fluency and pronunciation. An assessment based solely upon the structural and lexico-grammatical elements of speaking might not be adequately assessing more universal speaking abilities; it might also prove to be too prescriptive, by not allowing students to experiment, or to have sufficient confidence to be spontaneous with language. As such, the syllabus design in further

cycles of action research needed to find ways to integrate formative assessment, and also to remove teacher-dominated summative assessments that led to nervousness and frustration.

Forms of self-assessment might remove feelings of nervousness and frustration by correcting the power imbalance between teacher and students in interactions, as well as creating the necessity for making criteria for assessment explicit and clear to students.

5.5 Post-Syllabus Student Dialogue

Section 3.5.1 gave an example of students' recount texts before the syllabus intervention. The dialogue below provides a representative example of students' recount texts after completing the syllabus:

A: Teacher

B: Student

1. A: Hello!
2. B: Hello.
3. A: How are you?
4. B: I'm alright. And you?
5. A: Yeah, I'm good! What have you been doing recently?
6. B: During Golden Week my friends and I went to Marinoa City by car. First we went somewhere and ate some food, then my friend gave me some present.
7. A: Really? What did you get?
8. B: It was a key holder and some chocolate. I was so happy!
9. A: Why did your friend give you a present?
10. B: She came back from America.
11. A: Ah, it was a souvenir?
12. B: Yes! Next we went to karaoke and sung many songs.
13. A: Did you sing in English?
14. B: ...
15. A: Did you sing an English song?
16. B: No...Next, we went to café and eat cake. It was delicious. Next went to the shop and I bought this top.
17. A: Oh, it's nice. Was it expensive?
18. B: It was 1500 yen. Next, we went to game centre. It was fun. Next, we went to another shop and bought a key holder, it was very cute bear.
19. A: Oh, what colour?
20. B: Black... Next, we went to restaurant and ate omelettes and rice. It was delicious, and I was, I was full. At the end we went to car and came back my house. I had a great time.
21. A: Sounds like it was really fun.

In line 6 the student fully orientates the listener with key information at the start of the conversation, detailing who, what, where, when and how an event took place. For this reason, the student scored two points on the rating scale described in section 5.4, for the criterion “orientation”. In the same line, the student then begins with a transition phase to introduce the first in a sequence of events, quickly followed by another. This unprompted monological structure highlights some of the issues identified in this chapter: that this student has probably prepared and memorized a written text for spoken output. The student accurately uses past tense verbs and transitions, but without opportunities for a joint construction of the conversation via turn-taking. In lines 7, 9 and 11, the teacher has a chance to ask questions and a short period of turn-taking is established; however, in line 12 the student quickly and abruptly re-establishes the monological nature of the text by introducing a new event, despite interlocutor B’s apparent interest in the souvenirs.

To an unsympathetic stranger, the abruptness of the interaction might appear rude, or an attempt to change the subject. From line 12 onwards, any responses to interlocutor B are answered with one or two-word responses, as the student appears to be determined to develop the memorized monologue without interruption. A mistake in verb tense is in line 16 when the student says that they “eat” cake. For this reason, the student received two points in the criterion of “past tense”. Although transition phrases are accurately used, the word “next” is used multiple times to introduce the subsequent sequence of events, adding to the sense that a memorized monologue is being shared; and one point is earned on the rating scale for the criterion “temporal conjunctions”. The repeated one-word answers mean the student scored one point for the criterion “answers questions”.

In comparison to the dialogues presented in section 3.5.1, the student’s spoken utterances are much more detailed, grammatically and lexically sophisticated and logical in their structure. Overall, however, the transcript shows that the spoken output does not really

represent a casual conversation, and is more similar to a monologue that is occasionally interrupted by a second person. There is brief spontaneity between lines 7 and 12, but in response the student then appears to revert to a monologue that appears memorized and pre-planned. The one-word responses to questions suggest the student is committed to sharing the monologue, and that any diversion in the conversation is a distraction, and possibly a threat to completing planned responses. In casual conversation outside the classroom, conversation would probably break down, and the goal of establishing social bonds via the recount text (Slade, 1997) would be unfulfilled.

5.6 Reflection

Sections 5.3 and 5.4 went some way to answering the questions proposed by Burns (2010):

- Do these data answer my questions? If so, how?
- What are the main messages so far?

In terms of the research question, “In what ways can a genre-based approach assist the teaching and development of Japanese students’ speaking abilities”, the main messages so far showed some enjoyment and satisfaction with the syllabus; but Japanese students’ nervousness and frustration with their speaking abilities was very apparent. Reasons for this included the use of summative assessment and speaking with a teacher, but also a misunderstanding of turn-taking roles in speaking, and a reliance on written output to inform spoken output. Idealized texts and transcripts exacerbated students’ reliance on written output and did not give a clear indication of, or opportunity to fulfil, the learning goals of the syllabus. Data showed that students were benefitting from, and successfully achieving, the lexico-grammatical and structural elements of recount texts, but universal speaking skills were not being addressed in the current syllabus. The native speaker model did not assist in making clear these universal speaking skills, and created within students a notion of pronunciation and fluency that was beyond their abilities, thus resulting in a gap between

student and teacher expectations. Figure 5.3 showed two distinct groups of items on the assessment: lexico-grammatical aspects; and more general speaking skills that included turn-taking criteria. Students found the turn-taking criteria extremely difficult. In that respect, there seemed to be two distinct types of tests in progress: a grammar test and a speaking test. The speaking part of the test was proving to be far more difficult than the lexico-grammatical test, and indicated potential failings in the syllabus in preparing students for spoken output. Triangulation between qualitative and quantitative data showed frustration with speaking elements and confusion over “native speaker” model texts.

Cycle One raised important questions about students’ understanding of the syllabus in terms of learning goals and what models of speaking they could deconstruct in class in order to create their own texts independently. By triangulating qualitative and quantitative data, it became clear that assessment was another major theme in the data. Section 5.3.1 illustrated how students expressed satisfaction and enjoyment when they realized that they could now speak for extended periods beyond the one-word answers exemplified in the pre-syllabus dialogue in section 3.5.1. This satisfaction could be enhanced through greater targeted feedback after their spoken assessment. Section 5.3.1 also highlighted how factors of intrinsic motivation (Ryan and Deci, 2000) could be enhanced by giving students greater control over their learning goals, and enhancing their belief in their ability through this targeted feedback. Hattie (2007) posits that an essential aspect of successful feedback is explicit and clear learning goals, which were somewhat lacking in Cycle One of the syllabus design. Section 5.3.2 shows how idealized, transcribed and native speaker texts were not appropriate in making learning goals explicit. Sections 5.3.2 and 5.4 also emphasized the need to make learning goals explicit to students, to prevent misunderstanding of criteria, and also to close an apparent gap revealed in the data between teacher and student expectations for their speaking. In the planning phase of Cycle Two, this issue would need to be fully addressed via the sharing of assessment rubric and rating scales for assessment (see section 5.4). Section

5.4 also emphasized the need for alternative forms of assessment beyond the summative, due to the unidimensionality of a test of speaking, which might not present an entirely accurate measurement of overall performance.

Such issues could be ameliorated through the introduction of self-assessment, whereby criteria are made clear, explicit and relevant to students. In order to make criteria explicit, the role and selection of model texts would also need to be thoroughly planned, in order to make learning goals and teacher and student expectations synonymous. Student nervousness and frustration would also need to be addressed in order to increase enjoyment and satisfaction. In Cycle Two, greater opportunities for enjoyment could be achieved through reducing the role of the teacher in conversations, particularly during assessment. Model texts could also more accurately reflect learning goals in class, which would not require students to believe that they had to speak like a “native speaker” or with an “American accent”. Chapter 6 will introduce the “planning” and “action” stages of Cycle Two, based upon responses and reflection on the observations generated by the data analysis undertaken in Cycle One.

Chapter 6: Cycle Two: Planning and Action

In Cycle One, three main interrelated themes that arose from the data needed attention in Cycle Two. These comprised student nervousness and frustration; unclear learning goals as a result of problems with the model texts and the use of summative assessment as the only form of assessment; and further issues of assessment concerning the unidimensionality of a test that assesses both grammatical and communicative skills. Chapter 6 describes the planning and action process that attempted to address these problems in Cycle Two.

6.1 Planning: Syllabus Design

Based on observations and reflections in Cycle One, two key areas of the syllabus described in section 4.3 required attention: the selection of model texts and assessment. Idealized and transcribed texts were leading to memorized written dialogues with unclear learning goals. Native speaker audio models were creating learning goals in the minds of students that were either too vague or unachievable in one semester, and which differed from the learning goals of the teacher. Closely tied to this were issues of assessment. I felt that learning goals needed to be made explicit to reduce student anxiety, and also to narrow the gap between student and teacher expectations. Section 6.2 describes assessment modifications to the syllabus, and section 6.3 describes the adaptation of the model text.

6.2 Assessment for Learning

Reflection on Cycle One in section 5.5 raised issues of assessment, including seeking alternative forms of assessment in Cycle Two, beyond the summative; and also making learning goals more explicit, in order to close the gap between student and teacher expectations, and enhance intrinsic motivation. After consultation and cross-referencing these themes with the literature, the notions and definitions of assessment for learning appeared especially relevant, and this section now illustrates this process of triangulating my Cycle

One reflection with this literature. Hood (2000) claims that assessment is a vital consideration in syllabus development. Feez (2002) states that assessment within a genre-based syllabus needs to be linguistically principled, criterion-referenced, explicit and continuous. These criteria for assessment mirror principles of teacher-based assessment (Davison and Leung, 2009) or assessment for learning (Black and William, 1998). This can be seen in a definition provided by The Assessment Reform Group (1999, p. 2), which defines “assessment for learning” as having the following characteristics:

- Embedded in a view of teaching and learning of which it is an essential part
- Learning goals are shared with pupils
- Helps students recognize the standards they are aiming for
- Provides feedback which helps pupils recognize their next steps and how to take them
- Both teachers and pupils review and reflect on assessment data
- Pupils are involved in self-assessment

Cumming (2009) identifies the following three issues in assessment that he argues are addressed by assessment for learning approaches: the lack of training and professional knowledge that TESOL educators receive in regard to assessment; connecting classroom assessment to relevant syllabus criteria; and utilizing assessment to promote learning. Using these references to the literature in this study, my classroom documents and lesson plans highlight the attempt to integrate such assessment criteria into my syllabus design, as informed by the data analysis conducted after the pilot study.

6.2.1 Assessment for learning in the Asian context

Despite the apparent necessity to incorporate assessment for learning principles, it was nevertheless important to consider the impact on my students. Activities such as self-

assessment were not approaches that they had necessarily encountered previously, and a great deal of thought would need to be applied to ensure that such practices were effective.

According to Davison and Leung (2009), teacher-based assessment has become institutionally adopted in a number of education systems across the globe, including Asian contexts, such as Hong Kong, China and Singapore, where assessment for learning procedures have become policy-supported. Hill and Sabet (2009) conducted a study in Japan involving speaking assessments that utilized “Dynamic Assessment”, which suggested significant cumulative improvement in learners’ speaking performance. Ishihara (2009), also in Japan, considered that there was potential for using teacher-based assessment to develop learners’ productive skills to communicate. However, widespread adoption of assessment for learning practice has not enjoyed the same institutional recognition in Japan as in other Asian countries (Watanabe, 2004). This would have clear implications in my own classroom, given that we were also introducing students to new concepts in a genre-based approach.

In Japan, summative assessment procedures such as university entrance exams, or TOEIC, remain the primary recognized measurement of student achievement (Cohen and Spillane, 1999; Mulvey, 2010; Watanabe, 2004). As discussed in Chapter 2, attempts by MEXT to address this situation in 2004 (Takayama, 2008), and introduce learner-centred methodologies, were abandoned as “misguided”. This does not, however, mean that there is no future for these assessment procedures in Japan beyond individual teachers. Takayama (2008) points out the homogenizing effect of the PISA rankings, and the strong regional competitiveness Japan holds with high-ranking PISA nations such as Singapore, Shanghai and Hong Kong. These regions have adopted policy-supported assessment for learning procedures, and have performed well in PISA rankings; consequently, concerns of homogeneity and regional competitiveness, which Takayama describes as highly influential on educational policy, may compel Japan to follow suit. The Central Education Council

(1999) claims that changes to the entrance examination system have been under way for some time. Mulvey (2010) predicts that changes are inevitable due to rapidly falling admission rates for universities, which make the entrance examinations redundant.

Adoption of assessment for learning practices does not necessarily entail abandoning traditional Japanese assessment procedures, or creating a dichotomy between summative and formative assessment; rather, it requires integrating them with new methods.

Black (2009) argues that assessment for learning practice can enable the formative use of summative assessment tasks by treating them as an occasion for formative feedback. This can be done via peer- or self-assessment activities that require students to think about the purposes of the work to be tested, or mark each other's test responses to focus attention on criteria of quality (Blanche and Merino, 1989). Kennedy et al. (2006) argue that the polarization of formative and summative assessment is not useful, and that we should look at summative assessment as productive learning opportunities. Davison (2008) prescribes summative assessment as an integral part of assessment for learning in the classroom, provided that results are used formatively to guide future learning and syllabus design. He promotes summative tests at different stages of a syllabus, from a level focused on criteria that help students decide what to do next, conducted by students and peers themselves, to system-wide published scales, standards and formal tests. If implemented correctly, therefore, assessment for learning could provide opportunities for an assessment methodology that complements existing practices in Japan, rather than offering an alternative or contrasting view.

6.2.2 Issues in assessment for learning

Davison and Leung (2009), Black (2009) and Cumming (2009) point out that despite the widespread acceptance of assessment for learning in various forms, a number of issues remain:

- Much variability in how assessment for learning is implemented
- Very little research in the field of ESL/EFL
- Lack of information about the impact of assessment for learning
- Mechanistic criterion-based approaches that undermine teachers' assessment processes

Their final point raised particular resonance with my own anxiety, that the criteria I had created for assessment were becoming far too prescribed to promote casual conversation. Davison and Leung (2009) call for a more public and mainstream discussion of implementing assessment for learning. They do, however, argue that assessment for learning has powerful potential to improve learning and teaching, and that issues raised by this form of assessment are of central interest to English teaching. It seems relevant, therefore, to examine ways in which a genre-based approach assists assessment for learning practices (and vice versa).

Self-assessment allows students to develop the capacity to become life-long learners (Boud and Falchikov, 2007) and to appreciate their role in learning. For example, McDonald and Boud (2003) claimed in their study of self-assessment in Australian high schools that the vast majority of students found the experience empowering; it gave them greater independence; improved their analytical and critical skills and general studying habits; and allowed them to prepare more effectively for their final exams.

Some commentators have raised concerns in regard to self-assessment. Boud (1999), for example, distinguishes between self-assessment and formal assessment, and proposes that

self-assessment should primarily be part of the learning process rather than a substitute for other types of assessment. Teachers and learners also may question the value of self-assessment: the reliability of such assessment is often not trusted, and its validity doubted (Noonan and Duncan, 2005). These issues need to be taken into account, particularly when the substantial amount of time and effort required to produce meaningful self-assessment procedures is considered.

6.2.3 Integration of assessment for learning into the syllabus

Table 6.1 summarizes ways in which I attempted to integrate assessment as suggested by The Assessment Reform Group (1999). The right-hand side of the table highlights criteria suggested by The Assessment Reform Group (1999), and the left-hand side indicates strategies planned for integration into the syllabus for Cycle Two:

Table 6.1 Summary of the integration of assessment for learning into the syllabus

Shared learning goals	Based upon the deconstruction of model texts, analytic rubrics would be created jointly by teachers and students.
Recognizable standards of achievement	Model texts would illustrate standards of achievement by using Japanese English-speaking peers (see section 6.3).
Step-by-step feedback to inform students how to improve	Formative use of summative assessment. Assessment criteria directly linked to syllabus content. Weekly reference to rubrics and criteria.
Student and teacher reviews of assessment data	Feedback following summative assessments. Meetings following final assessment between teacher and student, to discuss differences or similarities between awarded teacher and student assessments.
Self-assessments	Students self-assessed their performance after each class and after the final assessment
Non-mechanistic criterion-based approach	Analytic rubrics to be jointly created by student and teacher. Student feedback to teacher at the end of each class.
Promotion of independent learning	Self-assessment procedures. Explicit goals for students on a weekly and semester-long basis.
Opportunities for student reflection	Weekly reflective journals. Feedback to teacher. Self-assessment procedures.

Following Cycle One, it was clear that students required greater access to the criteria used in syllabus design and assessment. At first, to avoid a mechanized approach to criteria selection, a holistic rubric was developed, designed according to the level descriptors from the Common European Framework (see Appendix 16). At first, I translated potential holistic rubrics into Japanese, to reduce the need to process complicated meta-language; however, this holistic approach reduced opportunities to jointly construct criteria with the students, as advised by The Assessment Reform Group (1999). It seemed clear that to fulfil the expectations of an assessment for learning approach as described in the literature, and ensure

that goals and expectations were clear to students, a jointly constructed analytic rubric was necessary.

Table 6.1 illustrates an attempt to address problems of a mechanistic criterion-based approach as highlighted by Davison and Leung (2009), by examining model texts with students during the deconstruction phase of the teaching and learning cycle, with teachers and students using this deconstruction to jointly prepare criteria for assessment. This meant that individual criteria for final assessment could often vary from class to class, and even change as the semester progressed. An example of an analytic rubric generated by teacher and students can be seen in Appendix 15. It was hoped these analytic rubrics would allow students to formulate clear goals of achievement that were recognizable and obtainable by the students themselves. It gave students some ownership of the assessment process, and was intended to reduce feelings of anxiety and frustration.

To further address the criteria and issues associated with assessment for learning (The Assessment Reform Group, 2009; Davison and Leung, 2009; Black, 2009; Cumming, 2005, 2009), numerous decisions were made based on reference to the literature. This led to innovations in the syllabus in a principled way, based on the criteria of The Assessment Reform Group for conducting assessment for learning. New syllabus components, therefore, ensured that students were encouraged to reflect on weekly classes in student journals, and also by self-assessing their understanding of key criteria or weekly activities. Students were also encouraged via classroom documents to provide feedback to the teacher about criteria or activities they found difficult or troubling. Students could also reach a consensus on whether specific criteria needed to be changed in the rubric, or descriptions amended (although it was noted in my teacher journal that such requests were never made).

Self-assessment and reflective journals also allowed students to reflect on what they had studied and what steps they needed to take to improve. Opportunities were given via

homework for study outside of class, and students were encouraged to study independently, in order to address any problems they felt they had with any of the ideas or criteria we discussed in class.

Immediately after the final assessment, students also made their own self-assessments, together with written feedback to the teacher on their performance. These scores and comments were collected and compared with the teacher's scores and comments. Student and teacher then had one final conversation together, to negotiate a final score for the speaking assessment based on these two sets of grades. In this way, the summative assessments became formative, and gave students concrete goals for how they should improve. Such a radical overhaul of the syllabus was not a simple procedure, however, and a number of issues arose. I was concerned about the volume of work needed to create lesson plans, but also the amount of explicit teaching that needed to be conducted in class. Together with introducing a new genre-based approach, students were also being asked to adopt new ideas about assessment, and to understand assessment rubrics. This greatly reduced time in class for actual speaking activities, and this concern is raised in the teacher journal at numerous times:

Today I felt that I achieved the objectives of the class, but it cannot be ignored that for the past two lessons the majority of the class time has been teacher-focused

Ultimately, it was decided that the negatives of preparation time and explicit teacher-focused classroom instruction were outweighed by the benefits afforded to the students. However, this is an important point to emphasize when generalizing the findings, as the volume of work may not be viable in some classrooms.

6.3 The Model Text

As discussed in section 5.3.2 and section 5.5, during Cycle One, idealized and transcribed models of speaking proved problematic. The introduction of a native speaker model introduced additional problems. In response to these observations, this section describes the adoption of new model text types in the syllabus design. The main issues with model text type included students relying too strongly on written output to inform spoken output, and text types providing students with learning goals that were unclear, or not relevant to their goals and proficiency levels. In the literature, Nunan (2002), Van Lier (1996), Herron (1991) and Vandergrift (2004) claim it is vital that authentic texts give students access to language where they understand only part of what is being said, to provide an experience of language that is much closer to real life. Nunan suggests two key aspects of authentic texts that are important for language learning, namely naturalness and real-life experiences. However, he qualifies his insistence on authentic materials by stating that students should be told in advance that they do not need to understand everything. Nunan also recommends that students transcribe authentic texts.

Based on these ideas, it seemed reasonable that I should maintain the native speaker model in the syllabus as a model text. In the planning phase of Cycle Two, I made a transcription of an authentic audio recount text between two native speakers. However, I decided that transcriptions of the audio would be avoided for two main reasons, one of which is illustrated in the teaching journal during the planning stage of Cycle Two:

Transcriptions of the dialogues were impossible as they were not suitable for student consumption. They required a hand-out of 3 to 4 pages of A4 and the hesitations, recasts, interruptions and overlaps that might be easy to ignore when listening are bewildering when written on paper. It will also be very difficult for students to sift through all of this language and select the pertinent elements that will help them construct their own conversations.

Introducing low-proficiency students to four pages of language would have been detrimental, in terms of motivation and focusing on key learning goals during the class. Another reason why transcripts were avoided was to exclude any opportunities for written monologues, which the students felt most comfortable in using. Therefore, I attempted to make a simplified transcription of the native speaker authentic text. The majority of vocabulary, idioms and structures that were new to students were removed. The aim of this new form of model text was to introduce the idea of turn-taking in conversation, and the necessity of spontaneity when constructing longer utterances with a partner. The new model and transcript also would also allow classroom discussion about the differences between spoken and written language as an explicit element of instruction. However, transcription and simplification of dialogues also meant that language input was primarily focused on reading rather than listening. The teaching journal reflects the difficulty of constructing listening activities when using simplified transcripts:

Today I intended to record another teacher and myself speaking the prepared transcripts so that students could listen in class and answer comprehension questions before I introduced the written version of the transcript, but the conversation sounded so unnatural and textbook-like I realized that this was as much a problem as the written model dialogues.

Simplification of model texts proved to be a major problem in equipping students with the skills necessary to process and decode casual conversation outside the classroom and in real situations: my teacher journal echoed concerns in Cycle One that students were merely being prepared to conduct short spoken utterances with a sympathetic listener in classroom situations, and to rely on memorized written dialogues. Simplification of language via transcription also limited opportunities to introduce authentic listening activities, as the simplified transcripts differed from the recording of a real conversation.

During further reflection on Cycle One in the planning stage of Cycle Two, it seemed reasonable to set learning goals based on a notion of mutual intelligibility, rather than a native speaker model. This might address students' feelings of anxiety and frustration in not being able to meet perceived learning goals, and also allow the creation of speaking models that make standards of pronunciation, fluency and turn-taking clear and achievable to students. Derwing and Munro (2005) and Jenkins (2006) argue that there is no universally agreed definition of what constitutes "mutual intelligibility". Smith and Nelson (1985) attempt to define intelligibility in three parts: the ability of the listener to recognize individual words or utterances; the listener's ability to understand the meaning of the word or utterance in its given context; and the ability of the listener to understand the speaker's intentions behind the word or utterance. Classroom materials attempted to utilize this definition as a concept of intelligibility that would be shared with students in an attempt to quantify issues of pronunciation and fluency, and achieve concrete learning goals relevant to my students' needs. With permission from participating students at a range of proficiencies, I collated a range of Japanese English-speakers from Cycle One to use as model texts. Thus, students were provided with examples of Japanese English-speakers from their own institution who fulfilled the criteria for a range of scores on their own spoken assessments, thus instilling the idea that "native-speaker" speech was not a necessity.

I felt that to assist the discussion of mutual intelligibility, and to narrow the gap between student and teacher expectations, it is essential to illustrate via modelling exactly what is expected of students: this is enhanced by providing models that are directly relevant to their own needs and expectations. The benefits of using student models as examples of recount texts was reflected in the data at various points, and will be discussed in detail in Chapter 7. Bearing in mind the arguments of Nunan (2002), Van Lier (1996), Herron (1991) and Vandergrift (2004), I also planned to introduce the native speaker audio models to students, but to simplify the task. Rather than deconstruct these models, students would

answer simple comprehension questions that would not require full comprehension of the text. A fuller description of the model text speakers can be seen in Table 6.2.

6.3.1 Subtitled videos

The integration of grammar and vocabulary instruction still remained a key aspect of the syllabus as part of a genre-based approach, and this raised issues in a lesson plan with no written transcripts. In my teaching journal I noted:

While planning potential grammar points without the transcripts, they feel somewhat decontextualized.

For this reason, I decided that subtitles would be added to the video recordings of the model students. The video recordings could then be paused, with single sentences analysed for their grammatical characteristics, but within the context of an entire conversation; this would also avoid students' compulsion to write and memorize, instead of engage in the micro-aspects of conversation. Such a video is illustrated in Figure 6.1:



Figure 6.1 Student-based model text with subtitles

Videos with subtitles were intended to address the problem of students relying on written instead of spoken texts during the deconstruction phase of the teaching and learning cycle, whilst also allowing a focus on grammar and vocabulary when appropriate. I could also pause the video at intervals to focus on difficult vocabulary, and phenomena such as recasts, interruptions and errors, and turn-taking, which were problems highlighted in Cycle One.

6.3.2 Storyboarding

It was also necessary to address the issue of giving students the framework for creating their own longer utterances, while eliminating the habit of writing and memorizing monologues. In the literature, Massironi (2001) suggests that graphic representations of ideas hold huge potential for storing information and then transferring that information to another person. Massironi (2001) also proposes that the pictorial representation of ideas facilitates greater access to a student's verbal system, thus assisting language learning and acquisition. It seemed possible in this syllabus that storyboarding a recount text would allow students to deconstruct texts, and also construct their own spoken recount text in subsequent lessons without the need for writing. Appendix 17 shows an example of a student's storyboard during the deconstruction of a recount text shown in a subtitled video.

6.3.3 Turn-taking

It was hoped that changes in assessment, and in the modelling, presentation and deconstruction of model texts, as described in sections 6.2 and 6.3, would assist students with the difficult problem of turn-taking and the micro-aspects of conversation. Additional conversation strategies were also introduced into the syllabus, and these worksheets can be seen in Appendix 18. Micro-aspects of conversation were emphasized during text deconstruction, and I ensured that I focused feedback on these aspects during classroom practice. Worksheets shown in Appendix 18 focused on the use of "WH" questions to ask for

additional information during conversations, and a long list of example sentences was shared. However, as illustrated in the observation and reflection phases of Cycle Two, micro-aspects of conversation remained problematic. Appendix 23, discussed in greater detail in section 7.8, illustrates that assessment procedures and the issues of Tenor between student and teacher contributed to issues of turn-taking; and additional consultation of the literature was also necessary in preparation for the planning phase of a third cycle of action research, beyond the scope of the present thesis.

6.3.4 Collaboration

Greenwood and Levin (2005) stress the cooperative nature of action research. According to Panitz (1996), collaboration involves learning about and engaging with the abilities and contributions of one's peers, and it highlights the importance of consensus-building and cooperation in research; particularly research that seeks to inspire social change through the paradigm of critical theory. Nunan (1992) states that collaboration is key to a number of disparate philosophical viewpoints and research traditions; and that in language education, teachers and researchers should collaborate to experiment with different ways of teaching, learning and researching, so that members can learn from each other equitably, and through cooperation rather than competition. Nunan (1992) argues that action research can help language teachers, learners and researchers make a collaborative reflection that will lead to an overall understanding of language classrooms. Burns (2010) also encourages the involvement of others in action research projects, claiming that it is a preferable way of conducting such research, as it opens avenues for deepening insights during observation and reflection. By contrast, Edge and Richards (1993) argues that collaborators can prove detrimental to insightful thinking, by offering limited advice and suggestions without a genuine cooperative understanding. It is important, therefore, to approach data generated by collaborating teachers objectively. Three additional teachers volunteered to collaborate in my

proposed syllabus intervention. Data collected via these sources are described in greater detail in section 6.5.1.

6.3.5 Self-assessment

A recurrent argument in the literature on assessment is that teacher evaluation should draw on multiple sources of information (Grissom and Youngs, 2016). Introducing collaborative teacher assessors to the collection of data creates increased opportunities for triangulation, examining reliability, and fit statistics. Another source of potential evaluation data is students themselves. An integral component of assessment for learning practices is self-assessment, which is concerned with revision and improvement. It enables students to independently assess their own and other students' progress with confidence, rather than always relying on teacher judgment (Benson, 2013). When students self-assess, they are actively involved in the learning process, and their independence and motivation are improved. A great deal of research has illustrated the educational benefits of self-assessment, with calls that self-assessment should be included in every major assignment a student undertakes (Wiggins, 1993).

Boud and Holmes (1981) highlight these benefits as threefold: firstly, improving feedback to the students by increasing its provision during the learning process in a more timely manner; secondly, providing additional reinforcement of teaching aims and objectives, with the criteria and materials undergoing consideration several times more than usual; and finally, self-assessment is one of the few teaching strategies that allows a reduction of teacher workload, in conjunction with a corresponding increase in educational benefits. McDonald and Boud (2003) point out that students themselves experienced the benefits of self-assessment: the vast majority found it empowering; that it gave them greater independence; improved their analytical and critical skills and general studying habits; and allowed them to prepare more effectively for their final exams. Students who had undergone self-assessment also outperformed similar students who had not experienced such assessment procedures. An

additional benefit was that it allowed students to develop the capacity to become life-long learners (Boud and Falchikov, 2007).

Nevertheless, some teachers have raised concerns in regard to self-assessment. Boud (1999) explains that care is needed if self-assessment also becomes a form of formal assessment. He goes on to propose that this form of assessment should focus primarily on being part of the learning process rather than as a substitute for other types of assessment. Similarly, for many teachers, the value of self-assessment is often questioned; the reliability of such assessment is often not trusted, and its validity questioned (Noonan and Duncan, 2005). Therefore, it is often absent from many syllabi and classrooms, as these issues may appear to many to negate any potential benefits, particularly when the great deal of time and effort required to produce meaningful self-assessment procedures is considered.

However, to regard self-assessment purely in terms of the learning process without maintaining a focus on actual grades, or to discard it entirely on the grounds of validity and reliability issues, ignores important potential benefits. By examining the correlations or differences between student and teacher grades, it is possible to increase understanding of issues with syllabus design or teaching techniques, and to emphasize opportunities for effective feedback on an individual level. Exploring the issues of validity and reliability in a positive manner, and exploring differences in student and teacher grades, could impart valuable knowledge about students' perceptions of their spoken performances. The first of these benefits is the potential to highlight opportunities for targeted feedback.

Sadler (1989) and Butler and Winne (1995) have highlighted the need for strategies to provide higher-quality feedback to students, and to encourage students to develop less reliance on teachers for their learning. McDonald and Boud (2003) have argued that formative assessment does not provide sufficient opportunities for feedback, beyond that which is the sole province of the teacher. They therefore argue that students themselves have an essential role to play in shouldering the responsibility for assessment. According to Black

and William (1998), the gap must be closed between students' lack of knowledge, as revealed by feedback, and the teacher's aims and objectives. Students following a prescribed dictation of assessment by the teacher without understanding its purpose are unable to learn. In this regard, Black and William (1998) describe self-assessment as an essential component of learning. Self-assessment provides an opportunity for highly detailed feedback, closing the "gap" in knowledge that may exist between teacher and student expectations, which was highlighted in Cycle One.

Of course, more accurate self-assessment between student and teacher would allow feedback to be a continuous process during the learning experience, rather than a summative process after the assessment event. If students are well aware of the criteria before assessment and these criteria are unambiguous, with classroom aims and objectives made clear, students' self-assessment accuracy could be improved. Boud (2000), like Ross (2006), describes the necessity for a criterion-based framework, so that learners assessing themselves are aware that their achievements are the result of meeting an acceptable standard, as opposed to simply doing better than other students. Brown (2004, p. 19) also highlights the need for an adequacy of construct definition in tests of second-language communicative competence. Brown explains that communicative competence is often "an abstraction that is rarely defined with any precision in terms of actual test performance". Citing Bachman (1990, p. 50), he quotes that in order to "maximize the reliability of test scores and the validity of test use, we should ... provide clear and unambiguous theoretical definitions of the abilities we want to measure".

Some research has explored the accuracy of student self-assessment. According to Blanche and Merino (1989), although errors do occur, considerable research has shown that learners can be accurate in assessing their own abilities. Blanche and Merino claimed that the key to consistent overall agreement between self-assessments and rating is that the skills to be assessed in foreign languages should be clear and detailed, echoing the views of Boud and

Falchikov (2006) and Ross (2006). In a range of different educational contexts, Black and William (1998) found correlations between self-assessments and teachers' assessments of 0.71; and this success was mirrored by Merrett and Wheldall (1992), Griffiths and Davies (1993), Powell and Makin (1994) and Meyer and Woodruff (1997). Boud and Falchikov (1989) reviewed 48 studies reporting self-teacher assessment agreement, and found that self-assessments agreed with teachers' ratings in most cases. They did, however, highlight numerous shortcomings regarding the quality of many of the studies, and found that there was extensive variation in what constituted student and teacher agreement. Other research has highlighted issues with the accuracy of self-assessment. Ross (2006) found mixed results for self-teacher agreement in studies of second-language learning, with wide variation between studies, and a mean correlation of .64. It has been argued that in the field of language learning, there has been a failure to examine student self-assessment in relation to criterion-referenced assessment, whereby the criteria used by teachers and students are adequately defined (Ross, 2006).

The aim of this study, therefore, is to generate clearly defined criteria based upon theoretical understandings of language performance. These criteria are to be made clear and quantifiable to students. Feez and Joyce (1998, p. 2) argue that the foundations for specifying abilities and outcomes in a genre-based approach are laid in the course design: "learners cannot learn everything about a subject at once nor can they learn effectively from a random collection of unrelated items. Therefore the teacher needs to develop a systematic plan for course content which will lead to desired outcomes". They explain that this systematic plan is the "syllabus", which provides a map for the teacher and the learners, and is based upon explicit objectives: "it is a public document, usually prepared by teachers and negotiated with learners. It specifies what is to be taught in any particular course of study". Self-assessment procedures were integrated during the planning phase of Cycle Two by introducing a self-assessment component to the final assessment, as well as weekly self-assessments of student

performance during each class, to be recorded in their student journals, as described in section 4.6.3.

6.4 Ethical Considerations

As stated in section 4.4, the PhD Student Research Ethics Approval Form (REC1) was submitted twice: this included the submission for data collection in Cycle Two, taking into account additional student- and teacher-generated data.

6.5 Action: Data Collection

Data collection followed the same pattern as Cycle One, as illustrated in Table 4.2; but with additional data collection points for Cycle Two, based on the inclusion of quantitative student self-assessments, qualitative student self-assessment reflections; and data from collaborating teacher researchers via teacher journals, weekly research meetings and quantitative data collection of assessments. As opposed to the 27 students of Cycle One, Cycle Two reflects data collected from 240 students. This section will explain these additional data collections in more detail.

6.5.1 Data from collaborative researchers

During a staff meeting, I presented my ongoing research, and additional teachers who were interested in the syllabus asked if they could use it in their own classes. I provided them with the syllabus, lesson plans, worksheets and assessment sheets. However, teachers were free to use the materials as they felt appropriate. Three teachers agreed to provide feedback on these materials at the end of the semester, and during “Action Research Group” weekly meetings that we already held in our institution: this assisted teachers with carrying out smaller-scale interventions that corresponded to their own research interests. Time constraints were a major consideration in data collection from other teachers. An adapted list of reflective questions was provided to teachers, to reduce the administrative burden on participants (Appendix 19).

Open-ended questions were employed due to the small sample size, and to avoid questions that were not necessarily representative of participants' beliefs, such as might be found in closed questions (Farrall et al., 1997). Farrall et al. argue that open-ended questions are far more useful for discovering teachers' beliefs. Werner and Schoepfle (1987) also refer to them as questions that presume, i.e. presuming that participants are aware of an abstract concept, to which they may reply with an invented response. Questions were also avoided that made participants uncomfortable, such as personalized or embarrassing questions. For example, I did not simply ask teachers, "Did you like my lesson plans?". I also avoided using questions that presented two areas of enquiry as one; using words that were loaded with stereotypes; or prestige questions that led to a certain answer, such as "I dislike the current way of teaching speaking, do you?".

In section 1.2, it was described how gaps were perceived in the current curriculum of the institution in which this research takes place: this led to the English department brainstorming ideas for new speaking syllabi that would assist students to develop greater fluency in spoken performance; and in the future, a potential university-wide curriculum would address perceived inadequacies in English language instruction and learning. My research was seen by other members as a chance to develop a speaking syllabus and assessment instrument that could measure students' spoken performance. At this time, any language assessment at the institution was purely receptive in nature. The research collaborators and I thus planned to use my syllabus to initiate discussion on how we could assess speaking abilities at our institution in the future. From our respective classes, following the syllabus intervention, 24 students were selected for an assessment exercise: six students were chosen by each teacher. A video of each student had been recorded during the final speaking assessment portion of the syllabus, with the permission of the student. These videos were then rated again by each of the four teachers during the assessment exercise. Prior to the assessment exercise, criteria for the assessment were negotiated in a meeting of

the four teachers. These assessment criteria would be used by all four teachers, and followed those highlighted in Table 5.1 of Cycle One, but with four additional criteria requested during negotiations between the four teachers. These additional criteria included a “Vocabulary Bank” that aimed to measure the frequency of vocabulary used, based on vocabulary banks that are distributed to students as part of the overall curriculum beyond the syllabus in this thesis. Another criterion included was “Dictation”. One teacher felt that if students were unable to construct a recount text during the final assessment and sat in silence, they should be afforded an opportunity to dictate a recount text provided by the teacher, to gain at least one criterion score. If students did not need the prepared text, then they would score full points on that criterion. Another teacher argued for the inclusion of a “Clear Voice” criterion, based upon how audible the students were in conversation. A final additional criterion negotiated amongst teachers was “New Text”, whereby a student would be able to generate a second recount text based on additional questions from the teacher. Criteria were scored on a scale of 1–5. Rasch measurement was used for item, person and rater analysis, based on scores awarded to each of the 24 students by all four teachers, using Facets software (Linacre, 2007a). Appendix 20 shows a copy of the rubric negotiated by teachers, with fuller descriptors of the criteria and scoring scale. The adoption of Facets (Linacre, 2007a) for quantitative analysis is explained in greater detail in section 6.6.

6.5.2 Student self-assessment and reflection

The criteria for self-assessment in this study were developed via negotiation with students during a textual deconstruction of the modelled recount texts. Structural and grammatical components of recount texts were highlighted and exemplified, together with discussions of what students thought made a “good speaker”. These negotiated criteria broadly correlated with the Cycle One criteria exemplified in Table 5.1 in Cycle One, but underwent small changes depending on the results of negotiation in individual classes. A range of speakers were used for the model texts described in Table 6.2, so that students could begin to construct

a scale of measurement for their own speaking in terms of pronunciation and fluency. The model texts were chosen from students who undertook the assessment in Cycle One, and were of a range of abilities, as indicated by the variable map in Figure 5.3. The profiles of the model text speakers are outlined in Table 6.2. These speakers became the basis of a 5-point Likert scale based upon the criteria negotiated with students in class.

Table 6.2 Profiles of persons used as model texts and to inform the 5-point Likert scale for pronunciation and fluency

Model Text:	Speaker Profiles
1	Japanese speaker. Female, 26. Student. TOEFL iBT score: 110
2	Japanese speaker. Female, 22. Student. TOEIC Score 580
3	Japanese speaker. Male, 19. Student. TOEIC Bridge Score: 140
4	Japanese speaker. Male, 19. Student. TOEIC Bridge Score: 125
5	Japanese speaker. Male, 19. Student. TOEIC Bridge Score: 115

Table 6.2 shows how model texts were based on Japanese students with varying degrees of spoken English proficiency. By watching model texts produced by their peers, it was hoped that students could immediately identify the levels of proficiency required in a range of negotiated criteria, but particularly the criteria of pronunciation and fluency, which had been implicated in concerns about a gap between teacher and student expectations in Cycle One (section 5.5). Appendix 21 shows an assessment rubric negotiated with students at the start of the syllabus intervention after a recount text deconstruction, with the same criteria as that dictated to students during Cycle One. Students assessed their pronunciation and fluency based on the model texts, with Model 1 correlating to 5 points on the Likert scale and Model 5 correlating to 1 point on a 5-point Likert scale. Students could thus judge their own pronunciation and fluency abilities by comparing them to their peers and creating a notion of mutual intelligibility that was achievable by students, as opposed to an unachievable notion of the “native speaker”. Appendix 22 shows a completed self-assessment sheet that was completed by a student immediately following the spoken assessment at the end of the semester. Appendix 22 also shows that during self-assessment, students also recorded their

immediate reflections on their performance in the spoken assessment, in English and Japanese. English comments were used as a component of qualitative data collection.

6.6 Preparing for Data Analysis and Presentation

Data analysis and presentation for Cycle Two followed the same framework as Cycle One, as outlined in sections 5.2, 5.3 and 5.4. However, one major change was the replacement of Winsteps (Linacre, 2007b) software with Facets (Linacre, 2007a) software. Facets (Linacre, 2007a) would allow the inclusion of additional facets during Rasch analysis: namely, the ratings of collaborative teachers and student self-assessments. Self-assessment analysis would allow exploration of the differences between teacher and student perceptions of their speaking performance on the criteria level, thus highlighting any gaps in understanding between student and teacher expectations, or a disparity in lesson aims and goals; it would also show what skills students were developing. An analysis of rater differences between teachers would also signify differences in the understanding of syllabus goals, and would provide important observations on the generalizability of a genre-based approach.

Linacre (2007b) recommends the use of Winsteps before embarking on Facets, due to the former software being a conceptually and operationally simpler Rasch-measurement program. Linacre (2007a) describes Facets as a many-facet Rasch measurement, a computer program for the construction of linear measures. In the case of the data in this thesis, students, criteria and raters are facets. For each facet, Facets provides a measure in logits, its standard error and fit statistics. The fit statistics enable diagnosis of aberrant observations and idiosyncratic items. Facets is also able to quantify discrepant interactions between elements of different facets. In this way, a judge's bias regarding one student, or a student's bias against a group of criteria can be identified, and its size and statistical significance estimated. Linacre (2007a) argues that Facets is ideally suited for judged performances, including speaking.

6.7 Summary

In this chapter, the adaptation of a genre-based syllabus was presented, which included major changes to assessment procedures and model texts based on the systematic observation of and reflection on data in Cycle One. New data collection procedures and analysis were outlined. Chapter 7 will describe the observational and reflective phases of Cycle Two.

Chapter 7: Cycle Two: Observation and Reflection

7.1 Introduction

In this chapter, data analysis undertaken during Cycle Two will be explained in detail, with reference to data analysis that was undertaken in Cycle One as a means of triangulating data through time. In this chapter, the ways in which spoken and written qualitative data were categorized and coded will be explained, together with the presentation of quantitative data and other processes of data triangulation. This chapter will also describe when and how reflection took place in the action research cycle, and how and why action research cycles will go on developing in future cycles beyond the scope of this thesis. The chapter will begin in section 7.2 with a description of qualitative data analysis.

7.2 Student Emotions

As with Cycle One (section 5.3.1), student emotion was a central theme of the data in Cycle 2. Figure 7.1 shows that students once again felt nervous and frustrated during the syllabus intervention; however, these feelings were not as frequent in the data as in Cycle One. Evidence of enjoyment and satisfaction was found to be much more frequent in Cycle Two. Figure 7.1 also shows data sources in Cycle Two that comprise the comments of collaborative teachers.

Time period	Source	Student Emotions			
		Nervousness	Frustration	Enjoyment	Satisfaction
Cycle Two Weekly	Student Journal		<i>I can't keep conversation</i>	<i>It was a lot of fun speaking English</i>	<i>I did it!</i>
	Teacher Journal	<i>Seeing the student model eased their nervousness about the task ahead</i>		<i>They really enjoyed the opportunity to speak in class</i>	<i>He was pleased to learn this new way of greeting</i>
Cycle Two Post-Assessment	Teacher reflective comments	<i>Students became less nervous as the interview progressed</i>	<i>Students lack strategies for coping with gaps in vocabulary</i>	<i>Some students seemed genuinely pleased</i>	<i>Students showed satisfaction in answering questions</i>
	Anony-mous student feedback	<i>I thought I would die, but I didn't die</i>	<i>I was not smooth</i>	<i>I want to do it again one more time</i>	<i>Finally I could speak and understand it</i>
End of the semester	Teacher 1	<i>Students were scared to talk one-on-one in the test</i>	<i>Students had a negative attitude to testing and talking to the teacher</i>	<i>They enjoyed telling each other their stories</i>	<i>They were surprised they could do it</i>
	Teacher 2	<i>The test made them very nervous</i>		<i>They really enjoyed talking to each other</i>	<i>They felt like they'd accomplished something</i>
	Teacher 3			<i>It gave them life skills I think, they learned how to enjoy a conversation with friends</i>	<i>They could put language into practice</i>

Figure 7.1 Matrix display to examine patterns in student emotions

Time triangulation of the data showed that instances of nervousness were far more apparent during the pilot study phase of the data collection. Further analysis of the data

suggests that this was due to a change in the model texts used in the syllabus. I noted in my research journal how the use of a student as a model of speaking eased students' nervousness about speaking English. For example:

After watching the student video they were far more relaxed about the task. They were engaged with the video, laughing at jokes and commenting on how good the student was at English. In comparison to the (native speaker) transcript in the pilot study the response was far more optimistic, one student commented in Japanese: “kantan” (easy)

Researcher triangulation emphasizes this finding. Other teachers also mentioned the use of the student model, and Teacher 2 remarked:

After watching the video they knew exactly what they had to do

Cycle One had used a native speaker dialogue as a model speaking text. This created a far more nervousness-inducing prospect for students when it was their turn to speak. In Cycle One, the written transcripts I had adapted from an “authentic” conversation between me and another teacher set a language model that was beyond the proficiency of the students. The introduction of a student model gave them a tangible target for conversational goals, and reduced nervousness. The aim of the student model in the design of the syllabus was not necessarily to reduce anxiety, but it had this effect.

The nervousness highlighted in the data when the student interacted with the teacher, and the change in the frequency of nervousness reported when the model text was changed to a student Japanese speaker, appears related to whether assessment was involved. Assessment created a formal situation that inspired nervousness in students. Even during casual conversation with the teacher, when there was no formal assessment, nervousness still occurred. This was perhaps created partly by the Tenor and Mode of the text generation (sections 2.6.2 and 2.6.3), with immediate feedback from an interlocutor in an unequal

continua of power, contact and affect (Figure 2.2). When feedback came from a teacher, students became nervous. Even with grades being awarded through self-assessment, this nervousness remained. When students were provided with a model text that exemplified a Japanese student as a target for learning, nervousness was reduced. This suggests that when students are taking part in a casual conversation, they are assessing themselves in relation to their interlocutor's level of English. This discussion of the model text and assessment is further examined in section 7.4. An associated negative emotional experience that arose from the data was frustration; this theme was closely connected with nervousness. Turn-taking aspects of the syllabus continued to promote frustration in students. Teacher 3 mentioned the lack of fluency in student responses during the micro-aspects of conversation, adding that:

Students took a very long time to answer questions, they could answer, but sometimes it took so long compared to others.

The teacher sensed that conversation had broken down when students took too long to answer. I noted in my journal that rather than using strategies to continue the conversation, some students would take extended periods of time to ensure they gave the correct answer to whatever question or idea they were confronted with, at the expense of fluency. In a natural conversation this would usually create feelings of discomfort; thus, I noted in my journal my own feelings of frustration that the syllabus was not addressing issues of fluency. Students often commented in feedback with such examples as:

I could understand the questions but I couldn't find an answer

Students spent uncomfortable periods of time trying to find answers instead of applying strategies for continuing the conversation. As well as nervousness, frustration also increased when students engaged in casual conversation with the teacher, as in Cycle One. One reason for this was due to the teacher introducing micro-aspects of discourse to the

conversation, which increased the difficulty level of the speaking interaction. This was reflected in students' self-reflection in their journals. Teachers also expressed frustration with the genre-based syllabus; the impact of the syllabus on teachers is discussed in greater detail in section 7.8.

In Cycle Two, however, there was greater evidence of satisfaction and enjoyment. Students enjoyed telling each other their recount-genre texts, not only to practise English but also to fulfil the role of casual conversation; i.e. to construct an interpersonally motivated interaction, and thus build social relationships with classmates. Two teachers mentioned that they learned a great deal about their own students. This helped them to build a rapport with the students that they had not experienced before: according to Teacher 3,

It helped definitely, build a little bit of a rapport with the students

The data suggests that satisfaction was tied closely to having a purpose for speaking; and that during interactions, the building of social bonds with other students and the teacher was enjoyable.

Satisfaction included data from student feedback that illustrated instances of students evaluating their own skills and emphasizing a desire to improve them through a range of strategies that they formulated for themselves, such as practising listening, studying vocabulary, or not using a dictionary. Students were also setting themselves targets relevant to individual goals we had discussed in class; or beyond this level, to “mastering” English and using it outside class. These comments correspond with ideas of intrinsic motivation (Ryan and Deci, 2000), which were discussed in section 5.3.1 and section 5.5 as potential benefits of the syllabus design in Cycle Two. One student commented anonymously at the end of the syllabus:

I want to let all things that I learned until now live in my real life

Students were becoming interested in learning English beyond the classroom and were showing signs of autonomy in developing those skills, which was one of the key goals in introducing assessment for learning strategies into the syllabus.

7.3 Cross-Referencing Student Emotion with Quantitative Data

Rasch analysis of students, based on self- and teacher assessments conducted during a final assessment, allows the targeting of individuals who may be in need of remedial attention. For example, unexpected results can be triangulated with qualitative data, to explore conclusions of nervousness and frustration, or enjoyment and satisfaction. The fit analysis in Table 7.1 shows a summary of scores that students awarded themselves during peer assessment, and scores that show “misfit” (Linacre, 2007a). These are examples of what Dörnyei (2007) describes as “outlier” or “extreme” responses (p. 272), which seem to run counter to common belief or even contradict it.

Below is a table that represents students’ quantitative data (n=240), showing those who were either inconsistent or more consistent than other students in scoring themselves, to a degree that warranted further qualitative investigation.

Table 7.1 Person Fit statistics.

Person Number	Score	Standard Error	Infit MnSq	Outfit MnSq
134	29	1.14	0.31	0.16
26	28	0.88	0.32	0.26
2	26	0.71	0.39	0.34
28	18	0.65	0.59	0.56
141	15	0.69	2.16	2.03
55	14	0.71	0.54	0.47
132	8	1.00	0.37	0.27
Mean	22.8	0.73	0.98	0.93
Standard D. (all students n=240)	4.7	0.12	0.54	0.73

Fit statistics greater than 1.0, termed “misfit”, indicate more “noise” in the data than expected, while those less than 1.0, termed “overfit”, indicate data that is more consistent than expected. For example, Person 141, with an infit of 2.16, shows 116% more randomness than is expected by the Rasch model. By contrast, Person 134, with an infit of 0.31, shows only 31% of the randomness expected. Following the guidelines of Bond and Fox (2007), mean-squared values of 1.4 and 0.8 were used as limits of acceptable fit in this study, meaning that persons with less than 80% or more than 140% of randomness are investigated more closely.

Students 134, 26, 2 and 28 have low infit, meaning they were very consistent in their grading. Students 134, 26, and 2 rated themselves highly, and the teacher grades show that they were also rated highly by the teacher. Student 28 (all pronouns are figurative) was consistent with the expected responses calculated by the Rasch model in their self-rating, but gave himself a score below the average of the class. Student 28’s self-assessment sheet shows that he simply marked themselves three points for each of the criteria. Student 134, meanwhile, rated himself highly in all of the criteria, which suggests he was very satisfied with his performance in the conversation with the teacher.

Student 141 was extremely inconsistent with his scoring, and awarded himself a low grade. This suggests that he awarded himself low scores even for easy criteria, perhaps even scoring himself higher in an easy criterion than a difficult one. After identifying this student’s self-assessment sheet, it appears that he awarded himself a higher grade for micro-aspects of dialogue than for formulaic aspects of discourse such as greetings. This suggests he may have randomly assigned himself scores for his conversation. After identifying this student’s attendance data, I noticed that he had only attended three out of fifteen classes. It is therefore likely that he had attended the final assessment, but was unprepared for the test, as he had missed the majority of classes.

Students 55 and 132 show very high consistency in their self-assigned grades, but very low grades. Student 132 had awarded himself just one or two points for each of the criteria. After identifying his teacher scores, it was apparent that the teacher had also awarded him low scores for the majority of criteria. This provides an example of the usefulness of Rasch as a diagnostic tool. Students who do not perform as expected in the class are immediately identifiable, and remedial help can be offered to them (Engelhard, 2009).

Student 55 had also awarded himself very low grades for almost all of the criteria, but interestingly the teacher had awarded him high grades (25 out of a possible 30 points). This would suggest perhaps a high level of nervousness or frustration. Fit statistics identify extreme ends of student performance, in terms of either high-performing or low-performing students in the final assessment; however, it also provides an interesting opportunity to explore these students in more detail through their qualitative feedback, as proposed by Engelhard (2009).

During data collection, students had submitted student journals that I could cross-reference with their teacher- and self-assessments. Students had also recorded their own anonymous comments on feedback sheets after the final assessment. Below are extracts from the student journals of student 134, who had been awarded high grades by both himself and the teacher:

Now that I have studied a range of greetings I do not think they are very difficult.

I could explain in detail what I had done in my weekend.

I couldn't recall some words, however I understood everything.

My partner and I could use questions and answers fluently.

I could do everything so smoothly and I was really relaxed in the class.

I could question and response smoothly to my teacher.

I could make a good story.

Using temporal conjunctions I could smoothly explain about my weekend.

I gave good answers to teacher questions I developed the content well.

Student 134 is extremely confident of his ability to conduct a conversation in English, as evidenced by the regular use of “*I could*” to describe mastery of criteria. Student 134 mentions being “*relaxed*” in class, and often uses the adverbs such as “*smoothly*” to describe his conversation. This student did not appear to be nervous or frustrated, and the syllabus seems to have been an enjoyable and satisfying experience for him. The student describes tackling micro-aspects of dialogue easily, and references grammatical aspects of the syllabus that assisted his speaking ability.

Student 132, on the other hand, had been awarded low grades by both himself and the teacher, and some of the underlying reasons for this may be found in the student journal:

I could make a summary in my mind, but I couldn't speak, my voice was small and poor

I couldn't speak so deeply

I had to take a long time to connect sentence

I couldn't do well answering questions

I thought I have to use English more in usual time

Student 132, in contrast to student 134, uses many negative modals in his reflections, which are primarily focused on speaking “*deeply*” and a lack of fluency. This student appears

to have struggled with the micro-aspects of conversation, particularly asking and answering questions in a timely manner. The student remarks that he could make summaries in his mind, but then struggled to produce them coherently. The comments appear primarily to portray a lack of confidence in oral output and a frustration with not being able to put ideas into words. In large classes of students, this kind of identification of students with low confidence gives the teacher some valuable tools in identifying and assisting remedial students, or students who need extra attention.

7.4 The Model Text

Major themes emerged from the data when the types of model text that were used in the syllabus mirrored those in Cycle One. Figure 7.2 shows how, as in Cycle One, the main themes that emerged from the data under the concept of modelling the text were *text authenticity*, i.e. the mode of the text, and *the model speaker*, i.e. who modelled the text. Within these themes, sub-categories included the impact of the new model texts that were introduced in the planning stage of Cycle Two (section 6.3).

Time	Data Source	Text Authenticity		The Model Speaker	
		Transcribed Texts	Spoken Texts	Pronunciation and fluency	Learning goals
Cycle Two	Teacher Journal	<i>Storyboarding has taken away the idea that they should write and memorize</i>	<i>They understood that it's ok to make "mistakes" when speaking</i>	<i>I think that mutual intelligibility is the key to their speaking</i>	<i>The video gave a clear idea of expectations</i>
	Classroom Documents	<i>Subtitles worked well</i>	<i>The model gave spoken strategies for when they didn't know what to say next</i>	<i>The student model is much easier for them to understand</i>	<i>The student model let them know exactly what was expected of them</i>
	Teacher 1	<i>A written version might be a good idea</i>	<i>It was more like a real conversation with the video</i>	<i>They found the student talking really funny</i>	<i>Seeing the student talking gave them confidence</i>
	Teacher 2	<i>The subtitles was a good thing to have</i>	<i>They struggled with listening to you guys speak</i>	<i>They don't need to speak like us</i>	<i>When they saw the video they knew what to do right away</i>
	Teacher 3	<i>I think a written transcript would be a good thing for them to have, to look at</i>	<i>The video supported participation in class</i>	<i>I think more fluency and pronunciation assessment is needed</i>	<i>The video gave focus on what was needed</i>

Figure 7.2 Matrix display showing reactions to the model texts

During the cycles of action research, there was a clear progression of levels of authenticity in the model texts that students interacted with in the classroom. This progression is clearly mapped out in teacher journals and comments, and in classroom documents. Figure 6.2 summarizes this progression:

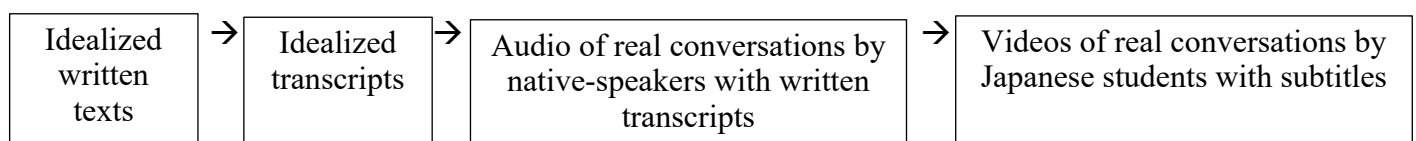


Figure 7.3 Progression of text authenticity

During the planning phase of Cycle Two, I had decided to keep the native-speaking model text, but to simplify any tasks associated with it (section 6.3). As in Cycle One, students struggled with the native speaker text:

I did not understand the particulars of it, so I could not answer the questions

The talking speed is so fast I cannot understand the conversation

Teacher 2 made the comment in their journal:

They struggled with listening to you guys speak

Even with added subtitles and transcripts, the native-speaker text still proved difficult for students, as noted in one of my lesson plans:

although the videos are great for listening, students are struggling to identify appropriate targets for their own output, after listening to the video there and discussing them in class it is clear the students are not as confident about being able to produce such conversations themselves

Even with greatly simplified tasks, authentic tasks still proved to be difficult for low-proficiency students. They could not understand the gist of conversations without breaking down the audio into shorter chunks, which began to de-contextualize the language and create further complications in understanding the overall message.

The impact of the Japanese model speaker texts was very apparent in the data collected via the teaching journal and classroom documents:

Students seem to be thinking about speaking instead of writing for output

They understood that it's ok to make "mistakes" when speaking

Videos with subtitles addressed the problem of students' output being written instead of spoken, whilst also allowing a focus on grammar and vocabulary when appropriate. The teacher could also pause the video at intervals to focus on difficult vocabulary, or phenomena such as recasts, interruptions and errors. Teacher 2 noted:

The subtitles was a good thing to have

Section 6.3.2 presented the rationale for the introduction of storyboarding into the syllabus during the planning phase of Cycle Two, in order to discourage the writing and memorization of written texts for spoken production. Appendix 17 shows an example of a student's storyboard during the deconstruction of a recount text shown on a subtitled video. Appendix 17 shows that when constructing their own recount texts during the joint-construction phase of the teaching and learning cycle, storyboarding allowed students to think about their experiences without having to simultaneously consider how they would relate those experiences in English. This was advantageous to low-proficiency students, who often preferred to revert to memorized English phrases and talk about experiences they were already confident in relating in English; they had practised such phrases numerous times before in textbooks, as they had in Cycle One. By storyboarding, they could formulate ideas for talk, and also plan new grammar and vocabulary to later describe in English; this assisted longer utterances by promoting the use of greater detail in storytelling, and by varying instances of recounted events beyond those included in textbooks. This was exemplified in a note in one of my lesson plans:

Storyboarding allowed them to plan in advance what they wanted to say, and even brainstorm appropriate verbs in the appropriate tense

Storyboarding also allowed a focus on the grammatical aspects of recount texts. Classroom documents show students identifying English verbs and conjunctions for each of

the pictures in their storyboards, which they could later use in conversation (See Appendix 17). In terms of preparing them for authentic casual conversation, students writing an individual English verb next to a picture was preferable to finding instances of whole transcribed monologues in their worksheets.

Storyboarding also allowed students to formulate extended monological spoken ideas in advance, without writing. The benefit of this is that it allowed students to participate in turn-taking strategies without becoming lost or confused in regard to the main monological structure of their talk. By referring back to the images, students were reminded of what they were talking about; this also allowed the conversation to stray from this monological plan at random moments. Storyboarding also allowed a more dynamic type of interaction than the memorized written monologue they had been constructing in Cycle One, as noted in my teacher journal:

Storyboarding has taken away the idea that they should write and memorize

In Cycle One, it was reflected that students misunderstood the goals of the class and thought that they had to be able to speak like me to be “English speakers”. In Cycle Two, I attempted to introduce a notion of mutual intelligibility by using a Japanese-speaking peer as a goal for their speaking development during the course of the semester. I introduced the notion of mutual intelligibility by using their peer as an example of mutually intelligible English spoken by someone like themselves. I explained that my mother, who does not live in Japan and does not speak Japanese, would be able to understand the English of Model 1 (Table 6.2), and that this was a level of English proficiency that they were capable of achieving: a level of mutual intelligibility. The benefits of using Japanese student models as examples of recount texts was reflected in the data from various sources. It was a familiar theme in teacher journals; for instance, Teacher 1 noted that:

The student model let them know exactly what was expected of them

Teacher 2 remarked:

When they saw the video they knew what to do right away

Teacher 3 commented:

The video gave focus on what was needed

Whereas previously, students were expressing frustration with their lack of speaking skills, they were now able to quantify more clearly what level of English proficiency they were expected to attain, and which I defined in the classroom as mutual intelligibility.

7.5 Cross-Referencing Qualitative and Quantitative Data on the Model Text

An analysis of quantitative data strengthens the finding that using Japanese English-speakers as model texts narrowed the gap between student and teacher expectations of their spoken performance. Following the introduction of Japanese English-speakers as models of a recount text, self-assessment procedures allowed a quantitative analysis of how students viewed their own test performance, compared with the teacher's assessment.

Quantitative analysis focused on scores awarded for an individual criterion by both teacher and student, after their final speaking assessments based on a recount text.

Assessment results based on criteria from two classes in 2011 ($n=60$) showed a Pearson Correlation of 0.85 between student and teacher ratings of spoken performance, suggesting strong positive rater agreement on the success or otherwise of items. Figure 7.4 shows that differences in rater leniency and severity between student and teacher were negligible, with the teacher being very slightly more severe a rater than students:

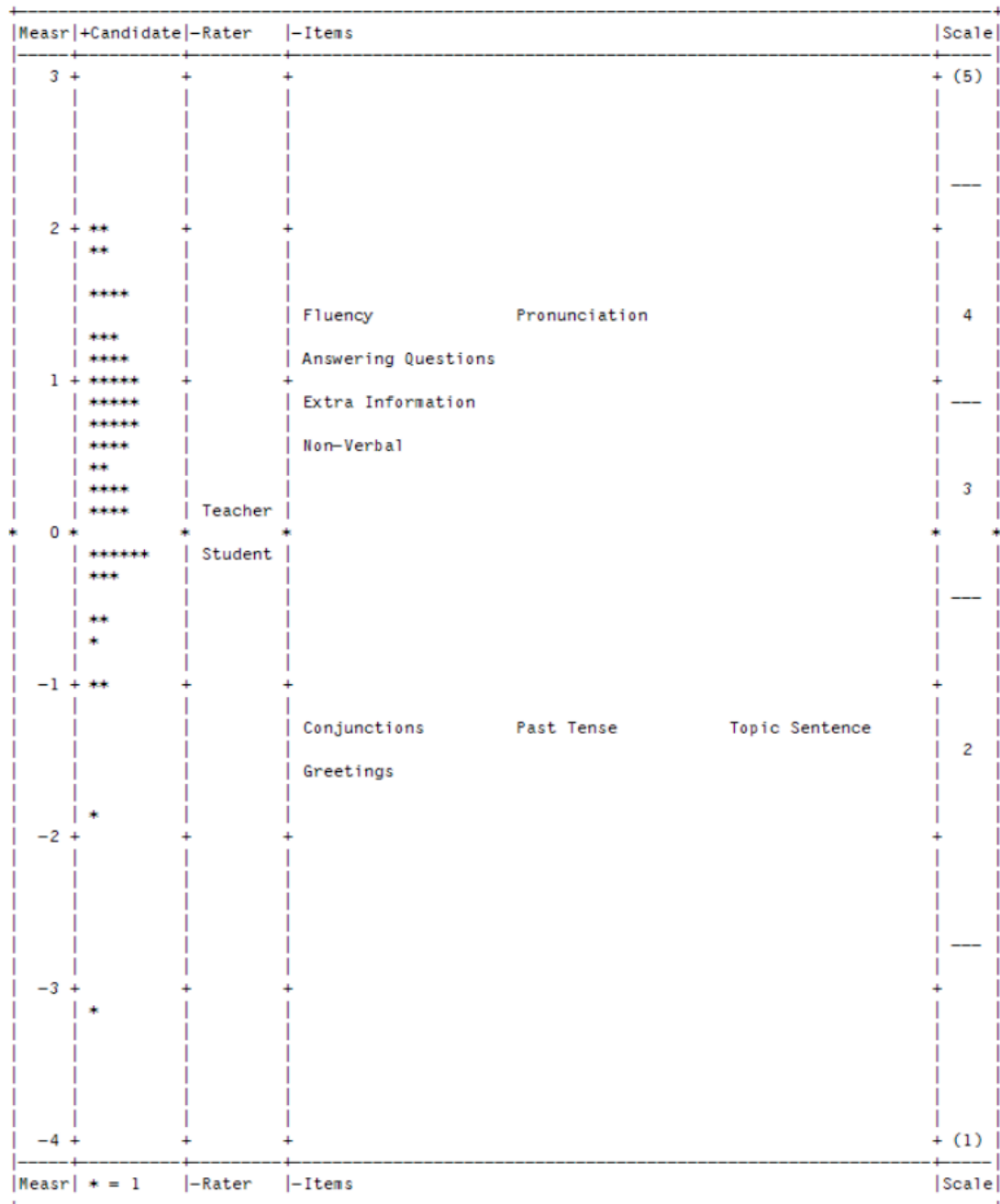


Figure 7.4 Person / Item / Rater map showing persons on a common scale with items and raters

The higher the rater appears on the chart, the lower the student's probability (indicated on the chart by an asterisk) of completing a particular criterion successfully when scored by that rater. Both raters appear at virtually the same level. If overall summative grades were to be awarded for the speaking task, therefore, the student and teacher grades would differ only slightly, or not at all. When cross-referenced with qualitative data, it

appears that the gap between student and teacher expectations was relatively small after the introduction of Japanese English-speakers as model texts.

7.5.1 Model texts: Reflection-in-action

During Cycle Two, reflection-in-action prompted a number of comments in my teacher journal emphasizing the necessity to ensure students were being taught to interact competently with unsympathetic strangers, and not merely with a teacher or fellow student in the classroom. It was considered important that the feedback students get from teachers should be a good guide to their performance outside the classroom: i.e., that teachers are not teaching them to talk to a sympathetic, familiar teacher, but to an unsympathetic, untrained stranger who is not used to the characteristics of a Japanese accent:

When I chose students as model texts I tried to imagine my mother trying to understand them if they were having a conversation with her, this really narrowed down the students I could use as models for the highest scores in the final assessment

Whilst keeping in my mind ethical considerations of notions of the “native speaker” and ensuring that students have realistic goals for their speaking proficiency, it is also necessary to consider levels of mutual intelligibility when students have to use their English language skills outside the classroom.

In an ESL context, demands of pronunciation and fluency can be shaped by the students’ immediate English needs outside the classroom, from speaking on the phone to utilities companies, to socializing in public areas. In an EFL context, ethical issues based upon the “native speaker” may be exacerbated. Students’ experiences and interactions with English may rely entirely on their experiences in the classroom and the choices of the teacher, with limited opportunity to use English in other contexts. In my teaching journal I noted the opinion:

It would be important therefore for EFL teachers to think carefully about the authenticity of model texts that students encounter and also notions of “native speakers” and mutual intelligibility. While planning syllabi based around a text-based approach a balance must be reached between preparing students to speak outside of the classroom, but also designing learning goals that are attainable and meet students’ needs.

Further systematic investigation of this opinion is necessary in order to propose any definitive reflective conclusion; but quantitative data presented in Section 7.6.1 did highlight a potential problem caused by not preparing students adequately to speak to unsympathetic strangers.

7.6 Assessment

As in Cycle One, assessment was another major theme in the data of Cycle Two. Additional teachers who shared data that were used in this study adopted the syllabus in their classrooms voluntarily, after a presentation of the thesis proposal in a teacher-training workshop initiated by the institution in which the study takes place. Teachers were particularly interested in assessment procedures for speaking tasks, so an assessment exercise was conducted by the four teachers who had adopted the genre-based syllabus. The exercise aimed to assess selected students’ spoken recount texts, as a means of starting a discussion on how we could assess the speaking abilities of students at our institution in the future.

Following a genre-based syllabus of instruction, 24 students were selected for the assessment exercise: six students were chosen by each teacher. A video of each student had been recorded during the final speaking-assessment portion of the syllabus, with the permission of the student. These videos were then rated again by each of the four teachers during the assessment exercise. Prior to the assessment exercise, criteria for the assessment were negotiated in a meeting of the four teachers (section 6.5.1 and Appendix 20). These assessment criteria would be used by all four teachers, and can be seen in Figure 7.5. Criteria

were scored on a scale of 1–5. Rasch measurement was used for item, person and rater analysis, based on scores awarded to each of the 24 students by all four teachers, using Facets software (Linacre, 2007a).

7.6.1 Assessment Results

Preliminary analysis showed that despite the small sample size, item reliability at .99 and person reliability at .95 was within acceptable limits for a speaking assessment, as cited by Hughes (2003). Preliminary rater reliability (not inter-rater) was measured at .98.

Exact inter-rater agreements were measured at 46.1%, with an expected rater agreement of 41.6%, which suggests that inter-rater reliability was higher than expected by Facets (Linacre, 2007a). This provided empirical evidence to suggest that the genre-based approach provided opportunities for inter-rater reliability during assessment. Facets (Linacre, 2007a) analysis, however, does not insist on “rating machines”, but instead calls for raters who behave like “independent experts”; it thus accounts for rater variation and differences in agreement.

Rater leniency and severity was also examined. Figure 7.5 shows a person, item and rater map on the common scale and highlights rater leniency and severity.

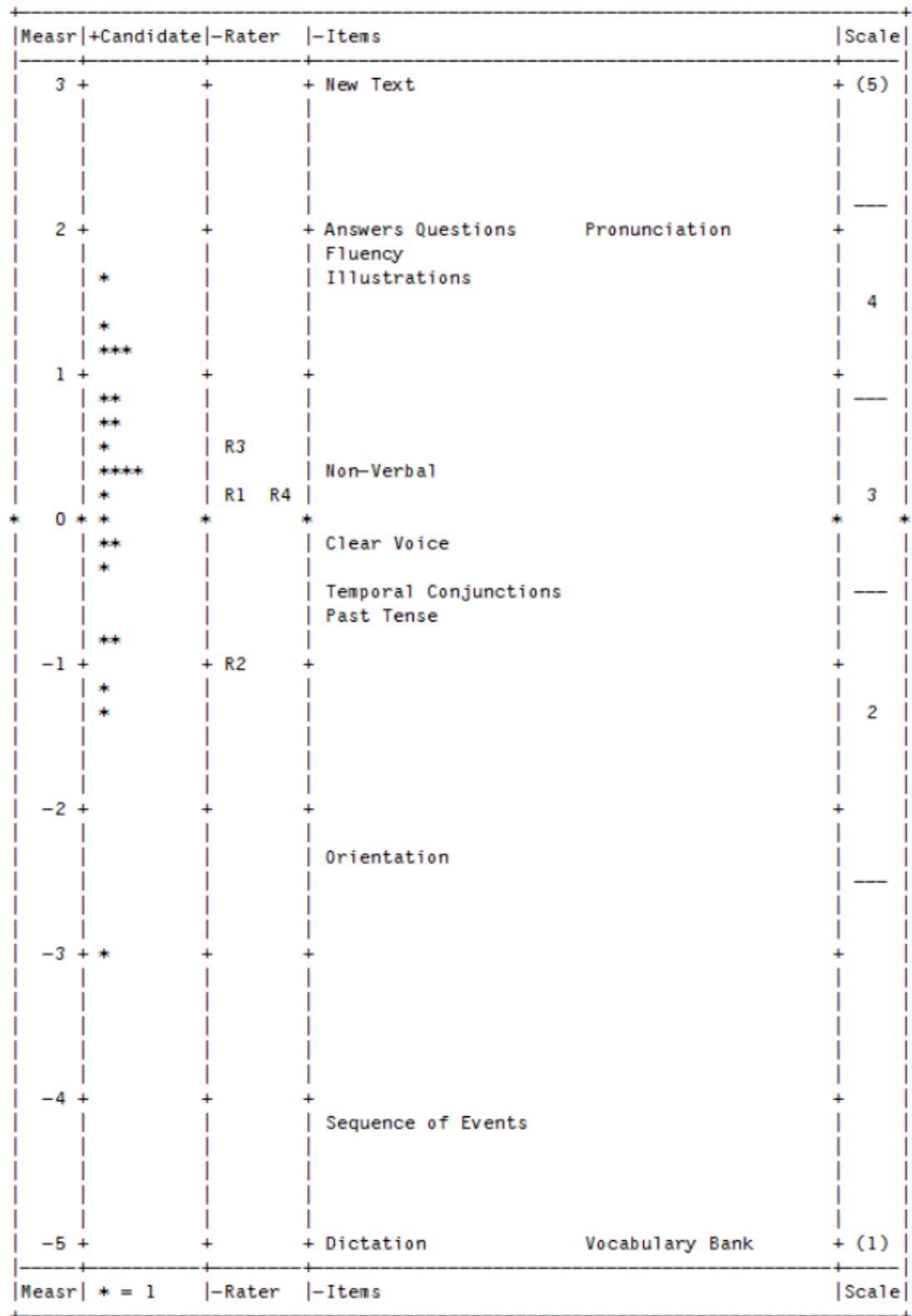


Figure 7.5 Person-Rater-Item map showing persons on a common scale with items and raters

The raters higher on the scale are stricter in their assessment. Rater 3 is the most severe, whereas Rater 2 is the most lenient. Raters 1 and 4 have been very similar in their assessments. The position of Rater 2 is interesting, in that it may call into question the inter-rater reliability of the assessment. Although the diagnostic information provided by the rater map allows us to account for this rater's leniency when distributing scores to students in summative tests, it suggests that raw scores for speaking assessment alone would not be a suitable form of student feedback. The very large substantive difference in severity is much more important than the reliability coefficient (stated as .98), as .30 logits is commonly used as a threshold for substantive significance (Bond and Fox, 2007). Here we have a difference of about 1.5 logits, which changes the probability of student success on an item from 50% to around 80% should they be interviewed by that rater. A closer look at fit statistics helps to diagnose the root of this leniency.

Table 7.2 Rater Fit Statistics.

Rater	Score	Count	Infit MnSq	Outfit MnSq
3	913	312	.89	.88
4	953	312	.99	1.09
1	963	312	.97	.84
2	1101	312	1.23	1.19

Table 7.2 shows that all raters are well within acceptable infit and outfit boundaries of .75 and 1.33 (Wilson, 2005), which demonstrates promising performance of raters on the test and highlights reliability. However, Figure 7.5 shows high leniency from Rater 2.

It was important to further investigate this leniency, for reasons discussed in 7.5.1. If the rater is rating familiar and unfamiliar students differently, then there is a question regarding the use of this instrument as a classroom formative test, because the feedback students get from teachers may not be a good guide to their performance outside the

classroom (i.e., raters are teaching them to talk to a sympathetic, familiar teacher, not an unsympathetic, untrained stranger, as discussed in previous chapters). As previously explained, during the assessment exercise, raters rated six of their own students. The key in Figure 7.6 shows that Rater 1 taught students C1–C6; Rater 2 students C7–C12; Rater 3 students C19–C24; and Rater 4 students C13–C18.

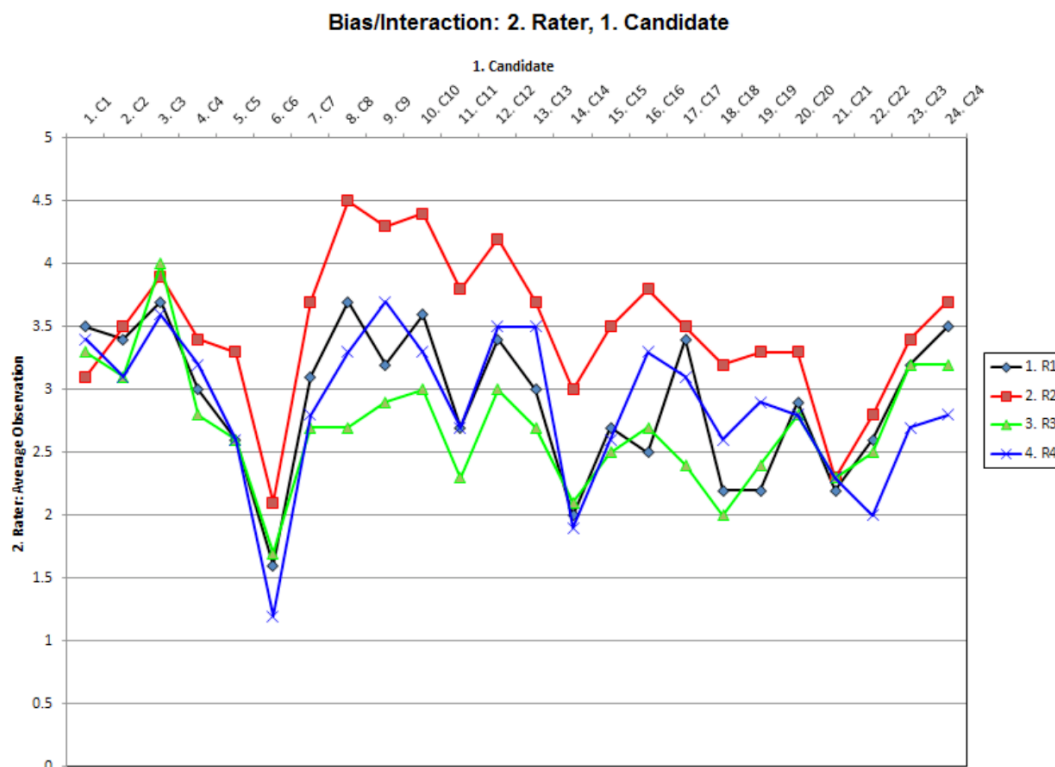


Figure 7.6 Chart showing bias/interaction between raters and students

Figure 7.6 shows interaction between raters and the different candidates. The higher mean ratings by Rater 2 are apparent from this chart, showing that Rater 2 is consistently lenient. However, it can also be seen that this rater is consistently more lenient when rating their own students, but assigns ratings more similar to other raters when rating unknown candidates. Rater 4, by contrast, appears to rate their own students more strictly than unknown candidates. Unfortunately, due to the small sample size, definitive conclusions are

not possible; but this result suggests that feedback from teachers may not accurately reflect how learners' performances are perceived by strangers, and thus that formative assessment may be improved by supplementing teacher feedback with feedback from strangers. These findings also have important implications for summative assessments, namely that emphasis must be placed on rating only the observed performance, with feelings or intuitions about what candidates might be capable of disregarded. This highlights a crucial difference between the role of teachers who aim to be sympathetic and assist learners in improving their performance and those who aim to be unsympathetic, and merely elicit samples that illustrate candidates' current ability to perform (Woolfolk et al., 1990).

When cross-referenced with data in previous chapters, we see a repeat of the idea that teachers should strike a balance between providing students with model texts, activities, feedback and assessment that are appropriate for their level, and to keep the aims and objectives of syllabi realistic. However, it is also important to consider that students should be prepared for situations in which they are not speaking with or listening to a sympathetic teacher, or a classmate with similar language proficiency. This would suggest the need for greater use of authentic texts, and perhaps summative assessments conducted by teachers who are not familiar with the student.

The data also highlights the limitations of summative assessments that provide summative scores for speaking assessment: without consistent and informed formative feedback being integrated into syllabi, students can be deceived by raw scores, and are likely to have no idea of how to progress in their speaking abilities and become autonomous learners. This reflects the argument by Black (2009) and Kennedy et al. (2006), that there is a place for summative assessment, but that such tasks should be treated as an occasion for formative feedback. It also strengthens the argument for self- and peer-assessment procedures, so that students can engage with the criteria of assessment and identify their strengths and weaknesses. These arguments suggest an inherent danger in summative

assessment designs that only provide raw scores in feedback. False notions, either positive or negative, may be accepted by the students as representing their true proficiency levels, and hinder identification of what they need to do to improve in the future.

7.7 Cross-Referencing with Qualitative Data

Assessment still caused anxiety in students, particularly when conducted face-to-face with the teacher during speaking. Considerable time was taken in the planning stage of Cycle Two to introduce assessment for learning strategies into the syllabus. Themes emerged in the data that were developed inductively and independently of the descriptions reported in the literature; these indicated some success in achieving the goals of assessment for learning, highlighted in section 6.2. Figure 7.7 highlights instances of the theme of assessment in the data, and from multiple sources, with examples provided from the raw data. The data generated themes associated with a recognition of learning goals, self-evaluation, independent learning and feedback.

Source	Assessment			
	Recognition of learning goals	Self-evaluation and reflection	Independent learning	Feedback
Anonymous student generated data	<i>I think I can make a perfect conversation by answering questions, I will try to make a routine to a higher level</i>	<i>I could answer the questions that you made, although I didn't know if I was correct or not, but you understood me</i>	<i>Here's a lot of foreigners working at part-time job place so I have chance, I want to try to make a talk with them</i>	<i>My target is making more conversations with the teacher because he helps me with my mistake</i>
Teacher journal and classroom documents		<i>Students are very honest about how they performed in class</i>	<i>A couple of students commented that they would like to use English more outside of class, I told them to make friends with overseas students!</i>	<i>I could use the criteria to give targeted and specific feedback on their speaking rather than simply saying "good", or "nice job"</i>
Student journals and classroom documents	<i>Temporal conjunctions are important I will use more in future</i>	<i>Unexpectedly I didn't know the English word I used, so I used only "went", I will check the meaning</i>	<i>It's very important to ask the question in conversation so I want to do more, I will practise with my friends</i>	<i>I was happy because my teacher, after seeing my face, he praised my talking</i>
Teacher 1	<i>The scaffolded approach provided good support on what to say and why.</i>	<i>The syllabus helped the students focus on what to do</i>	<i>The genre approach provided a good scaffolded for the direction they should go.</i>	
Teacher 2	<i>It provides a clear structured approach to both teaching, studying and assessing teaching skills.</i>		<i>The syllabus helped students with general life skills</i>	
Teacher 3	<i>Each of the steps were helpful and the students liked it.</i>		<i>I think a programme like this will help students get more confident and less fearful.</i>	

Figure 7.7 Matrix display to examine themes of assessment

A key objective of integrating assessment for learning into the syllabus during the planning phase of Cycle Two was to ensure that students knew what we were studying, why we were studying it, and how we were going to achieve success. Through joint deconstruction of texts and joint construction of assessment rubrics, students could feel

ownership of the assessment procedures. Student journals reflect that students engaged with the learning goals when they were made explicit:

Temporal conjunctions are important I will use more in future

Students were also very honest about the scores they awarded themselves, which I was pleased to note in my teacher journal:

Students are very honest about how they performed in class

During interviews with students after the final assessments, students were grateful for the feedback, and negotiating their final speaking grade went smoothly:

I was happy because my teacher, after seeing my face, he praised my talking

I also found in Cycle Two that I was able to give more targeted and specific feedback to assist speaking goals:

I could use the criteria to give targeted and specific feedback on their speaking rather than simply saying “good”, or “nice job”

One major concern was raised in my teaching journal regarding the integration of assessment for learning:

Today I felt that I achieved the objectives of the class, but it cannot be ignored that for the past two lessons the majority of the class time has been teacher-focused

Together with introducing a new genre-based approach, students were also being asked to adopt new ideas about assessment, and understand assessment rubrics. This greatly reduced time in class for actual speaking activities. My teaching journal notes that ultimately, I decided that:

the negatives of preparation time and explicit teacher-focused classroom instruction were outweighed by the benefits afforded to the students.

However, this is an important point to emphasize when generalizing findings, as the amount of planning necessary to conduct assessment for learning may not be viable in some classrooms: this concern is described further in section 7.8.

7.8 Impact on teaching

With the integration of collaborative teachers into data collection, during Cycle Two an additional major theme arose from the data, concerning the impact of a genre-based approach on teaching. Chapter 3 highlighted some of the limitations of relying too heavily on “insider”-generated data. Therefore, where possible, conclusions are based upon cross-reference with other data sources that were explored in previous chapters; with other teachers; and with reference to the literature. Four main themes were identified in the data that showed an impact of the syllabus on teaching, as exemplified in Figure 7.8: these comprise effects on the teaching of grammar, on syllabus design, on the teacher’s role in the classroom, and also anxiety or concerns raised by the syllabus design.

Impact of a genre-based approach on teaching				
Source	Grammar teaching	Syllabus design	Teacher's role in the classroom	Concerns
Researcher	<i>Students seemed more conscious of the importance of grammar when speaking</i>	<i>The genre-based approach assisted me in thinking about the needs of my students and how I was going to get there</i>	<i>I could become more of a facilitator in later classes as students grew more confident in the procedures</i>	<i>Storyboarding worked well, but something entirely different would be needed to solve this problem for other genre types</i>
Teacher 1	<i>Conjunctions helped them sequence their ideas</i>	<i>It was nice to try something other than "practice and memorize this conversation"</i>	<i>I remember students commenting that they really enjoyed the opportunities to talk in English about their stories with their friends. I think if that was all they did and didn't have to face the teacher in the last class, they would be very content with it. Facing the teacher was very scary for them</i>	<i>Yes, if I really wanted to focus on teaching students how to have a conversation about something they did recently, I think it's great for that. Obviously it could be used for talking about future plans as well. However, if I were teaching students, say, how to offer an opinion and defend it, then I would not use this format</i>
Teacher 2	<i>There was a purpose for teaching them grammar</i>	<i>It provides a clear structured approach to both teaching, studying and assessing teaching skills.</i>	<i>I was more involved in the class than usual, the extra energy that this approach takes was tiring.</i>	<i>I think it is a lot of work to keep up with all the prints and steps</i>
Teacher 3	<i>It helped contextualize grammar point and made it easier to teach.</i>	<i>Step-by-step framing a story was very helpful. Doing a step per class for 20 minutes helped develop their ability to construct a story and to have something to say in English. Although students were still quite nervous to take the test at the end.</i>	<i>I did a lot more talking at the front of the class than usual, especially at the start, but it balanced out a bit at the end</i>	<i>I think it's a good way to practise talking about a past event, but I am not sure if this syllabus could be used to teach other speaking situations (giving directions, making suggestions, declining request, etc.)</i>

Figure 7.8 Matrix display to examine the impact on teaching

Each of the themes in Figure 7.8 is now discussed in greater detail with reference to representative examples drawn from the raw data.

7.7.1 Grammar teaching

Academics have pointed out that grammar teaching is often neglected or intentionally omitted entirely from communicative syllabi (McDonough and Shaw, 1993; Feez and Joyce,

1998). In a Japanese context, Yamaoka (2010) describes an apparent dichotomy between communicative approaches and grammar, where on the one hand grammar is ignored entirely when the objective is to get students speaking, whereas it is given primary focus when students are prepared for university entrance exams. When grammar is taught, it is primarily done in a non-contextualized manner. In the local context of this study, this dichotomy is also prevalent, and summarized by a comment in the teacher journal:

After observing my class, one of the observers told me that I shouldn't be teaching grammar points such as past tense verbs, as this is already covered in Japanese (nationality) teachers' classes

This comment was made by a senior member of the English department at a school that was not included as part of this study, but it mirrored other anecdotal observations. As previously described in section 1.1, compulsory English classes in the institution where the study was conducted are broadly divided into two different types: “Reading and Writing” classes taught by Japanese teachers, and “Conversation classes” taught by “native speakers”. Beyond the title of the class there are no syllabus or general curriculum guidelines to follow, as syllabi are constructed by individual teachers and there is no curriculum. It is generally expected that native speakers should be facilitators of fluency-based tasks that involve some form of spoken output where the Japanese language is not used. Japanese teachers then discuss grammar rules and examples in the Japanese language.

It is important to place the findings of this research in this context in order to illustrate the implications of a genre-based approach for English teaching in Japan, particularly the integration of grammar into “conversation classes”. Adoption of this approach may face resistance from educators who believe in a dichotomy of fluency and accuracy, in a context where the role of grammar has been undervalued (Joyce and Burns, 1999), or where grammar

points need to be addressed in a non-contextualized way in L1, particularly in preparation for tests such as entrance exams.

Teachers involved in this study were largely positive towards the increased integration of grammar into their communicative classes. The data show that teachers viewed the integration of grammar as making grammar teaching simpler and learning easier, due to the following aspects:

- Assisting textual cohesion
- Making students aware of the need for grammar to assist fluency as well as accuracy
- Creating a purpose for grammar instruction that is applicable to students' real-life experiences
- Structuring the teaching of grammar in a logical manner, rather than the ad hoc sequencing of grammar points often found in textbooks
- Presenting grammar points in a simple and clear format

One teacher noted that by exploring different types of temporal conjunction, students were able to sequence a string of different ideas in their talk, rather than simply uttering one short summative sentence that lacked sufficient detail. This observation is reinforced by a student's comment in section 7.7, that temporal conjunctions were important and that they desired to use them more in the future. Students were thus developing an awareness of the need to use grammar to assist their speaking fluency, as well as for accuracy. The syllabus also made students aware that grammar was directly applicable to using English to describe their own experiences, rather than in the non-contextualized formats they may have been accustomed to in textbooks that prepared them for university entrance exams.

It was noted in the teaching journal that in textbooks there is often no clear logical reason for the sequencing of grammar points, and no explanation as to why they are ordered

in such a way. In the local context of this study, teachers were encouraged to choose from a list of prescribed textbooks if one was to be selected for use in the syllabus. One such textbook orders its grammar instruction in its index of units in the following sequence:

- contractions
- past tense verbs
- prepositions
- adjectives
- countable/non-countable nouns
- irregular verbs

No explanation is given as to why grammar points are sequenced in this way, and it is not easy for the classroom teacher to discern if the grammar points complement each other or are leading to a specific goal in grammar instruction. Teachers commented that in a genre-based approach, grammar was sequenced in a logical format, with a clear progression of goals; and that a contextualized focus on directly relevant grammatical points made instruction easier to teach and for students to understand:

It provides a clear structured approach to both teaching, studying and assessing teaching skills

There was a purpose for teaching them grammar

It helped contextualize grammar point and made it easier to teach

There were, however, some issues raised regarding the scope of grammar instruction that was viable in such an approach; this point will be discussed in more detail in section 7.7.4.

7.7.2 Syllabus design

Feez and Joyce (1998) argue that language teaching approaches have approached different aspects of language as distinct building blocks, such as grammar, vocabulary, pronunciation; and as such, they have influenced the creation of syllabi by convincing teachers to select items from lists of these building blocks as discrete units according to difficulty. They argue that items such as grammatical forms were presented in contrived, isolated sentences, with memorization of forms and rules. A grammar-based syllabus, meanwhile, they view as handling language as a communicative resource, so that learners should be dealing with extended stretches of language in authentic contexts of use. The raw data presented in 7.7.1 suggest that via a genre-based approach we can sequence units of work in a clear and principled way that ensures integration of grammar. A genre-based approach allows a range of different approaches to be integrated into its design, countering previous criticisms that communicative language teaching followed fads and trends (Feez and Joyce, 1998). Aims and objectives can easily be shared with students and made explicit, thus creating a cycle of learning and teaching that creates a step-by-step approach to improving language production.

“*Framework*” and “*scaffolding*” (Teacher 1, Figure 7.7) were words that appeared frequently in the raw data, and emphasized both the way in which a genre-based approach provided scaffolds for students’ learning, and also frameworks for classroom content and lesson planning, as exemplified by a teacher comment:

It provides a clear structured approach to both teaching, studying and assessing teaching skills.

7.7.3 The teacher’s role in the classroom

In this study, integration of grammar into the syllabus required variance in the teacher’s role in the classroom: this was another theme that was prevalent in the data, as exemplified by teachers’ comments:

I was more involved in the class than usual, the extra energy that this approach takes was tiring.

I did a lot more talking at the front of the class than usual, especially at the start, but it balanced out a bit at the end

It was nice to try something other than “practise and memorize this conversation”.

According to Paul (2003), the prevalent notion in Asia is that the teacher's role in the classroom is as an authority who transmits to students knowledge that they do not possess. Such a situation has led to students who are not used to adopting the role of inquirers and self-directed learners. In high power-distance cultures such as Japan (Ryan, 2000), students tend to be more passive and may be reluctant to participate in communicative activities, as they are not used to speaking in front of their superiors.

Grasha (1994) outlines five teaching styles that are prevalent in classrooms: the expert (transmitter of information); formal authority (sets standards and defines acceptable ways of doing things); personal model (teaches by illustration and direct example); facilitator (guides and directs by asking questions, exploring options, suggesting alternatives); and delegator (develops students' ability to function autonomously). Robertson (2004) claims that the adoption of instructional strategies is closely related to teachers' perceptions of their roles, and perspectives about teaching and learning. Ertmer and Newby (1993) claim that in European and North American teaching traditions, there has been a shift from teacher-orientation to student-orientation, originating from ideas in constructivism. There is clear potential for tension between what Japanese students expect from a teacher and how a European or North American teacher would expect the class to be conducted.

The genre-based approach, therefore, has potential to address this tension by incorporating both teacher-orientated and student-orientated approaches into syllabus design.

It was noted in my teacher journal that although initial classes were heavily teacher-oriented, in later classes the focus switched to classes that were primarily student-centred:

I could become more of a facilitator in later classes as students grew more confident in the procedures

At the same time, however, the role of facilitator has the potential to alienate both students and teachers if aims and objectives are not made clear and understood by both parties. Feez and Joyce (1998) suggest that stages of the teaching and learning cycle in the genre-based approach should be based upon modelling, deconstruction, joint construction and independent construction of texts. In this process, the teacher can utilize all five of Grasha's proposed teaching styles in the classroom, thus allowing the teacher to consider the cultural context in which the students are studying, whilst gradually acclimatizing them to function autonomously. Data discussed in section 7.7 show that during this study, students were able to transition to autonomous learners, with the teacher taking the role of facilitator as the lessons progressed. However, as Richards and Lockhart (1996) point out, research shows that whole-class instructional methods are most commonly used in public-school teaching. Teacher 3 in this study, meanwhile, illustrated that they did much more work at the front of the class than usual when adopting a genre-based syllabus:

I did a lot more talking at the front of the class than usual

These various notions of what is expected of a teacher by students and by individual teachers adds another level of complexity to the widespread adoption of a genre-based approach. Other issues that appeared in the data indicate that the genre-based approach did not have an entirely positive impact on teaching in all aspects.

7.7.4 Issues with the syllabus

Numerous issues and considerations have been raised when considering the adoption of a genre-based approach, in this chapter and in previous chapters. One that has already been discussed in this chapter is the workload associated with a genre-based approach, particularly when careful consideration is given to integrating formative assessment practices. All teachers involved in the implementation of the syllabus in this study noted the tiring aspect of conducting such an approach in class. The process of preparing materials for the syllabus and adapting them during the cycles of action research was also extremely time-consuming. Additional to this was the exploration of students' individual speaking abilities, which in this study culminated in the teacher conducting face-to-face interviews with up to 90 students a day for two consecutive weeks, whilst consecutively performing assessment procedures, and finally providing individualized feedback. For teachers who feel they are already overburdened with teaching and research demands, the additional demands of a genre-based syllabus may ultimately have detrimental effects on both teaching and learning. This is summed up by a comment made by Teacher 3:

With doing research projects and being busy outside of school, I think I just haven't had the extra energy that this approach takes, and I'm not sure I will continue in next semester

A second issue raised in the data was a concern from teachers that such an approach was not applicable to different genres. Although teachers saw the benefits of using this approach to teach past and present events, they were unable to formulate ideas as to how they would utilize the same approach in different genres. In the teaching journal, I also noted that I was finally able to overcome numerous obstacles that I encountered in the action research cycle by using this approach; but other unique issues are certain to be encountered following the design of syllabi that tackle additional genres. Indeed, due to time constraints and student timetables dictated by the university, it was only possible to explore the use of one genre in

this study. A major drawback of this research is the limited exploration of different types of genres beyond superficial references exemplified in worksheets, such as that shown in Appendix 1.

A final issue that proved problematic for teachers was micro-aspects of conversation; this mirrored findings in Cycle One, which are summarized in section 5.5. Micro-aspects of conversation elicited feelings of frustration in students and difficulties in teaching and assessment; these were exacerbated when having a conversation with a teacher. The “native speaker” fluency in asking and answering questions was not an achievable learning goal for many of the low-proficiency students, who needed a greater amount of time to develop questions, or to answer questions beyond the monologic chunks of their conversations. During the planning stage of Cycle Two, attempts were made in the syllabus to address these issues. Appendix 18 shows some of the worksheets that attempted to engage students with the micro-aspects of conversation; however, as notes written in the lesson plans describe:

We need more practice with this kind of stuff

This needs to take up a considerable amount more lesson time

Issues of students being unable to achieve the micro-aspects of conversation were very difficult to resolve in the time-frame allotted to class. It seemed clear that the best approach to developing these aspects was through practice, and this was one of the major drawbacks of a genre-based approach in an EFL setting. Although the genre-based approach provided a framework for constructing monologic chunks of conversation and preparing students for expected structures in casual conversation, there was not enough classroom time to create the necessary opportunities to simply practise conversation. In an ESL setting, time outside class is perhaps the best opportunity to practise casual conversation; whereas in an EFL setting, this is the only time that the majority of students will hold a conversation in

English. Time demands entailed by such a syllabus are therefore exacerbated. The necessity of explicit instruction reduces the amount of time for conversation practice, particularly with unsympathetic interlocutors. In an EFL setting, it is imperative that additional opportunities for casual conversation are introduced to students. Discussion of such strategies lies beyond the scope of the present study; nevertheless, in further cycles of action research, the syllabus did include homework activities that required students to find partners for conversations outside the classroom. A website was created that allowed students to upload videos from their smartphones. The teacher asked students to interview parents, friends and classmates about what they had been doing recently, or similar topics that initiated the use of a recount conversation.

7.8 Post-Syllabus Student Dialogues

Sections 3.5.1 and 5.5 provide representative examples of students' spoken recount texts before and after Cycle One of the action research process. Appendix 23 shows two examples of student recount texts after the completion of the Cycle Two observation phase. Transcript A is very similar to the dialogue in section 5.5; we see a clear orientation and sequence of events that provide a basic recount text, with accurate use of past tense verbs and temporal conjunctions. Answers to questions are more developed than in Cycle One, but there is still evidence of a memorized monologue, such as in lines 12 and 16, where the teacher asks a question and the reply is somewhat abrupt, before the student quickly proceeds to the description of a new event. In line 17, we can see that the teacher attempts to understand the term "*Bon* dancing", which the student appears reluctant to elucidate. The same avoidance process occurs in line 29, when the teacher again asks for clarification of "*Bon* dancing", which the student attempts to ignore in order to complete their monologue. Overall, the students' spoken output appears to be a clear improvement on the examples provided in section 3.5.1, with longer utterances and accurate lexico-grammatical elements and structure. However, the dialogue once again does not really exemplify a successful attempt at casual

conversation. Slade's (1997) description of casual conversation as a tool for social bonding is not fulfilled in this interaction. In fact, such an interaction might be construed as uncomfortable, and the use of avoidance techniques might appear rude.

Transcript B in Appendix 23, however, indicates that this phenomena may not necessarily be entirely a result of memorized dialogues at the expense of social bonding techniques: it might be due to the unequal relationship between the interlocutors, and the uneven balance of power that is exacerbated by the assessment procedures. Transcript B is a recount dialogue that students video-recorded on their phones in a part of the classroom away from the teacher. As part of the reflection-in-action process, at the end of Cycle Two I began to consider ways I might in future remove myself from the face-to-face aspect of the final assessment procedure, and experimented with asking some students to record another recount text at the end of the syllabus and upload the videos to my website. Transcript B shows students sharing the key lexico-grammatical and structural components of the recount text that formed the syllabus aims and objectives, but with far greater adoption of turn-taking aspects of casual conversation. In the representative dialogues illustrated in section 3.5.1 and 5.5, we see conversations with a brief greeting followed by an orientation and sequence of events. In Transcript B this initial greeting process is much more detailed and spontaneous, with a greater sense of a social interaction taking place. In line 13 we see an orientation, which is followed by a sequence of events in lines 19 and 23 with accurate use of past tense verbs and temporal conjunctions. Moreover, there is clear evidence of spontaneous turn-taking, and none of the avoidance of questions displayed in conversation with the teacher. In line 26, the students begin a non-recount genre interaction about the type of food they prefer, which continues until line 37. The video-recording itself is filled with laughter and gesturing, to aid understanding and strengthen the interlocutors' social bonds. When Student B begins their recount text in line 38, some lexico-grammatical aspects of the interaction are not as proficient as those of Student A, but the conversation continues as a successful social

interaction, and communication does not break down. In line 52, Student B talks about her money “fly(ing) away” and there is much laughter; she repeats this joke in line 58, and it is echoed by Student A. The jokes in lines 49 and 50 about the “handsome” Exile pop group member and his “face” are interactions that would probably not occur with the teacher, particularly during assessment. This highlights the tension between developing students’ skills in casual conversation and the necessity for assessment procedures with the teacher; such assessments make the conversation less authentic in terms of their social goals and the students’ abilities to illustrate their turn-taking skills.

7.9 Reflection

The planning phase of Cycle Two appeared to address some of the major problems that arose during the reflection stage of Cycle One. The adoption of assessment for learning removed some of the nervousness and frustration students felt with the syllabus and changes to the model text; and additionally, subtitling and storyboarding removed students’ reliance on written output to inform spoken output. However, issues with turn-taking and micro-aspects of conversation still proved to be persistent.

7.9.1 Reflections on student emotion

During the planning stages of syllabus design, student nervousness and issues with language proficiency were considered. One reason why student assessment was integrated into syllabus design was to reduce student nervousness in test-taking. Tenor was also explored with students, in order to lessen the formality of the final test. However, ongoing observation and subsequent data-analysis illustrates the difficulty that the students in the dataset have with the micro-aspects of casual conversation. An interview between student and teachers as a form of summative assessment at the end of the syllabus was included, to explore students’ abilities to engage with the micro-aspects of dialogue. The data illustrate that the inclusion of this final interview with the teacher increased formality and nervousness in the spoken

interactions, which negatively affected student emotions. I developed Figure 7.9 to summarize the data in abstract terms, to express how student nervousness and frustration increased with the formality of the conversation, based on the purpose of the talk and the relationship between the interlocutors:

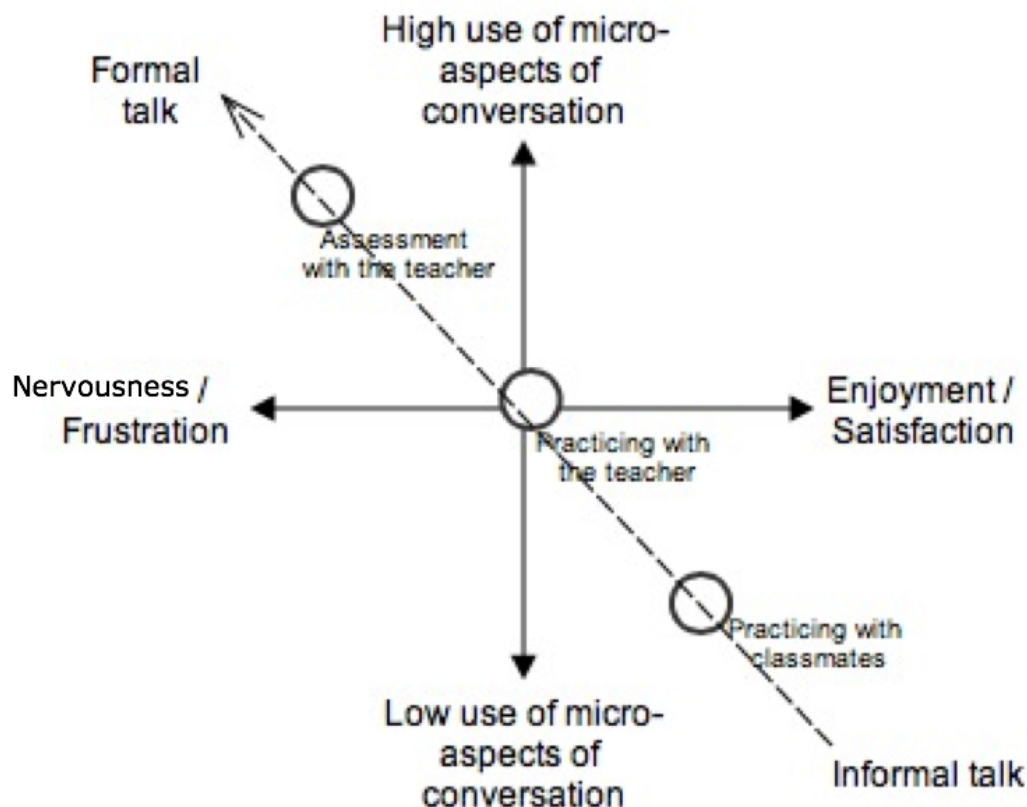


Figure 7.9 Student emotions and formality of talk in the classroom context

The formality of the interaction increases depending on who is speaking with the student and the purpose of the speaking; the use of micro-aspects of conversation also increase with this formality. With the increase in formality and micro-aspects of conversation comes an increase in student nervousness and frustration.

Assessment and interaction with the teacher increased inequality in the power relationship between the interlocutors, and increased student nervousness and frustration.

This may be due to the teacher being a model of “native speaker” English, which induced anxiety. This raises the question of validity in the assessment when performed with the teacher. If the aim of the assessment is to measure students’ ability in casual, informal conversation, perhaps teacher involvement in the conversation is not appropriate. However, micro-aspects of conversation might then be neglected, as students are then unable to create these aspects of conversations for themselves. There is also the issue that when students leave the classroom to interact with other speakers of English, they might face difficulty. Their classmates are likely to be sympathetic to their mistakes or lack of fluency; whereas when they meet an unsympathetic interlocutor, students may not be prepared for an appropriate level of conversation.

Possible ways of addressing these issues in the cycle of action research and syllabus design may be to create three-way conversations between two students and a teacher for assessment. The teacher would observe the conversation between two students and contribute only when conversation breaks down, such as moments when asking questions is appropriate. A further alternative would be for all students to converse in pairs while the teacher walks around and interjects where appropriate in different conversations. In addition, all conversation could be self- and peer assessed. In this way, the teacher could then ensure that the formality of the conversations is reduced, but could also act as a facilitator by assisting students’ problems with micro-aspects of conversation in a non-intrusive manner.

A key component of the problems highlighted might be the educational context of the syllabus. Unlike ESL classes, where opportunities to practise and recycle conversational strategies exist in many other contexts outside the classroom, in EFL classes this is not the case. English interaction is largely restricted to the classroom (Reid, 1995). Repeated experience and exposure to micro-aspects of discourse might be the best way to strengthen these skills, despite the nervousness and frustration it creates in students.

With reference to the research question the data discussed in Cycles One and Two suggest that a genre-based approach affords opportunities to use English in class in purposeful ways that create enjoyment and satisfaction. However, student anxiety and frustration also shows that spoken output is a highly emotional experience for some students. Immediate feedback from a native speaker and micro-aspects of conversation create nervousness and frustration. The key to reducing student nervousness and frustration appears to be finding ways to enhance the enjoyable aspects of conversation, such as making it purposeful and informal, with the teacher acting as a facilitator rather than an assessor, wherever possible.

7.9.2 Reflections on the model text

The data in Cycle One raised important issues when considering the introduction of model texts for deconstruction in class. Idealized, transcribed or over-simplified model texts did not adequately prepare students to address issues of pronunciation and fluency. Japanese students' reliance on transcribing and memorizing utterances before speaking them might have enhanced grammatical abilities and produced longer utterances, but this did not adequately prepare them for turn-taking and the micro-aspects of casual conversation. Despite the difficulty of introducing students to authentic texts, it was apparent in the data that making tasks easier, rather than the texts, was the optimum approach. In Cycle Two, subtitled videos and storyboarding allowed students to experience and produce authentic spoken language without becoming reliant on writing as a form of output, whilst also allowing an explicit exploration of grammar and structure by pausing the subtitled video on key sentences (see Figure 6.1). The data suggested that it is important to highlight explicitly to students the difference between spoken and written English, and make them aware that memorizing texts that they have written will make them sound unusual to unsympathetic strangers: this will not prepare them adequately for the micro-aspects of conversation, such as quickly asking and answering questions.

Who is speaking is also a key consideration when introducing model texts. The EFL teacher is responsible for the majority of English students' experiences with the English language, and ethical issues that address pronunciation and fluency, as well as desired learning goals, should be considered. The idea of a "native speaker" sets a model of text that contains elements of speech that are unobtainable to students in this context. The data suggested the importance of introducing Japanese speakers of English as model speakers, particularly those they can relate to directly, such as their peers. For appropriate learning goals to be integrated into a syllabus that meets students' needs, the concept of mutual intelligibility needs to replace that of a "native speaker". It is important for students to consider the value of a Japanese speaker of English; this may even reflect some of the idiosyncrasies of the Japanese language, provided that definitions of mutual intelligibility are met. The notion of a native speaker also entails learning goals that are vague and abstract, with unobtainable criteria for assessment; this does not assist the creation of learning goals that students feel they can adequately address in the classroom.

7.9.3 Reflections on assessment

Issues with assessment suggested that the integration of micro-aspects of conversation with fluency and pronunciation is best achieved by providing a range of model texts that show Japanese speakers of English of varying proficiency. Students are then able to quantify, through observation of these model texts, a learning goal for their own proficiency that is relevant to them, and which reinforces the notion of mutual intelligibility rather than that of the native speaker.

In terms of assessment, different aspects of speaking ability should be considered, and different types of assessment are necessary. Structural and lexico-grammatical elements of speaking might not be measurable in a summative assessment that also includes universal speaking abilities such as pronunciation and fluency. Speaking assessment does not seem to

follow a unidimensional trait, but requires the measurement of a range of language skills. As such, a single summative assessment does not seem an adequate way of measuring students' speaking abilities, or provide opportunities for appropriate feedback. Assessment for learning approaches that include formative feedback and student and peer assessment should complement any summative assessments.

The data exemplified that assessment for learning procedures and the genre-based approach complemented each other very well. The genre-based approach provided clear and realistic goals for student achievement, in a way that students could share and understand.

Model texts combined with analytic rubrics provide opportunities for making assessment criteria explicit and understandable to students. Negotiating criteria together between student and teacher is one way of addressing speaking objectives that might become too prescriptive, and thus dissuade students from experimenting with language. Model texts of Japanese speakers of English provide quantifiable examples of the micro-aspects of conversation and levels of pronunciation and fluency, thus promoting mutual intelligibility, rather than an unobtainable or abstract notion of the native speaker.

It should be remembered, however, that the genre-based approach and assessment for learning strategies require additional teacher-orientated classroom time, in comparison with other communicative approaches. An increase in teacher-orientated instruction is a necessity that some teachers might need to consider before adopting such an approach. The data also raised issues of preparing students for speaking English outside the classroom with unsympathetic strangers.

7.9.4 Reflection on the impact on teaching

Cycle Two illustrated additional implications for teachers that a genre-based approach might entail in similar contexts. The genre-based approach addresses some of the issues of fluency versus accuracy that were explored in section 2.9.1. A genre-based approach appeared to allow a principled and logical integration of grammar into syllabi that are primarily focused on speaking (section 7.7.1). Students' comments showed that when grammar points were presented in context, they were able to connect knowledge of grammar with an ability to increase their fluency. Thus, rather than non-contextualized grammar points, a genre-based approach highlights the necessity of grammar in real-life situations, exemplified by comments such as:

Temporal conjunctions are important I will use more in future

Conjunctions helped sequence ideas

The genre-based approach provides a scaffold for sequencing classroom content and allows teachers to be eclectic in their choice of language activities, as shown in comments from Teacher 2:

The scaffolded approach provided good support on what to say and why.

The genre approach provided a good scaffolded for the direction they should go.

A teacher-centred approach is also appropriate, with explicit instruction being a necessary mode of learning, but with ample opportunity to allow more student-centred learning later in a semester, and with the ultimate goal of inspiring autonomous learning, as discussed in section 7.7.3.

Nevertheless, a number of issues were also raised. Time constraints associated with this approach involved the necessity for teacher training, in order for such an approach to be properly understood and implemented in the classroom. The perceived value of such an approach might also rely on the appropriate training of teachers regarding the complex theories of systemic functional linguistics in which the genre-based approach is based. The integration of assessment for learning strategies would also require teacher training. Summative assessments would also require negotiation and training, as well as a possible need to ensure that students experience assessment by a variety of raters, some of whom they are not familiar with, in order to prepare them for talking to unsympathetic strangers. Summative assessments also need to be supplemented with formative assessments, so that appropriate feedback is available. Once again, micro-aspects of conversation and issues of fluency and pronunciation proved problematic for teachers. A teacher-centred classroom, and explicit discussion of meta-languages that need to be made explicit to students in this approach, reduce classroom time that could be used to simply practise speaking. The necessity of time to practise and speak as much as possible is exacerbated in the EFL classroom, where opportunities to put speech into practice outside class are limited. These demands of teacher training and time place a very large burden on teachers, and might prohibit widespread acceptance of such an approach, particularly where the value of such an approach is not readily apparent.

Chapter 8: Final Reflections and Future Planning

8.1 Introduction

This chapter attempts to draw together the main conclusions from the two action research cycles; it discusses the generalizability of the research, and proposes future directions for a third cycle of action research beyond the scope of this thesis.

8.2 Summary of Findings

The data reflected that speaking English was a profoundly emotional experience for students (section 5.3.1 and section 7.2). These emotions were even able to manifest themselves through physical displays of nervousness, as well as through data that illustrated a strong emotional response (Figure 7.1). Talking to a teacher or talking for assessment exacerbated feelings of nervousness and frustration, to such a level that it was clear that such emotions would need to be taken into account in order to ensure effective teaching and learning. The data identify that the main triggers of nervousness and frustration stem from ideas of a native speaker of English, the tenor of talk when interacting with a teacher, talking for assessment purposes (section 5.3.2 and section 7.2), and the difficulty in the teaching and learning of micro-aspects of conversation (section 6.3.3 and section 7.9.2). These triggers also ensured that the gap between student and teacher expectations in the classroom was widened by learning goals being disrupted or obscured.

The concept of the native speaker, which is institutionalized in the context of this study and further afield in English language teaching in Japan, introduces the idea of a measurement of speaking ability that is abstract to students; unobtainable in terms of assessment and learning goals; poor in terms of providing opportunities for targeted feedback; and which invests the teacher with an superior power status that increases student anxiety in communication, to the detriment of classroom learning. The model of a native speaker eliminates the value of a mutually intelligible form of Japanese English, and makes

access to authentic listening texts extremely limited. Talking to a native speaker made students nervous and frustrated, as they were unable to mimic or reproduce micro-aspects of English communication or achieve similar levels of fluency and accuracy with their interlocutor (section 6.3.3 and section 7.9.2). Whilst it is also imperative to prepare students for communication with unsympathetic strangers in real-life situations (section 7.6.1), a balance is necessary between achieving learning goals, reducing anxiety, and achieving mutual intelligibility. It is also true that some nervousness and frustration is a natural part of casual conversation and the establishment of social bonds; nevertheless, unless these considerations are taken into account, classrooms may not be able to achieve learning goals (section 5.3.1 and section 7.2).

The data illustrated a number of strategies that proved successful in addressing issues of achieving appropriate learning goals in speaking, whilst also preparing students for communication in real-life situations. These strategies included the classroom teacher taking a more informal role in the classroom, by allowing students to have conversations amongst themselves and to assess themselves and each other (section 5.3.1, section 7.2 and section 7.8). At the same time, opportunities for explicit instruction and teacher-centred classrooms are not only available, but also encouraged, in a genre-based approach (section 7.9.4). However, this is balanced by giving some power of assessment to students, and ensuring formative feedback is integrated into the syllabus (section 5.5, section 6.2 and section 7.7). After explicit instruction, students should be allowed to learn autonomously, with the teacher taking the role of facilitator in speaking situations, and providing targeted help where needed, to encourage more conversation amongst peers (section 7.7). Removing the necessity for a native speaker also allows greater use of authentic texts that focus on Japanese speakers of English, where the aim of speaking is to achieve mutual intelligibility (sections 5.3.2, 7.4, 7.5.1, 7.9.1 and 7.9.2). Japanese models of whole texts, situated in different genres, provide learning goals that are accessible and achievable to students, as well as being directly

relatable to their everyday lives. The data showed that in the context of this study, the introduction of authentic texts removed students' reliance on memorization and written English as their primary means of output (section 6.3 and 7.4). Subtitling videoed model texts ensured that the mode of communication remained that of speaking, whilst still allowing a focus on the generic lexico-grammatical and structural features of spoken genres (section 7.4). Storyboarding of conversations allowed students to prepare chunks of talk for future conversations without the need to transcribe, and ensured that students were aware of the differences between spoken and written English.

The data in this study reflected the importance of planning and integrating different forms of assessment into a syllabus that aims to improve students' speaking abilities. In the context of this study, summative assessments proved to be an unsatisfactory indicator of students' spoken ability, and did not assist learning goals (section 5.6). The data suggested that spoken output is not measurable by a single unidimensional trait, and that a range of language skills are utilized when speaking: from lexico-grammatical to pronunciation and fluency (section 5.4). Summative feedback might be providing students with inaccurate assumptions about their language needs, although the role of feedback was essential. Formative, self- and peer assessment allowed students to identify learning goals, to situate their own speaking on those learning goals, and it promoted autonomous learning in order to achieve their own models of spoken texts (section 7.7).

Figure 8.1 summarizes some of the main findings of the research. The model shows factors that led to an increase in student anxiety and frustration, which widened the gap between student and teacher expectations. The items closer to the centre of the circle indicate lower levels of nervousness and frustration in student responses, and higher instances of student enjoyment and satisfaction with the syllabus. As the circle extends outwards, the gap widens between teacher and student expectations.

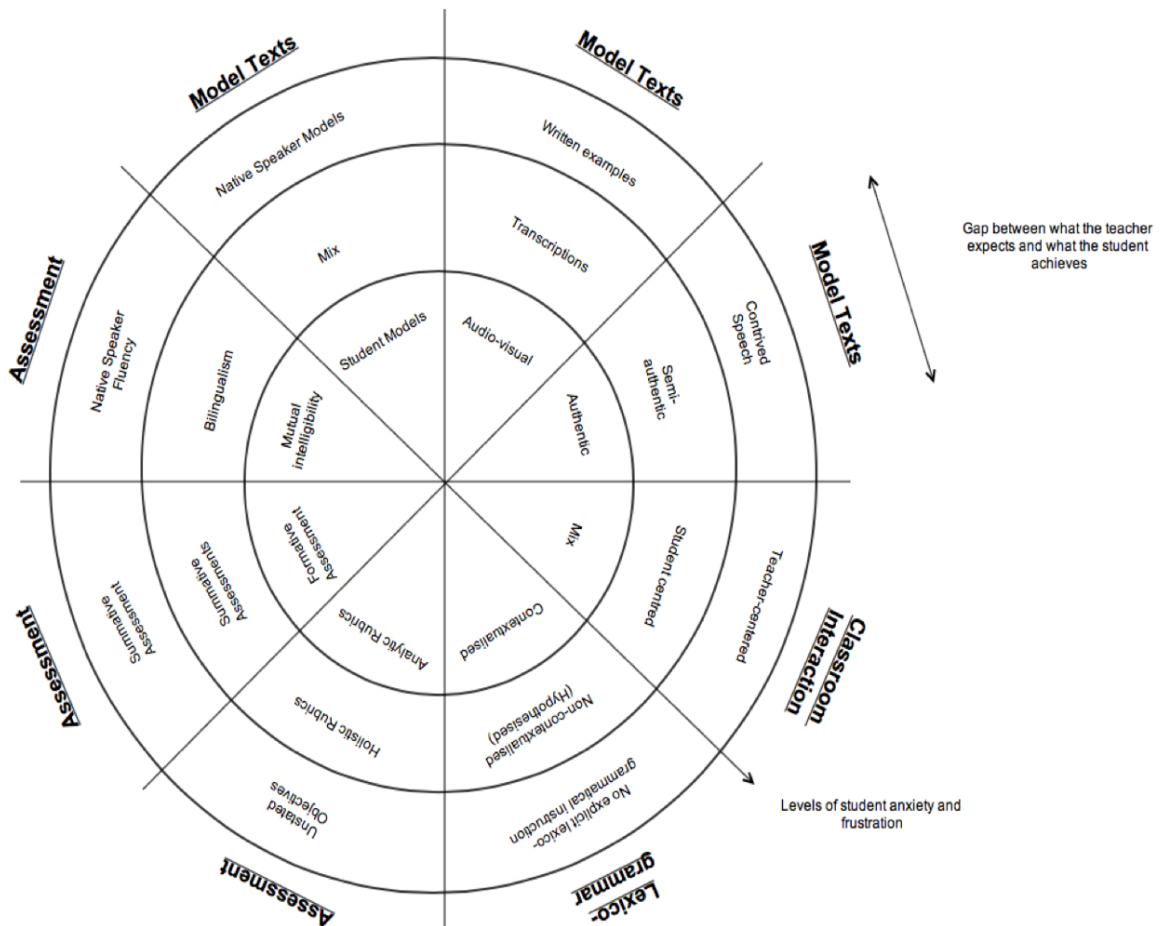


Figure 8.1 Model summarizing some of the main findings of the research

Figure 8.1 is an idealized abstraction within the teaching context of this study, illustrating that native speaker models or idealized written texts widened the gap between student and teacher expectations; and that this gap could be narrowed by using student models of spoken English that were authentic, but relied on mutual intelligibility rather than the concept of the native speaker. Such model texts immediately informed students of what was expected of them, and they reduced anxiety as well as illustrating learning goals. A hierarchy of student models based on a rubric of expected levels of pronunciation and fluency, as well as engagement with micro-aspects of casual conversation, also allowed students to scaffold their learning goals, and provided examples of what they were expected to achieve regarding otherwise abstract concepts. In this way, as Figure 8.1 shows, analytic

rubrics could be used to create explicit learning goals that could be negotiated with students through the deconstruction of model texts. During the deconstruction phase of classroom learning, a more teacher-orientated classroom could occur, which would later give way to student-centred learning. A mix of student- and teacher-orientated classrooms in this way led to optimum learning. Opportunities for teacher-centred classrooms also enabled explicit grammar instruction, which could be sequenced in a logical and contextualized manner, and thereby assist speaking. The general findings in this section address the research question identified in section 4.2; the next section applies these findings more explicitly.

8.3 The Research Question

In Chapter 1, issues concerned with the desire for increased oral proficiency in Japan were discussed, with the various problems associated with implementing communicative approaches in Japanese classrooms highlighted. The research question associated with this study was:

In what ways can a genre-based approach assist the teaching and development of Japanese students' speaking abilities?

In the reflection phases of the action research, I argued that the genre-based approach as adopted in this study provided a framework for sequencing of content in a principled manner, to address students' speaking abilities (section 7.7.2). The framework allowed a methodological approach that enabled the integration of explicit grammar instruction (section 7.7.1), as well as communicative activities that attempted to address spoken fluency (section 5.5, section 7.8 and Appendix 23). Such a syllabus can provide explicit aims and objectives that work towards improving students' spoken abilities and the integration of assessment for learning strategies (section 7.7). The use of whole texts is illustrated through models, that allow deconstruction, joint construction and self-construction of the generic structures of speaking; it also enabled students to speak at much greater length than they had done

previously, and with a greater understanding of the role of grammar in assisting fluency. The use of explicit grammar instruction and teacher-centred learning might go some way to addressing some of the cultural obstacles faced by communicative language teaching (section 3.4.4).

Such an approach also raised a number of issues, however. Micro-aspects of conversation, such as asking and answering questions, providing helpful information, expressing support or surprise, recasting information, and strategies for avoiding uncomfortable silences (section 4.8.2 and section 6.3.3), as well as pronunciation and fluency (section 5.3.2, section 6.3.3 and section 7.4), need careful consideration when planning speaking syllabi in similar contexts. These factors are exacerbated in an EFL context, where the only opportunities for spoken practice might be in the language classroom. Although students were able to lengthen utterances and engage in conversation more successfully by analysing generic structures of talk (section 5.5, section 7.8 and Appendix 23), this did not always prepare them for spontaneity in talk, particularly when talking with a teacher or for assessment purposes (section 7.8). The time demands of the genre-based approach in the classroom, including time for explicit instruction and processing of meta-language, reduced the time afforded to general practice, which they could also not achieve outside the classroom. In an EFL context, the demands of casual conversation and the time necessary to practise and experience authentic texts outside class is limited, and as such, the full possibilities of the genre-based approach may be harder to achieve. Due to the institutional environment, and the lack of motivation and opportunities to practise and listen to authentic conversations outside class, a focus on generic structures may become too prescriptive and not provide students with opportunities to experiment with language. In the present context, in order to fully utilize the genre-based approach, devoting far more curriculum time to English instruction would be extremely beneficial. The genre-based approach also requires a

great deal of teacher training and classroom planning, which are further demands on time in the curriculum (section 7.9.4).

Teacher training for such an approach would need to ensure the universal understanding of the main theoretical underpinnings of such an approach. In order to justify the amount of time required to plan and execute a genre-based approach on a wider scale, teachers must appreciate its value and its wider implications. A misunderstanding of the main theoretical perspectives may also lead to incorrect applications of such syllabi. The data in this study also suggest that a genre-based approach needs a considered and integrated approach to assessment, thus further necessitating the requirement for teacher training, and also increasing tensions associated with traditional assessment procedures in the research context.

The validity of the observations in this study and their applicability to the wider field also requires a discussion of the limitations of the study itself. Therefore, the next section discusses these limitations, before the final section discusses the wider implications for language teaching.

8.4 Limitations of the Findings

In this section, the limitations of the findings of the research are discussed. Due to the action research process, changes were made in the intervention, the assessment tools, and the research design itself, based on reflection and the observed demands of my students' learning. As a result of these changes, the value of triangulating various data sources over time and between groups is open to debate. Also, the participants in this study were low-proficiency, and the vast majority were Japanese students, which has implications for the generalizability of the research. The collection of data might also have benefited from the inclusion of questionnaires and interviews that aimed to focus more deeply on some of the findings garnered from reflective journals and classroom documents.

The study also focused solely on the recount genre, which represents only a fraction of the different types of genres that encompass casual conversation, all of which may overlap in a single conversation. In different genres, alternative or unique findings might have been observed, and different conclusions might have been investigated. The recount genre was chosen on the basis that it was the most applicable and useful genre to students, and thus may have inspired more positive results than might have been the case for other genres. Huge pedagogical advantages could be gained by comparing genres in far greater detail than was undertaken in this study, which focused on just one. Unfortunately, due to the proficiency levels of the students in this study, this would require multiple semesters. Another drawback in this context is that students change their English teacher each semester, which requires far greater collaboration with other teachers. With higher-proficiency students, it might be possible to compare multiple genres in one semester, and thus enable much greater pedagogical benefits of this approach.

Data analysis also relied on students' feedback in English, which limited the complexity of their possible responses. Although students were able to provide feedback in Japanese, which allowed a cross-reference with their responses in English in the final analysis, only the English feedback was used, as I did not feel confident enough in my Japanese language abilities to provide objective translations. In the quantitative analysis described in section 5.4 and section 7.3, sample sizes were limited, with the analysis in section 5.4 relying on the data for just 27 students. Although Rasch analysis does not demand the massive data samples required in other factor analyses, conclusions are weakened by such a small sample size.

The choice of action research as a methodology in this study was based on the research perspectives and my own beliefs; it also reflected my own specific teaching context and the problematization of that teaching context. Although the wider implications were

considered as a rationale for the research, other instructors may not find relevance to their own teaching situations.

8.5 Implications for English Language Teaching

As section 8.4 illustrates, the generalizability of findings in this study is open to debate; nonetheless, the findings aim to provide resonance for similar teaching situations, rather than universal truths about the classroom and language learning. The main implications, which I feel have resonance in a number of language teaching classrooms, are best illustrated in Figure 8.1. The first involves considerations regarding the modelling of the English language to our students. The authenticity of model texts must be considered in order to prepare students with the skills necessary for encountering real English outside the classroom; text authenticity demands that students are aware of the differences between spoken and written English, and the importance of grammar in organizing and sequencing spoken communication. Who is modelling the text must also be considered: notions of a native speaker appear detrimental to language learning, and raise issues of the power imbalance between English speakers, as well as the relative value of various Englishes around the world. As teachers, we need to ensure that we are empowering our students and providing them with language models that are within their capabilities. This calls for a re-evaluation of the native speaker model, and a focus upon mutual intelligibility. In an EFL context this is particularly important, as the teacher may be solely responsible for students' interactions with English; and teachers' beliefs about language and power are, therefore, significant.

Another implication concerns the choices teachers make when selecting and sequencing content for our students in a principled and logical manner. It is important to consider a methodological approach that draws on a range of teaching ideas; and not necessarily to conclude that there is a right and wrong way to teach speaking, but that different contexts and different students demand different approaches within an organized

framework. A principled approach does not reject traditional approaches, such as explicit grammar instruction, or a teacher-centred language classroom, but draws on the possible advantages of a range of approaches. The importance of observation and reflection on current teaching practices is shown via the action research processes that took place in the present study, where assumptions in the first cycle of action research were modified by reflection on systematically collected data.

The research also raised the importance of integrated assessment strategies, which should be an important part of the planning of a syllabus. Aims and objectives in speaking can often be abstract and difficult to make explicit to students; furthermore, opportunities for feedback and explicit sharing of learning goals integrated into syllabus design can assist learning by closing the gap between student and teacher perceptions of spoken performance and learning goals, and can reduce students' anxieties. Summative assessments should reflect stated classroom goals, and to prevent misunderstandings, feedback should be provided after these summative assessments.

8.6 Future Planning

The process of conducting an action research study was challenging and time-consuming. However, having systematically collected data and analysed the results of changes in my classroom, I believe the process was worthwhile, and it has renewed my appetite for further professional development. The process also provided valuable insights into students' experiences of my language classroom. I was able to observe improved spoken fluency in my students, as well as an overall feeling of satisfaction and pleasure in their study. Since beginning this study, I have engaged much more reflectively in my teaching, and considered the needs of my students more fully.

In planning Cycle Three of this action research process, I would attempt to address three main problems that remain with the syllabus intervention used in this study. The first is

the inclusion of multiple genres, which requires greater collaboration with other teachers as the students' progress through the university curriculum. This collaboration would also involve the creation of new syllabi encompassing other genres in casual conversation. The second problem to address would be issues with turn-taking and micro-aspects of conversation. This could be tackled by taking a methodological approach to syllabus design and introducing approaches from Conversational Analysis, such as those proposed by Wong and Waring (2010), who argue that Conversational Analysis provides a solid understanding of what constitutes "talk-in-interaction" (p. 2). They propose that approaches drawn from Conversational Analysis involve a comprehensive and systematic introduction to the basic features of turn-taking, without which, they argue, there is no social interaction. Finally, I would like to completely remove the teacher from any interaction with students during their conversation, particularly during assessment. I believe that this would eliminate feelings of nervousness and frustration and promote enjoyment and satisfaction, which is surely the primary social goal of casual conversation. This could be done via activities already proposed in section 7.7.4. To provide students with more English practice outside the classroom, I asked students to have conversations with parents and friends, and upload videos of the conversations to my website using their smartphones. Students could also use this process to have casual conversations with their classmates during a planned assessment period. As the teacher monitors and assists where necessary, students could conduct their final recount conversation in class in pairs, and record it on their smartphone to be uploaded to a secure website. In this way, the teacher is completely absent from the conversation, but is still able to view the conversation afterwards. This process would also allow additional teachers to view other students' conversations and provide assessments, thus addressing the issues of rater reliability raised in section 7.6.1. Sharing the results of this study with colleagues and peers in other teaching contexts, via attending international conferences and publishing in

international journals, would also help me to explore the implications of a genre-based approach in other contexts.

References

- Agger, B. (1991). Critical theory, poststructuralism, postmodernism: Their sociological relevance. *Annual review of sociology*, 17(1), 105-131.
- Allwright, D., & Hanks, J. (2009). *The developing language learner: An introduction to exploratory practice*. London, U.K.: Palgrave Macmillan.
- Anderson, J., Gibson, A., Howarth, M., Pritchard, T., Williams, A. & Wilkins, S. (2012). Practical classroom approaches to vocabulary instruction: a longitudinal experiment. *KOTESOL International Conference Proceedings*, 19, 295-300.
- Assessment Reform Group, UK (1999). *Assessment for learning: beyond the black box*. Cambridge, U.K.: University of Cambridge School of Education.
- Barnlund, D. C. (1989). *Communicative styles of Japanese and Americans: Images and realities*. Belmont, CA: Wadsworth Thomson Learning.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford, U.K.: Oxford University Press.
- Bazeley, P. (2009). Analysing qualitative data: More than 'identifying themes.' *Malaysian Journal of Qualitative Research*, 2(2), 6-22.
- Beglar, D. (2010). A Rasch-based validation of the Vocabulary Size Test. *Language testing*, 27(1), 101-118.
- Benson, P. (2013). *Teaching and researching: Autonomy in language learning*. 2nd edition. Oxford, U.K.: Routledge.
- Black, P. (2009). Formative assessment issues across the curriculum: The theory and the practice. *TESOL Quarterly*, 43(3), 519-524.
- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7-74.
- Blanche, P., & Merino, B. J. (1989). Self-assessment of foreign-language skills: Implications for teachers and researchers. *Language learning*, 39(3), 313-338.
- Bond, T. & Fox, C. (2007). *Applying the Rasch model*. 2nd edition. London, U.K.: Lawrence Erlbaum Associates.
- Boud, D. (1999). Situating academic development in professional work: Using peer learning. *The International Journal for Academic Development*, 4(1), 3-10.
- Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher education*, 18(5), 529-549.
- Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31(4), 399-413.
- Boud, D. & Falchikov, N. (2007). *Rethinking assessment in higher education: Learning for the longer term*. Oxford, U.K.: Routledge.
- Boud, D. & Holmes, W. H. (1981). Self and peer marking in an undergraduate engineering course. *IEEE Transactions on Education*, 24(4), 267-274.
- Bonk, W. J. & Ockey, G. J. (2003). A many-facet Rasch analysis of the second language group oral discussion task. *Language Testing*, 20(1), 89-110.
- Borg, S. (2006). *Classroom research in English language teaching in Oman*. Sultanate of Oman: Ministry of Education.
- Borg, S. (2008). *Investigating English language teaching and learning in Oman*. Sultanate of Oman: Ministry of Education.
- Brown, H. D. (2003). Interviewer variation and the co-construction of speaking proficiency. *Language Testing*, 20(1), 1-25.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. White Plains, NY: Longman.
- Brown, J. D. (2005). *Testing in language programs: A comprehensive guide to English language assessment* (New Ed.). New York, NY: McGraw-Hill.

- Brown, H. D. (2007). *Teaching by principles: An interactive approach to language pedagogy*. 3rd edition. New York, NY: Pearson Education.
- Burns, A. (2005). Action research: an evolving paradigm? *Language Teaching*, 38, 57–74
- Burns, A. (2011). *Doing action research in English language teaching: A guide for practitioners*. Beijing: Foreign Language Teaching and Research Press.
- Butler, D. L. & Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research*, 65(3), 245–281.
- Butt, D., Fahey, R., Spinks, S. & Yallop, C. (2000). *Using functional grammar: An explorer's guide*. Sydney, NSW: NCELTR.
- Callaghan, M., Knapp, P. & Noble, G. (2014). Genre in practice. *The powers of literacy: A genre approach to teaching writing*, 179–202. Pittsburgh, PA: University of Pittsburgh Press
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied linguistics*, 1(1), 1–47.
- Carter, R. & McCarthy, M. (2006). Ten criteria for a spoken grammar. *Explorations in corpus linguistics*, 27, 27–52.
- Central Education Council. (1999). Improvement of admissions with an emphasis on the connection with the higher education and elementary and secondary education. Retrieved from <http://www.mext.go.jp/b_menu/shingi/old_chukyo/old_chukyo_index/toushin/attach/1309753.htm>.
- Chappell, C. (2007). *Research Perspectives*. Sydney, NSW: University of Technology.
- Charmaz, K. (2006). *A practical guide through qualitative analysis*. London, U.K.: Sage.
- Cheung, K. C. (1991). *Meaningful Assessment of Problem-Solving Activities in the Classroom: Some Exemplars*. Research Monograph No. 2. Singapore: National Institute of Education.
- Chomsky, N. (1967). Recent contributions to the theory of innate ideas. *Synthese*, 17(1), 2–11.
- Clark, U. (2013). *Language and identity in Englishes*. London, U.K.: Routledge.
- Cohen, L., Manion, L. & Morrison, K. (2013). *Research Methods in Education*. London, U.K.: Routledge.
- Cohen, D. K. & Spillane, J. P. (1992). *Policy and practice: The relations between governance and instruction*. *Review of Research in Education*, 18, 3–49.
- Collerson, J. (1994). *English grammar: A functional approach*. Sydney, NSW: Primary English Teaching Association.
- Collinson, D. (2006). Rethinking followership: A post-structuralist analysis of follower identities. *The Leadership Quarterly*, 17(2), 179–189.
- Cook, I. & Crang, M. (2007). *Doing Ethnographies*. London, U.K.: Sage.
- Coulmas, F. (2007). *Population decline and ageing in Japan – the social consequences*. London, U.K.: Routledge.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching and Assessment*. Cambridge, U.K.: Cambridge University Press.
- Council of Europe (2008). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, U.K.: Cambridge University Press.
- Cresswell, J. W. (1994). *Research Design: Qualitative and Quantitative approaches*. London, U.K.: Sage.
- Cresswell, J. W. (1998). *Qualitative Inquiry and Research Design Choosing Among Five Traditions*. London, U.K.: Sage.
- Cresswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches*. London, U.K.: Sage.
- Crotty, M. (2008). *The Foundations of Social Research: Meaning and Perspective in the Research Process*. London, U.K.: Sage.
- Crystal, D., & Davy, D. (1975). *Advanced Conversational English*. London, U.K.: Longman Publishing Group.
- Cumming, A. (2009). What needs to be developed to facilitate classroom-based assessment? *TESOL Quarterly*, 43, 515–519.

- Cumming, A., Kantor, R. & Powers, D. (2002). Decision making while rating ESL/EFL writing tasks: A descriptive framework. *Modern Language Journal*, 86, 67–96.
- Cummings, M. (2005). The role of Theme and Rheme in contrasting methods of organization in texts. *Pragmatics and Beyond New Series*, 140, 129.
- Dávila, A., Bohara, A. & Saenz, R. (1993). Accent penalties and the earnings of Mexican Americans. *Social Science Quarterly*, 74, 902–915.
- Davis, J. M. (2004) Writing an action research thesis: One researcher's resolution of the problematic of form and process. In E. McWilliam, S. Danby, & J. Knight (Eds.), *Performing Educational Research: Theories, Methods and Practices*, 15–30. Flaxton, Qld: Post Pressed.
- Davis, J. M. (2007) Rethinking the architecture: An action researcher's resolution to writing and presenting their thesis. *Action Research* 5(2), 181–198.
- Davison, C. (2008). *Assessment for learning: Building inquiry-oriented assessment communities*. Paper presented at 42nd Annual TESOL Convention and Exhibit, New York.
- Davison, C. & Leung, C. (2009). Current issues in English Language Teacher-Based assessment. *TESOL Quarterly*, 43, 393–415.
- Denscombe, M. (1998). *The Good Research Guide for Small-Scale Social Research Projects*. London, U.K.: Open University Press.
- Derwing, T. & Munro, M. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39, 379–797.
- Derwing, T. M., Rossiter, M. J. & Munro, M. J. (2002). Teaching native speakers to listen to foreign-accented speech. *Journal of Multilingual and Multicultural Development*, 23, 245–259.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford, U.K.: Oxford University Press.
- Edge, J. & Richards, K. (1993). *Teachers develop, teachers research: Papers on classroom research and teacher development*. Oxford, U.K., : Heinemann.
- Educational Testing Service. (2012). *TOEIC Program data and analysis 2012*. Retrieved from <http://www.toeic.or.jp/library/toeic_data/toeic_en/pdf/data/TOEIC_Program_DAA2012.pdf>.
- Eggs, S. (2004). *An introduction to systemic functional linguistics*. London, U.K.: Pinter.
- Eggs, S. & Slade, D. (1997). *Analysing Casual Conversation*. London, U.K.: Cassell.
- Ellis, G. (1996). How culturally appropriate is the communicative approach? *ELT Journal*, 20, 213–218.
- Ellis, R. (1985). Sources of variability in interlanguage. *Applied Linguistics*, 6, 118–131.
- Ellis, R. & Barkhuizen, G. (2005). *Analysing learner language*. Oxford, U.K.: Oxford University Press.
- Engelhard, G., Jr. (2009). Using item response theory and model-data fit to conceptualize differential item and person functioning for students with disabilities. *Educational and Psychological Measurement*, 69(4), 585–602.
- Ertmer, P. A. & Newby, T. J. (1993). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, 6(4), 50–72.
- Farrell, S., Bannister B., Ditton, J., & Gilchrist, E. (1997). Open and closed questions. *Social Service Update*, 17. Guildford, U.K.: University of Surrey.
- Feez, S. (2002). Heritage and innovation in second language education. In A. M. Johns (Ed.), *Genre in the classroom: Multiple perspectives*, 43–69. Mahwah, NJ; London, U.K.: Lawrence Erlbaum Associates.
- Feez, S. & Joyce, H. (1998). *Text-based syllabus design*. Sydney, NSW: NCELTR.
- Fennelly, M. & Luxton, R. (2011). Are they ready? On the verge of compulsory English, elementary school teachers lack confidence. *The Language Teacher*, 35(2), 19–24.
- Firth, J. R. (1950). Personality and language in society. *The sociological review*, 42(1), 37–52.
- Flege, J. E. (1988). The production and perception of foreign language speech sounds. In H. Winitz. (Ed.), *Human communication and its disorders: A review*, 224–401. Norwood, NJ: Ablex.
- Freire, P. (1976). *Education, the practice of freedom*. London, U.K.: Writers and Readers Collective.

- Fryer, L. K. & Ozono, S. (2014). LERC Report 2013. *Kyushu Sangyo University Journal*, 9, 50–60.
- Fryer, L. K., Ozono, S., Carter, P., Nakao, K. & Anderson, C. J. (2014). Instrumental reasons for studying in compulsory English courses: I didn't come to university to study English, so why should I? *Innovation in language learning and teaching*, 8(3), 239–256.
- Geuss, R. (1981). *The idea of a critical theory: Habermas and the Frankfurt School*. Cambridge, U.K.: Cambridge University Press.
- Gibbons, P. (1991). *Learning to Learn in a Second Language*. Portsmouth, NH: Heinemann.
- Gibbons, P. (2001). *Curriculum and methodology of TESOL*. Sydney, NSW: University of Technology.
- Glaser, B. G. & Strauss, A. L. (1967). *The Discovery of Grounded Theory*. Chicago, IL: Aldine.
- Goh, C. & Burns, A. (2012). *Teaching speaking: A holistic approach for English language learners*. New York, NY: Cambridge University Press.
- Gorsuch, G. J. (2000). EFL educational policies and educational cultures: influences on teachers' approval of communicative activities. *TESOL Quarterly*, 34(4), 675–710.
- Grasha, A. F. (1994). A matter of style: The teacher as expert, formal authority, personal model, facilitator, and delegator. *College Teaching* 42, 142–9.
- Greenwood, D. J., & Levin, M. (2005). Reform of the social sciences and of universities through action research. *The Sage handbook of qualitative research*, 3, 43–64.
- Griffiths, M. & Davies, G. (1993). Learning to Learn: action research from an equal opportunities perspective in a junior school. *British Educational Research Journal*, 19, 43–58.
- Grisson, J. A. & Youngs, P. (2016). *Improving teacher evaluation systems: Making the most of multiple measures*. New York, NY: Teachers College Press.
- Halliday, M. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London, U.K.: Hodder Arnold.
- Halliday, M. (1985). *Spoken and Written Language*. Waurin Ponds, Vic: Deakin University Press.
- Halliday, M. (1994). *Introduction to functional grammar*. 2nd edition. London, U.K.: Edward Arnold.
- Halliday, M. & Hasan, R. (1976). *Cohesion in English*. London, U.K.: Longman.
- Halliday, M. & Hasan, R. (1985). *Language, context and text: Aspects of language in a social semiotic perspective*. Waurin Ponds, Vic: Deakin University Press.
- Halliday, M. A. K. & Matthiessen, C. (2006). *Construing experience through meaning: A language-based approach to cognition*. London, U.K.: A&C Black.
- Hammersley, M. & Atkinson, P. (1983). *Ethnology: Principles in practice*. New York, NY: Tavistock.
- Harmer, J. (2007). *The practice of English language teaching*. Harlow, U.K.: Pearson Longman.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81–112.
- Herron, C. A. & Seay, I. (1991). The effect of authentic oral texts on student listening comprehension in the foreign language classroom. *Foreign Language Annals*, 24(6), 487–495.
- Hill, K. & Sabet, M. (2009). Dynamic speaking assessments. *TESOL Quarterly*, 43, 537–45.
- Holster, T. A. (2008). How High Can They Jump: An introduction to Rasch Measurement. *文芸と思想*, (78), 148–122.
- Holster, T. A. & Lake, J. W. (2015). From raw scores to Rasch in the classroom. *Shiken*. 19(1), 32–41.
- Hood, S. (2000). *Language programming and assessment*. Sydney, NSW: University of Technology.
- Horkheimer, M. (1972). *Critical Theory, Selected Essays*. New York, NY: Continuum.
- Hu, G. (2002). Potential cultural resistance to pedagogical imports: the case of communicative language teaching in China. *Language, Culture and Curriculum*, 15(2), 93–105.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge, U.K.: Cambridge University Press.
- Husserl, E. (1970). *The crisis of European sciences and transcendental phenomenology: An introduction to phenomenological philosophy*. Evanston, IL: Northwestern University Press.

- Hymes, D. (1967). Models of the interaction of language and social setting. *Journal of Social Issues*, 23(2), 8–28.
- Hymes, D. (1972). On communicative competence. *Sociolinguistics*, 269–293.
- Hyon, S. (1996). Genre in three traditions: Implications for ESL. *TESOL quarterly*, 30(4), 693–722.
- Ishihara, N. (2009). Teacher-based assessment for foreign language pragmatics. *TESOL Quarterly*, 43, 445–470.
- Japan Society for the Promotion of Science (2012). *KAKENHI Grants-in-Aid for Scientific Research*. Retrieved from <http://www.jsps.go.jp/english/e-grants/data/kakenhi_pamph_e.pdf>.
- Jenkins, J. (2006). Global intelligibility and local diversity: possibility or paradox? In R. Rubdi & M. Saroceni (Eds.), *English in the world: Global rules, global roles*, 32–39. London, U.K.: Continuum.
- John, W. and Johnson, P. (2000). The Pros and Cons of Data Analysis Software for Qualitative Research. *Journal of Nursing Scholarship*, 32(4), 393–397.
- John, W. & Johnson, P. (2004). The pros and cons of data analysis software for qualitative research. *Journal of Nursing Scholarship*, 32(4), 393–397.
- Johnson, K. (1981). Introduction: some background, some key terms and some definitions. In K. Johnson & K. Morrow (Eds.), *Communication in the classroom*, 1–12. London, U.K.: Longman.
- Johnson, R. B. & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher*, 33(7), 14–26.
- Joyce, H. & Burns, A. (1999). *Focus on grammar*. Sydney, NSW: NCELTR.
- Joyce, H. & Slade, S. (2000). The nature of casual conversation: Implications for teaching. *Teachers' Voices* 6, viii–xv.
- Kashima, E. S. & Kashima, Y. (1998). Culture and language: The case of cultural dimensions and personal pronoun use. *Journal of Cross-Cultural Psychology*, 29(3), 461–486.
- Kemmis, S. & McTaggart, R. (1988). *The action research planner*. Waurin Ponds, Vic: Deakin University Press.
- Kennedy, K. J., Chan, K. S. J., Yu, W. M. & Fok, P. K. (2006). *Assessment for productive learning: Forms of assessment and their potential for enhancing learning*. Paper presented at the 32nd Annual Conference of the International Association for Educational Assessment, Singapore.
- Kincheloe, J. L., & McLaren, P. (2011). Rethinking critical theory and qualitative research. In K. Hayes, S. Steinberg & K. Tobin (Eds.), *Key works in critical pedagogy*, 285–326. Leiden: SensePublishers.
- Kinginger, C. (2000). Learning the pragmatics of solidarity in the networked foreign language classroom. In J. Hall & L. Verplaetse (Eds.), *Second and foreign language learning through classroom interaction*, 23–46. London, U.K.: Routledge.
- Klopf, D. W. (1991). Japanese communication practices: Recent comparative research. *Communication Quarterly*, 39(2), 130–143.
- Kubota, R. (2011). Questioning linguistic instrumentalism: English, neoliberalism, and language tests in Japan. *Linguistics and Education*, 22(3), 248–260.
- Kunyot, T. (2005). *Developing the English oral competency of tourism industry students by means of a genre-based ESP approach*. Perth, WA: Doctoral Thesis, Edith Cowan University.
- Law, G. (1995). Ideologies of English language education in Japan. *JALT Journal*, 17, 213–224.
- Laufer, B. & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, 54(3), 399–436.
- Lerner, G. (1986). *The Creation of Patriarchy*. Oxford, U.K.: Oxford University Press.
- Linacre, J. M. (2007a). *Facets*. Retrieved from <<http://www.winsteps.com/facets.htm>>.
- Linacre, J. M. (2007b). *Winsteps*. Retrieved from <<http://www.winsteps.com/>>.
- Lincoln, Y. & Guba, E. (1985). *Naturalistic Inquiry*. Beverly Hills, CA: Sage.
- Lock, G. (1996). *Functional English grammar: An introduction for second language teachers*. Cambridge, U.K.: Cambridge University Press.

- Lunz, M. E. & Stahl, A. (1993). The impact of rater severity on person ability measure: A Rasch model analysis. *American Journal of Occupational Therapy*, 47, 311–317.
- Luke, C. & Gore, J. (2014). *Feminisms and critical pedagogy*. London, U.K.: Routledge.
- Malinowski, B. (1923). The problem of meaning in primitive languages. In J. Maybin (Ed.), *Language and literacy in social practice: A reader*, 1–11. Clevedon, U.K.: Multilingual Matters.
- Malson, H. (2003). *The thin woman: Feminism, post-structuralism and the social psychology of anorexia nervosa*. London, U.K.: Routledge.
- Martin, J. R. (1984). Language, Register and Genre. In F. Christie (Ed.), *Children Writing: A Reader*, 21–29. Warrn Ponds, Vic: Deakin University Press.
- Martin, J. R. (1992). *Macro-genres: The Ecology of the Page*. Sydney, NSW: University of Sydney.
- Martin, J. R. & Rothery, J. (1986). What a functional approach to the writing task can show teachers about ‘good writing’. In B. Couture (Ed.), *Functional approaches to writing: Research perspectives*, 241–265. London, U.K.: Burns & Oates.
- Martin, J. R. & Rothery, J. (1993). Grammar: Making meaning in writing. In B. Cope & M. Kalantzis (Eds.), *The powers of literacy: A genre approach to teaching writing*, 137–153. Pittsburgh, PA: University of Pittsburgh Press.
- Massironi, M. (2001). *The psychology of graphic images: Seeing, drawing, communicating*. Hove, U.K.: Psychology Press.
- Matsuura, H., Chiba, R., & Hilderbrandt, P. (2001). Beliefs about learning and teaching communicative English in Japan. *Japan Association for Language Teaching Journal*, 23(1), 67–82.
- McDonald, B. & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education: Principles, Policy & Practice*, 10(2), 209–220.
- McDonough, J. & McDonough, S. (1997). *Research Methods for English Language Teachers*. London, U.K.: Arnold.
- McDonough, J. & Shaw, C. (1993). The impact of the communicative approach, in *Materials and Methods in ELT: a Teacher’s Guide*. Hoboken, NJ: Blackwell.
- McNamara, T. F. (1996). *Measuring second language performance*. Harlow, U.K.: Pearson Education.
- Merrett, F., & Wheldall, K. (1992). Teachers’ use of praise and reprimands to boys and girls. *Educational Review*, 44(1), 73–79.
- Meyer, K., & Woodruff, E. (1997). Consensually driven explanation in science teaching. *Science Education*, 81(2), 173–192.
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. London, U.K.: Sage.
- Mills, C. W. (1963). *Power, politics, and people: The collected essays of C. Wright Mills*. Oxford, U.K.: Oxford University Press.
- Ministry of Education, Culture, Sports, Science and Technology, Japan (1947). *Gakushu shidou yoryu eigohen* [Study of course guideline for English education]. Retrieved from <www.nier.go.jp/guideline>.
- Ministry of Education, Culture, Sports, Science and Technology, Japan (1998). *Junior high schools’ courses of study*. Tokyo: National Printing Bureau.
- Ministry of Education, Culture, Sports, Science and Technology, Japan (1999). *Senior high schools’ courses of study*. Tokyo: National Printing Bureau.
- Ministry of Education, Culture, Sports, Science and Technology, Japan (2003a). *Action plan to cultivate “Japanese with English abilities”*. Retrieved from <http://www.mext.go.jp/english/topics/03072801.htm>.
- Ministry of Education, Culture, Sports, Science and Technology, Japan (2003b). *The course of study for foreign languages*. Retrieved from <http://www.mext.go.jp/english/shotou/030301.htm>.

- Ministry of Education, Culture, Sports, Science and Technology, Japan (2005). *References*. Retrieved from <<http://www.mext.go.jp/english/statist/05101901/008.pdf>>.
- Ministry of Education, Culture, Sports, Science and Technology, Japan (2007). *A survey for practical situation of elementary school English activities in 2006*. Retrieved from <http://www.mext.go.jp/b_menu/houdou/19/03/07030811/004.htm>.
- Ministry of Education, Culture, Sports, Science and Technology. (2008a). *Shogakkou gakushu shidou yoryo kaisetsu gaikokugo katsudouhen* [Description for study of course guideline for foreign language activities in elementary schools]. Retrieved from <www.mext.go.jp/a_menu/shotou/new-cs/youryou/syokaisetsu/index.htm>.
- Ministry of Education, Culture, Sports, Science and Technology (2008b). *Higher Education in Japan*. Retrieved from <http://www.mext.go.jp/english/highered/_icsFiles/afieldfile/2012/06/19/1302653_1.pdf>.
- Ministry of Education, Culture, Sports, Science and Technology (2009). *Koutougakkou gakushu shidou yoryo gaikokugo eigoban kariyaku* [Study of course guideline for foreign languages in senior high schools; provisional version]. Retrieved from <www.mext.go.jp/a_menu/shotou/new-cs/youryou/eiyaku/1298353.htm>.
- Ministry of Education, Culture, Sports, Science and Technology (2010). *Koritsu kotogakko ni okeru kyoi- kukatei no hensei jissai chosa no kekka (B Hyo)* [Survey of English education in public high schools (B)]. Retrieved from <www.mext.go.jp/a_menu/shotou/new-cs/1301650.htm>.
- Ministry of Education, Culture, Sports, Science and Technology (2011). *Kokusai kyotsugo toshite no eigoryoku kojo no tameno itsutsu no teigen togutaiteki shisaku* [Five Proposals to Improve the Proficiency of English as Lingua Franca]. Retrieved from <www.mext.go.jp/b_menu/houdou/23/07/1308888.htm>.
- Mitchell, R. & Lee, J. H. W. (2003). Sameness and difference in classroom learning cultures: interpretations of communicative pedagogy in the UK and Korea. *Language Teaching Research*, 7(1), 35–63.
- Moore, N. (2009). 12 Advanced language for intermediate learners: corpus and register analysis for curriculum specification in English for Academic Purposes. In H. Byrnes (Ed.), *Advanced Language Learning: The Contribution of Halliday and Vygotsky*, 246. New York, NY: Continuum.
- Mulvey, B. (2010). University accreditation in Japan: Problems and possibilities for reforming EFL education. *The Language Teacher*, 34(1) 15–24.
- Munro, M. J. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal*, 20, 38–51.
- Munro, M. J. & Derwing, M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49, 285–310.
- Neustupný, J. V. (1985). Problems in Australian-Japanese contact situations. In J. Pride (Ed.) *Cross-cultural encounters: communication and miscommunication*, 44–84. Melbourne, Vic: River Seine Publications.
- Nicholson, P. & Butterworth, A. (2000). *Listening to Australia*. Sydney, NSW: NSW Adult Migrant English Service.
- Nishino, T. (2011). Japanese high school teachers' beliefs and practices regarding Communicative Language Teaching. *JALT Journal*, 33(2), 131–155.
- Nishino, T. & Watanabe, M. (2008). *Communication-oriented policies versus classroom realities in Japan*. *TESOL Quarterly*, 42(1), 133–138.
- Noonan, B. & Duncan, C. R. (2005). Peer and self- assessment in high schools. *Practical Assessment, Research & Evaluation*, 10(17).
- Nunan, D. (1992). *Research Methods in Language Teaching*. New York, NY: Cambridge University Press.
- Nunan, D. (2002). Listening in language learning. In W. Renandya & J. Richards (Eds.), *Methodology in language teaching: An anthology of current practice*, 238–241. Cambridge, U.K.: Cambridge University Press.

- Okazaki, S. (1994). Stating opinions in Japanese: Listener-dependent strategies. In J. Alatis (Eds.), *Georgetown University roundtable on languages and linguistics*, 69–95. Washington, D.C.: Georgetown University
- Packer, M. J. (1989). *Entering the circle: Hermeneutic investigation in psychology*. New York, NY: Suny Press.
- Parsons, H. L. (1971). *Humanism and Marx's thought*. Springfield, MA: Charles C. Thomas.
- Patniz, T. (1996). *Collaborative Versus Cooperative Learning: A comparison of the two concepts which will help us understand the underlying nature of interactive learning*. Retrieved from <<http://home.capecod.net/~tpanitz/tedsarticles/coopdefinition.htm>>.
- Paul, D. (2003). *Teaching English to children in Asia*. Hong Kong: Longman Asia.
- Powell, S. D. & Makin, M. (1994). Enabling pupils with learning difficulties to reflect on their own thinking. *British Educational Research Journal*, 20, 579–585.
- Poynton, C. (1985). *Language and Gender: Making the Difference*. Waurin Ponds, Vic: Deakin University Press.
- Poynton, C. (1990). *Address and the semiotics of social relations: A systemic-functional account of address forms practices in Australian English*. Sydney, NSW: PhD thesis, University of Sydney.
- Preston, D. R. (2007). The uses of folk linguistics. *International Journal of Applied Linguistics*, 3(2), 181–259.
- Rapley. (2008). *Policy and reality: The teaching of oral communication by Japanese teachers of English in public junior high schools in Kurashiki City, Japan*. Retrieved from <<http://www.asian-efl-journal.com/Thesis-D-Rapley.pdf>>.
- Reid, J. M. (1995). *Learning styles in the ESL/EFL classroom*. Boston: Heinle & Heinle.
- Richards, J. & Lockhart, C. (1996). *Reflective teaching in second language classrooms*. Cambridge, U.K.: Cambridge University Press.
- Richards, J., Platt, J. & Weber, H. (1985). *Longman dictionary of applied linguistics*. London, U.K.: Longman.
- Richards, J. C. & Rodgers, T. (1986). *Approaches and methods in language teaching: A description and analysis*. Cambridge, U.K.: Cambridge University Press.
- Rivera, J. D. H. (2012). Using genre-based approach to promote oral communication in the Colombian English classroom. *Colombian Applied Linguistics Journal*, 14 (2), 109–126.
- Robertson, I. (2004). *Teachers at the interface: A model of implementation*. Paper presented at the AARE conference, Melbourne.
- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation*, 11(10), 1–13.
- Rothery, J. (1996). Making the changes: developing an educational linguistics. In R. Hasan & G. Williams (Eds.), *Literacy in society*, 86–123. London, U.K.: Longman.
- Ryan, J. (2000). *A guide to teaching international students*. Oxford, U.K.: Oxford Centre for Staff and Learning Development.
- Ryan, R. M. & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, 25(1), 54–67.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144.
- Sakui, K. (2004). Wearing Two Pairs of Shoes: Language Teaching in Japan. *ELT Journal*, 58(2), 155–163. Retrieved from <biblioteca.uqroo.mx/hemeroteca/elt_journal/2004/abril/580155.pdf>.
- Schiellerup, P. (2007). Stop making sense: the trials and tribulations of qualitative data analysis. *Area*, 40(2), 163–171.
- Schön, D. (1983). *The Reflective Practitioner: How professionals think in action*. London, U.K.: Temple Smith
- Scott, D. & Usher, R. (1996). *Understanding Educational Research*. London, U.K.: Routledge.
- Scriven, M. (1991). Pros and cons about goal-free evaluation. *Evaluation Practice*, 12, 55–76.

- Scrivener, J. (1994). *Learning teaching*. Oxford, U.K.: Heinemann.
- Slade, D. (1997). Stories and Gossip in English: The Macro-Structure of Casual talk. *Prospect*, 12(2), 43–71.
- Slade, D. & Widin, J. (2004). Teaching spoken English in the secondary school classroom. *The Zen-ei-ren magazine*, 42, 2–13.
- Smith, L. & Nelson, C. (1985). International intelligibility of English: Directions and resources. *World Englishes*, 4, 333–342.
- Sorell, T. (1991) *Scientism: Philosophy and the Infatuation with Science*. London, U.K.: Routledge.
- Spindler, L. (2014). *Interpretive ethnography of education at home and abroad*. Hove, U.K.: Psychology Press.
- Stapleton, A. J. & Taylor, P. C. (2004). *Representing the flow of R&D in a thesis: Diachronic structure & hyperlinking*. Paper presented at the 2004 Australian Association for Research in Education (AARE) International Education Research Conference, Parkville, Victoria.
- Strauss, A. & Corbin, J. (1998). *Basics of qualitative research*. 2nd edition. Newbury Park, CA: Sage.
- Stroh, M. (2000). Computers and qualitative data analysis: to use or not to use? In D. Burton (Ed.), *Research training for social scientists*, 226–243. London, U.K.: Sage.
- Swales, J. (1986). Citation analysis and discourse analysis. *Applied linguistics*, 7(1), 39–56.
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge, U.K.: Cambridge University Press.
- Tahira, M. (2012). Behind MEXT's new Course of Study Guidelines. *The Language Teacher*, 36(3), 3–8.
- Takayama, K. (2008). The politics of international league tables: PISA in Japan's achievement crisis debate. *Comparative Education*, 44, 387–407.
- Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2), 237–246.
- Ting-Toomey, S. (1986). Japanese communication patterns: Insider versus the outsider perspective. *World Communication*, 15, 113–126.
- Tobin, K. (2000). Interpretive research in science education. In A. E. Kelly and R. A. Lesh (Eds.), *Handbook of Research Design in Mathematics and Science Education*, 487–512. London, U.K.: Lawrence Erlbaum Associates, Publishers.
- Tollefson, J. (2000) Policy and ideology in the spread of English. In J. K. Hall and W. Egginton (Eds.), *The sociopolitics of English language teaching*, 7–21. Clevedon, U.K.: Multilingual matters.
- Tsuda, Yukio (1994). The diffusion of English: Its impact on culture and communication. *Keio Communication Review*, 16, 49–61.
- Van Lier, L. (1996). *Interaction in the language curriculum: Awareness, autonomy and authenticity*. London, U.K.: Longman.
- Vandergrift, L. (2004). 1. Listening to Learn or Learning to Listen? *Annual Review of Applied Linguistics*, 24, 3–25.
- Wagner, C. (1994). *Placing psychology: a critical exploration of research methodology curricula in the social sciences*. Retrieved from <[http:// upetd.up.ac.za/thesis/available/etd-06292004-123737/](http://upetd.up.ac.za/thesis/available/etd-06292004-123737/)>.
- Wallace, M. (1998). *Action research for language teachers*. Cambridge, U.K.: Cambridge University Press.
- Watanabe, Y. (2004). Teacher factors mediating washback. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and models*, 129–146. Mahwah, NJ: Erlbaum.
- Watson-Gegeo, K. A. (1988). Ethnography in ESL: Defining the essentials. *Tesol Quarterly*, 22(4), 575–592.
- Weber, M. (1978). *Economy and society: An outline of interpretive sociology*. Berkeley, CA: University of California Press.

- Welsh, E. (2002). Dealing with Data: Using Nvivo in the Qualitative Data Analysis Process. *Qualitative Social Research*, 3(2). Retrieved from < http://www.qualitative-research.net/index.php/fqs/article/view/865/1880&q=nvivo+manual&sa=x&ei=zah_t5pqoyubhqfe9swgbq&ved=0cc4qfjaj>.
- Werner, O. & Schoepfle, G. M. (1987). *Systematic Fieldwork: Foundations of ethnography and interviewing*. Beverly Hills, CA: Sage.
- Wetzel, P. J. (1988). *Are powerless communication strategies the Japanese norm? Language in Society*, 17(4), 555–564.
- Wiersma, W. (2000). *Research Methods in Education: An Introduction*. Boston, MA: Allyn and Bacon.
- Wiggins, G. P. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. San Francisco, CA: Jossey Bass.
- Wignell, P., Martin, J. R. & Eggins, S. (1989). The discourse of geography: Ordering and explaining the experiential world. *Linguistics and Education*, 1(4), 359–391.
- Wilkins, S. G. (2006). Language in Context: Text-based Language Learning. *Nagasaki University of Foreign Studies Journal*, 10, 343–354.
- Williams, M. (2000). Interpretivism and generalisation. *Sociology*, 34(2), 209–224.
- Wilson, M. (2005). *Constructing measures: An item response modelling approach*. Mahwah, NJ: Erlbaum Associates.
- Wong, J., & Waring, H. Z. (2010). *Conversation analysis and second language pedagogy: A guide for ESL/EFL teachers*. London, U.K.: Routledge.
- Woolfolk, A. E., Rosoff, B. & Hoy, W. K. (1990). Teachers' sense of efficacy and their beliefs about managing students. *Teaching and Teacher Education*, 6, 137–148.
- Yamaoka, K. (2010). The current situation and the future prospects of Japanese English teaching. *Ritsumekan Language Studies*, 22, 59–66.
- Young, M. F. D. (1970). *Knowledge and Control: New Directions for the Sociology of Education*. London, U.K.: Collier Macmillan.
- Zuber-Skerrit, O. (2002). Action research within organisations and university thesis writing. *The Learning Organization*, 9(4), 171–179.

Appendices

Appendix 1: Activity to raise awareness of differences between genres

Match the sentence or phrase with the genre

Stop	e-mail
Add a tsp of salt	Advertisement
Once upon a time	Newspaper
To whom it may concern,	Road sign
Man Killed in Accident	Label
Hey Paul! :)	Instruction manual
Clouds of cherry blossoms	Recipe
Before using this camera, please check the following . . .	Menu
Don't forget to take the garbage out!	Formal Letter
menthol2.0g	Dictionary
Hamburger ¥ 2 0 0	Website
Buy one today!	Fairy Tale
Click here	Poem
	Memo
—noun	
1.	
a young swine of either <u>sex</u> , especially a domestic hog	

Appendix 2: Syllabus entry Form

Syllabus Entry Form

Class name: English Conversation

- Class aims:
- To understand the concept of genre and genres for different purposes.
 - To be able to identify the structural and lexico-grammatical features of a spoken recount text.
 - To be able to construct their own spoken recount text

Weekly
class
structure
(Goals):

Building the context

1. Teacher and student introductions (including spoken recount).
2. Match the genres worksheet (movie / music / written genres). How do we identify a genre? Introduction of spoken genres.
3. Greetings. The types of greetings used depending on the roles and relationships of the people talking (Tenor of discourse). Why do we use recount in conversations?

Modeling, deconstruction and joint construction

4. Presentation of whole recount text. Structure of the recount text – sorting, matching and labeling of whole chunks of text. Vocabulary notes.
5. Orientation. What kind of information is given at the start of the text?
6. Sequence of events. What sequence of events do the models talk about? Storyboarding pictures to construct the story.
7. Cohesion. How does the speaker join the events together?
8. Grammar and vocabulary. How are events described?
9. What is the listener doing? Asking questions / supplying information / surprise and support.
10. Turn taking strategies. Answering questions.

11. Sequencing events to construct a story

12. Cohesion and grammar

13. Practice with a partner

14/15 Teacher and student conversation / self-assessment and reflection

Appendix 3: Study plan distributed to students in the first class

Study Plan	Criteria	Understanding 1-5 ご理解	Ability 1-5 能力
Step 1.	Greeting 挨拶	5	5
Comments コメント: 先生の質問にすぐに答えられ、自分なりの質問をすることもできました。 頭ではわかっていても、口ではなかなかでてこないの。そこをも、こでできるようになりたい。 I could answer your questions immediately. I think my questions were good. Because it's hard to say what I'm thinking in class, I want to learn more conversation phrases.			
Step 2.	Orientation 要約文	5	5
Comments コメント: 頭に浮かんでいた考えをすぐに文にすることができました。 117, 何と、どこで、どうや、と、たれと、などすべての情報を書けた。 I could make sentence from ideas inside my head. I wrote all information, when, what, where, how and who.			
Step 3.	Sequence of Events イベントのシーケンス	5	4
Comments コメント: 自分の絵を相手にわかりやすく英語で伝えるのがとてもむずかしい。 It's difficult to explain what the pictures are in English. でも一生懸命伝えようと思ってあげたら、伝わったので理解はできているが 伝える能力が少し足りなかったと思う。 I tried hard and my Partner understood. I lack ability to communicate.			
Step 4.	Past Tense Verbs 過去時制の動詞	5	4
Comments コメント: I could translate the sentence to Past tense. But I couldn't make my Partner understanding. 過去形に変えて、内容を理解することはできたけど、友達にそれを伝えるのが 難しかった。途中でつまってしまったり、わからなくなることもあったので、たすけようように がんばる。楽しかった。 While I was reading, I often stopped reading difficult words. I'll try to make it possible. It was fun.			
Step 5.	Temporal Conjunctions 接続詞	5	4
Comments コメント: Conjunctions. It's easy to insert into sentence. But difficult to read. I'll practice to read 接続詞を文と文の間に入れることはとても簡単にできたけど、それをよくめと読む。 ときたスラスラと読むことができなかった。もと文と文のつなまりをスラスラとよめる ように練習してがんばる。 more fluently.			
Step 6.	Answering Questions 質問に答える	4	4
Comments コメント: I couldn't answer questions immediately. But I could answer with details. 今日は、相手の質問に対してすぐに答えることができなかった。でも、すぐには答えられなかった分 しかりと質問に答えることができたのでそこはよかったと思った。もと質問に文としての の答えをスムーズに答えられるようになりたい。 I want to answer more smoothly.			
Comments:			

Appendix 4: "Greetings" worksheet used to explore context and tenor

Greetings

A: Hi

B: Hi!

A: How are you?

B: I'm good. and you?

A: I'm good.

What have you been doing recently?

1. Watch the video and fill in the blanks.
2. Practice saying these greetings with your teacher:

Good morning! Hello! Hi! Hey! Yo!

Friendly ----->Very friendly!

3. Unscramble the following sentences:

are How. you?

How are you

it going? How's

How's it going

things? How's

How's things

4. Practice these replies with your teacher:

Fine thanks! And you?

Not bad. You?

I'm good! And you?

I'm alright. You?

5. Create your own greetings dialogue with a partner.

Appendix 5: Series of worksheets building the concept of “orientation”

1WHEN 2WHO 3WHAT (went to) 4WHERE 5HOW 2WHO

e.g.

During summer vacation I went to Nakamura
University with my friends.

Or,

Last week my friends and I went to Canal City.

The sentence must be *general* (一般的な) and not too *specific* (特定な):

e.g.

Last week I ate ice-cream

X

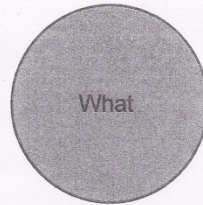
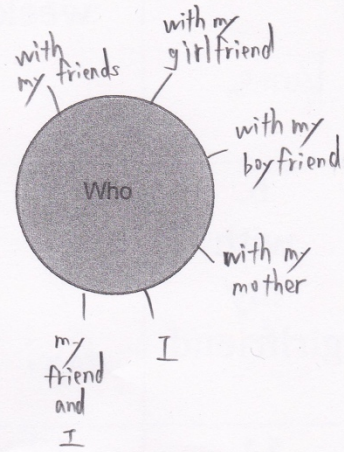
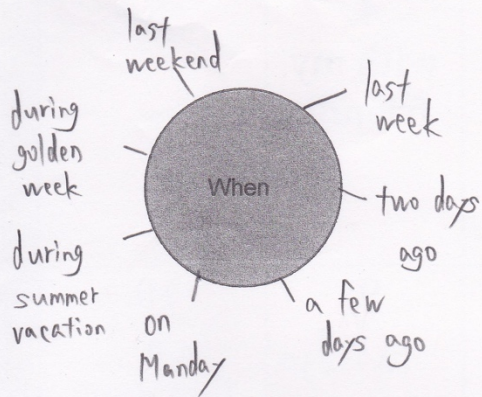
Yesterday I went to KFC

X

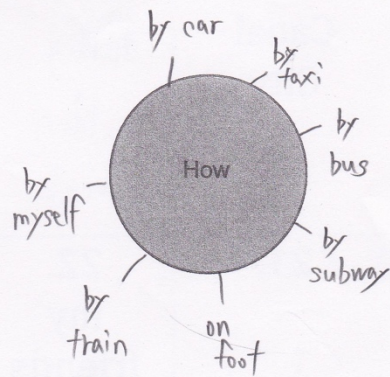
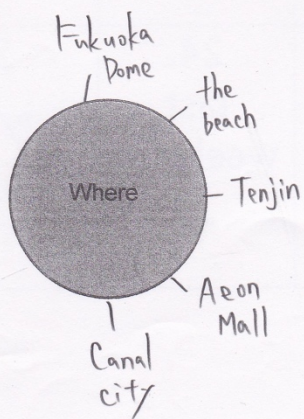
Make your own sentence:

During spring vacation I went to Space World
by train with my friends.

1. Fukuoka Dome	2. last week	3. the beach	4. with my friends	5. by car
6. with my girlfriend	7. by taxi	8. Tenjin	9. two days ago	10. a few days ago
11. by bus	12. Aeon Mall	13. on Manday	14. during summer vacation	15. with my boyfriend
16. Canal City	17. during golden week	18. with my mother	19. by subway	10. on foot
21. I	22. my friends and I	23. by train	24. last weekend	25. by myself



went to



Questions

1. When did she do something? *summer vacation*
2. Where did she go? *Nakamura university*
3. Who did she meet? *her friends*
4. What did she do first? *Japanese Archery*
5. What 3 things did she do after that? *Ban dance
went to friends house
went summer festival*
6. What did she do at the end?

自信	緊張	やる気
3 — 簡単にできると思う	3 — すごく緊張する	3 — 機会があれば絶対やりたい
2 — たぶんできると思う	2 — かなり緊張する	2 — 機会があればやってみたい
1 — あまりできないと思う	1 — 少しは緊張する	1 — 場合によりけり
0 — 絶対できないと思う	0 — ぜんぜん緊張しない	0 — できれば避けたい

How do you feel about talking like this?

自信: ① ② ③

緊張: ① ② ③

やる気: ① ② ③

Ask your partner "What have you been doing recently"?

How long could your partner speak for? 5

How long could you speak for? 17

Appendix 6: Temporal conjunctions worksheet

\Temporal Conjunctions

1. Listen to the audio. How many times do you hear the following words:

then	after that	that night

2. Listen to your teacher. Fill in the missing words:

During summer vacation I went to Nakamura University.

When I got there, I met my friends. *first*

Then I practiced Japanese archery for 2 hours. *next*

After that I practiced bon dancing with my friend's *then* grandmother.

Next, I took part in the local summer festival. *after dancing*

That night, I stayed at my friend's house. *finally*

3. Listen to your teacher and write the temporal conjunctions you can use in your story:

Beginning	Anytime	End
When I got there First	Then After that Next	That night Finally

Appendix 7: Asking and answering questions worksheets

Answering Questions

Your teacher may ask many types of questions:

WH Questions:

Who _____? (e.g. Who is your friend?)	Why _____? (e.g. Why did you buy that?)
What _____? (e.g. What colour was it?)	Which _____? (e.g. Which cinema did you go to?)
When _____? (e.g. When did you do that?)	How many / much _____? (e.g. How much did it cost?)
	How _____? (e.g. How was it?)

Yes / No Questions

Did you _____? (e.g. Did you like it?)	Have you _____? (Have you been there before?)
Was it _____? (e.g. Was it tasty?)	

1) Think of questions your teacher might ask you. Add details to the events on your storyboard. Some details you could add include:

PRICE (How much was it?) 2000 yen	COLOUR (e.g. What colour was it?) Red Blue Stripes
TASTES (e.g. What did it taste like?) It was tasty Strawberry flavour	APPEARANCE (e.g. Was it cute?) Cute Ugly Big
FEELINGS (e.g. How was it?) It was fun It was boring	

2) Listen to the video. What questions did I ask?

1. Was it good?

2. How long?

3. How far away is the shop from
here?

4. Which one?

5. What colour was it?

3) Listen to your friend's conversation. Ask the following question types. Check the box when you have asked:

<input checked="" type="checkbox"/>	Who?
<input checked="" type="checkbox"/>	What?
<input checked="" type="checkbox"/>	When?
<input checked="" type="checkbox"/>	How?

<input checked="" type="checkbox"/>	Did you?
<input checked="" type="checkbox"/>	Was it?
<input checked="" type="checkbox"/>	How much / many?
<input checked="" type="checkbox"/>	(your own question)

4) **If you include extra details in your story. Maybe you won't get any questions! See "Sequence of events" criteria 4 and 5 points. **

Practice your conversation again with as many details as possible for each event.

Appendix 8: Lesson plan with teacher's notes

Date: Week 3	Class Code: English Conversation 1
Materials: Worksheet 1, sample video with subtitles, distribute student portfolios	
Aims: Introduce the concept of spoken genres, and specifically the recount genre. Models of the genre and raise initial awareness of its key features.	
Objectives: <ul style="list-style-type: none"> - Greet students and ask them what they did last evening. Give my own account of last evening. - Check understanding through questions. Ask if students did similar. - Explain to students that this semester we will be conducting a speaking task aimed at developing a recount genre. - Review last week's genre work. Connect to spoken genres. - Show the sample video – explain that you will have to speak like that during a final interview with the teacher. - Listening Comprehension – watch video again and answer questions on worksheet 1. (Comprehension questions are based on key features of the recount genre). - Discuss answers in pairs. - Whole class answer discussion. - Ask your partner "What have you been doing recently?" - Discuss with the class how long they could speak for, what kind of problems did they encounter? - Group discussion: Why do we often use recount when we have a conversation? What are the lexico-grammatical features you noticed in the video? Refer to last week's genre study (structure, vocabulary, grammar of different genres) - Whole class discussion - Introduce next week's class: Deconstructing the recount genre and types of greetings. 	Notes: <p>Students were keen to use generic greetings</p> <p>familiar with what they seemed to be very relaxed with the video</p> <p>they primarily created short monologues</p> <p>go over this bit again next time</p>
Types of Assessment Planned: Listening comprehension (pair and whole class feedback). Teacher observations of pair discussions. Giving advice where appropriate. Student reflection in student journals.	

Students could identify the basic genre's criteria

Appendix 9: Questions to elicit content for the teacher and student journals

Teacher questions:

Questions about my teaching:

1. What did you set out to teach?
2. Did you achieve these goals?
3. What teaching materials did I use? How effective were they?
4. What techniques did I use?
5. What grouping arrangements did I use?
6. Was the lesson teacher dominated?
7. What kind of teacher-student interaction occurred?
8. Did anything amusing or unusual or memorable occur?
9. Did I have any problems with the lesson?
10. Did I do anything different from my usual practice?
11. Did I depart from the lesson plan? If so, why? Did the change make things better or worse?
12. What was the main accomplishment of the lesson?
13. Which parts of the lesson were most successful?
14. Which parts of the lesson were least successful?
15. Would I teach the lesson differently if I taught it again?
16. Were my beliefs about teaching reflected in the lesson? Reference to GBA
17. Did I discover anything new about my teaching?
18. What changes do I think I should make in my teaching?

Questions about students:

1. Did students actively contribute to the lesson?
2. Did I respond to different students' needs?
3. Were students challenged by the lesson?
4. What do I think students really learned from the lesson?
5. What did students like most about the lesson?
6. What didn't they respond well to?

Questions about Research:

1. Do these data answer my questions? If so, how?
2. What are the main messages so far?
3. What are the gaps in the message that I still need to fill?
4. Am I still asking the right questions or are the data telling me that something else is more important?
5. Do I need other kinds of data to help me really see what I am looking for?
6. To answer my questions, are some pieces of data more important than others?

Student questions:

1. What was the topic of today's class?
2. Did you learn anything new?
3. What activities did you do in class?
4. Did you have any chances to work with your classmates?
5. Did you have any chances to ask and answer questions in class?
6. Did you have any chances to use English in class?
7. Did you speak English in front of others?
8. How proficient do you think your spoken English was?
9. In what ways do you think your English is improving?

10. Did the teacher talk to you individually?
11. Did anything interesting or amusing happen?
12. Did you do something you've never tried before?
13. Did you have any problems with the lesson?
14. What changes do you think the teacher should make to the lesson?
15. Did the teacher help you when you had a problem?
16. Was the lesson challenging?
17. What was the most enjoyable part of the lesson? Why?
18. In what ways did the class help to develop your English speaking?
19. Did you understand the content of today's class?
20. How many times did you volunteer to answer questions in class?
21. How well did you contribute to class discussion?
22. How well did you contribute to group discussion?
23. What do you think you need to do to improve for next class, or for the final assessment?

Appendix 10: Teacher journal entry 4 – April 29th, 2013

This week I had 8 classes using the GBA. Total number of students present was (222) and the aim of the class was to introduce the concept of spoken genres and specifically the recount genre.

I prepared the class by selecting a good model speaker from last year so the students could see KSU students speaking well and to give them a realistic goal for their own speaking as compared to showing them native speakers and expecting them to emulate. This worked very well and I felt much better than the transcript I had used previously. I was worried that it would be difficult for students to understand as compared to having the transcript in their hand, but pausing the subtitles worked well and comprehension questions were answered accurately. Students seemed far more engaged with the video version, laughing at jokes and so on. It was more like a real conversation than the transcription and I think therefore students saw real value in it. I gave students a chance to talk in pairs, but still felt like I was dominating the class too much. But I think this will be a trend during these early stages of the course. Metalanguage was again a problem, giving students a chance to discuss and translate works well for getting deeper meaning I think, but it takes up valuable time and students were still asking if I could use more Japanese. I think for feedback and metalanguage Japanese is appropriate. I need to plan and predict what Japanese I will need before class.

The reaction from students when they were told that this is what they would do at the end of the semester made them laugh – partly because they felt it was a tall order, but also because they might have been motivated by it. This is reflected in student journals with comments such as “I want to learn greetings”, “I didn’t know this word, it was useful”. But also many negative comments saying that it was too difficult for them, and the model used was so great. It still felt to some that it was something beyond their capacity. Need to encourage them more that they can do it. The activity where I asked them to construct a recount text off the bat didn’t help, as they really couldn’t do it successfully and the activity floundered. I expected this however, and explained that this is the point of the course.

I may need to go over this lesson again, and I am a little concerned about the amount of material we have to cover for this level of student. I am hoping that as they grow accustomed to the approach and self-assessing themselves it will become easier. Students seem divided on self-assessment. Many are reluctant to write much, some write quite a lot. This is something I need to look at in more detail. Perhaps ask students directly what they think about these self-assessment procedures.

I feel students have a goal in these early classes and there is a point to the semester, they might not get learning random pages in a textbook. The main messages are models on video work better and students are better, more realistic models. Mutual intelligibility rather than native speaker fluency is the goal. Language problems again is a theme and authentic texts. Self-assessment needs addressing

OBSERVED PROBLEMS AND POTENTIAL RESOLUTIONS

Problems Observed

Potential Resolutions

Initial classes focusing on the transcript and deconstruction of texts did not provide many opportunities to produce spoken output.

The “chat segments” proved to be extremely difficult for all levels of student. Students recorded that they understood the criteria receptively, but could not produce it during the independent construction phase.

Some students were adept at lexical-grammar and structure, but utterances:

- were too quiet, or too loud
- filled with long pauses in communication, or even inappropriately short pauses
- L1 interference meant that the listener could only understand certain words with L1 phonological knowledge (particularly loan words). E.g.
- . or saazdei – “Thursday”
- Students did not make eye contact, or adopted body language such as sitting bolt upright during talk, making the communication feel unnatural or even unpleasant

These different factors to varying degrees added to misunderstanding and also an innate feeling of successful communication.

During assessment there were flooring effects, meaning some criteria were far too easy and didn’t provide useful items for assessment. Very low ability students still could not issue long utterances but relied on single vocabulary items.

Ceiling effects. With the removal of the “chat segments”, some students would be able to complete all criteria too easily.

The addition of further information gap activities to provide chances to talk with a purpose.

It became clear that, the teacher might need to take the role of one of the interlocutors during the final assessment in order to facilitate “chat”. Chat criteria would need to be removed during assessment, but hopefully replaced in subsequent semesters as students built on experience and would remain in the syllabus throughout.

The introduction of new criteria, in particular such criteria as:

- Clarity of voice
- Fluency
- Pronunciation
- Non-verbal communication

A study of other spoken assessments may provide useful ideas for developing these criteria.

The introduction of new criteria such as “vocabulary”, or even “dictation”, so very low ability students could still achieve some criteria.

The introduction of additional criteria such as “fluency” may address this.

Some students had clearly constructed and memorized written transcripts outside of class.

According to Hughes (2003) assessments must be able to rank order candidates, or they cannot make valid measurements of the construct in question. It was impossible to observe accurately the levels of difficulty of different criteria and how successfully different students achieved different criteria.

The 3-point Likert scale seemed inappropriate in distinguishing between abilities effectively.

Some students were able to achieve all criteria by saying far less than other students. Students who illustrated their talk needed to be accounted for in assessment

Problems Observed

Small sample size.

Bias shown towards own students.

A complicated assessment rubric.

It is still unclear whether students are aware of what is expected from them as students of the text-based approach.

Greetings and endings to conversation are absent.

Definitions of Pronunciation and Fluency continue to be inadequate.

The data in the research journal was difficult to organize and

Problems Observed

Students over-estimated their ability in non-lexical-grammatical criteria

The use of storyboards with pictures worked well to deter written transcripts, but a “fluency” criterion and the teacher taking the lead with the “chat” segments could deter memorized transcripts further.

An analysis of student scores for each criterion should be analyzed using an item response model, in order to contribute to construct validity. The Rasch model would also allow an analysis of individual students.

5-point Likert scale to be constructed.

A new criteria “Illustrations” added.

Potential Resolutions

Larger sample size and a second study.

Assessment may benefit from some input by strangers, to provide more objective feedback.

Simplification of rubric, or increased negotiation of syllabus content and rater training.

An interview or questionnaire to determine student perceptions of the text-based approach. Self-assessment may help to identify aspects of the syllabus students find difficult or unclear.

New criteria to be added to the syllabus.

Reference to the literature and adaptation of assessment rubric for these items.

Specific interviews and questionnaires could be designed to gather data from teachers and students about their thoughts and feelings about the text-based approach in their classes.

Potential Resolutions

Criteria need greater modelling in class, particularly criteria such as pronunciation and fluency. Negotiation of criteria with students would also address this problem

by closing the gap between what is expected and what is to be achieved.

Students underestimated their ability in lexical-grammatical criteria

As above and also the need for targeted and specific feedback to individual students.

Who is the “authority” on judgment in fluency and pronunciation?

Negotiation of criteria with students and also increased modelling along the full range of scores, from a poor speaker, to a high level speaker.

Small sample size

A larger scale study to take place

Affective factors such as anxiety and confidence not controlled for.

Strategies for measuring students’ confidence and anxiety before speaking to be explored.

Only one text type has so far been employed. Do other text types work?

A new text type will be explored after negotiation with students.

Appendix 11: Student journal (undated)

Voice was small and poor. Didn't say "and you" in reply

I could make summary


I could speak so deeply

I could get pp's verb

I had take a long time to connect sentence

I really glad to understand basic English conversation, When I speak English, I couldn't remember the many words so I thought, I have to use English in usual time and I wanna take opportunity to speak English in this class.

Appendix 12: Spoken assessment sheet

		Student ID number entered here
		Student or teacher scores entered here
1.	Greeting	① ② ③ ④ ⑤
自己評価の理由を説明して下さい Student or teacher comments entered here		
2.	Orientation	① ② ③ ④ ⑤
3.	Sequence of Events	① ② ③ ④ ⑤
4.	Past Tense Verbs	① ② ③ ④ ⑤
5.	Temporal Conjunctions	① ② ③ ④ ⑤
6.	Answering Questions	① ② ③ ④ ⑤

Structural Elements:	Person A	Person B	Grammatical Features:
Greeting	Hello!	Hello!	
	How's it going?	Not bad. You?	
	I'm good!	What have you been doing recently?	
Topic Sentence	Chunk Segments:	Chat Segments:	
Sequence of events	On Saturday my friend and I went to Canal City by bus.		
	When we got there we went to KFC and ate some food.	Expresses surprise or support	Past Tense Verbs
	Then we had some ice-cream.		
	First we went to P2 and saw the pets. The dogs were very cute.	Supplies Helpful Information	
	Next I bought a new black sweater in Comme Ca.		
Illustrations	It was expensive.		Conjunctions
	After that we went to see a movie in the cinema. We lined up for a long time to pay for tickets.	Asks Questions	
	Next we both went for dinner. We ate yakiniki.		
Answers Questions	At the end we went to the games arcade and took purikura.		
Conclusion	Then we got on the bus and came back to Kashii.		
	It was fun!		

Appendix 14: Data control file for Winsteps with example of raw data

TITLE = "Speaking Test"
 PERSON = Person ; persons are ...
 ITEM = Item ; items are ...
 ITEM1 = 5 ; column of response to first item in data record
 NI = 13 ; number of items
 NAME1 = 1 ; column of first character of person identifying label
 NAMELEN = 3 ; length of person label
 XWIDE = 1 ; number of columns per item response
 CODES = "123 " ; valid codes in data file
 UIMEAN = 0 ; item mean for local origin
 USCALE = 1 ; user scaling for logits
 UDECIM = 2 ; reported decimal places for user scaling
 &END
 New Text
 Answers Questions
 Illustrations
 Orientation
 Pronunciation
 Fluency
 Temporal Conjunctions
 Past Tense
 Non-verbal
 Clear Voice
 Sequence of events
 Vocabulary Bank
 Dictation
 END LABELS
 001 1323223333333
 002 1223213323333

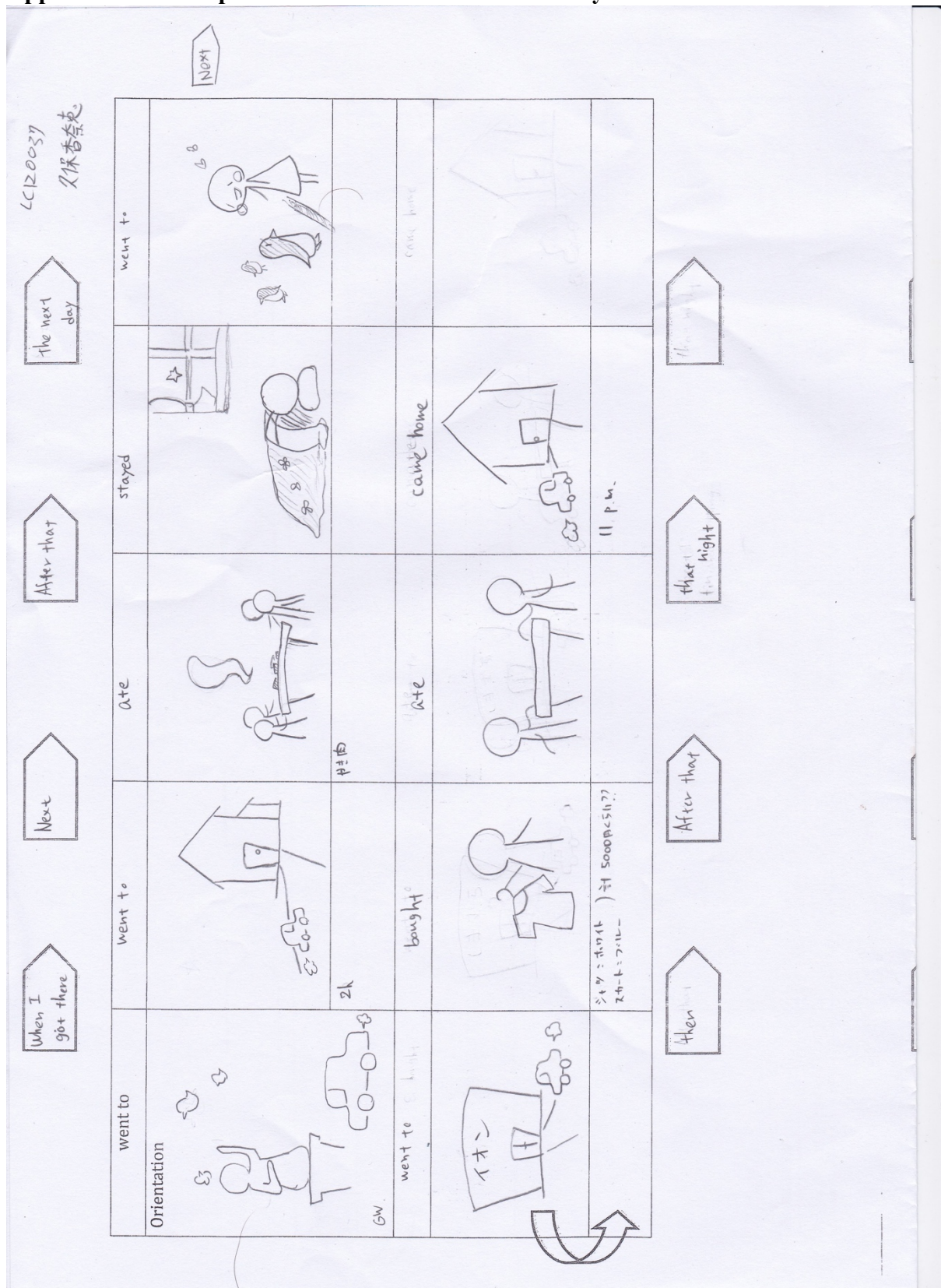
Appendix 15: Example of analytic rubric negotiated with students

	1	2	3	4	5
Greeting 挨拶	Did not respond to greeting 挨拶に反応しませんでした	Misunderstood the greeting and / or no reciprocation 挨拶を誤解と／または見返りに質問をしませんでした	Responded with a <u>one word</u> answer and no reciprocation 1 ワードの答えと見返りに質問をしませんでした	Responded and reciprocated 100% of the time 正確に答えたと100%を見返りに質問をした	Responded and reciprocated <u>with some originality</u> 正確に答えたと独創てきな見返りに質問をした
Orientation 要約文 Key information = WHO, WHAT, WHERE, WHEN 重要な情報＝誰、何、何処、何時	No key information 重要な情報なし	Included only 25% of key information and / or too specific 重要な情報の 25% と／または具体的すぎる	Included only 50% of key information and / or too specific 重要な情報の 50% と／または具体的すぎる	Included only 75% of key information, but not too specific 重要な情報の 75% でも具体的すぎるない	Included 100% of key information and not too specific 重要な情報の 75% と具体的すぎるない
Sequence of Events イベントのシーケンス	Only 1 event 1 つのイベントだけ	Only 2 or 3 events and very basic details 1 つまたは 2 つのイベントだけと非常に基本的な詳細	4 or 5 events with basic details 4 または 5 のイベントけど基本的な詳細	5 events or more, with some details 5 イベント以上といくつかの詳細	5 events or more with a lot of detail 5 イベント以上と多くの詳細
Past Tense Verbs 過去時制の動詞	Never used appropriate verb tense 正しい動詞の時制なし	Appropriate verb tense in only 25% of cases 25% 正しい動詞の時制	Appropriate verb tense in 50% of cases 50% 正しい動詞の時制	Appropriate verb form in 75% of cases 75% 正しい動詞の時制	Appropriate verb form in 100% of cases 100% 正しい動詞の時制
Temporal Conjunctions 接続詞	Never used temporal conjunctions 接続詞なし	Conjunctions used appropriately only 25% of the time and / or repeated 25% 接続詞は適切に使用と／または繰り返し	Conjunctions used appropriately only 50% of the time and / or repeated 50% 接続詞は適切に使用と／または繰り返し	Varied conjunctions used appropriately 75% of the time 75% 接続詞は適切に使用	Varied conjunctions used appropriately 100% of the time 100% 接続詞は適切に使用
Answering Questions 質問に答える	Could not answer any questions 答えることができませんでした	Only answered 25% of questions and / or one word answers 質問の 25% に答えることができると／または 1 ワードの答え	Answered 50% of questions and / or one word answers most of the time 質問の 50% に答えることができると／または 1 ワードの答え	Answered 75% of questions some using detail 質問の 75% に答えることができるのといくつかの明細	Answered 100% of questions and used details in many cases and sometimes reciprocates question 質問の 100% に正しい答えるの上でたくさん明細があると時々質問を与える

Appendix 16: Holistic rubric developed from the Common European Framework

	Pronunciation	Vocabulary	Structure	Fluency	Strategy
10 9	Excellent grasp of English sound system. Intonation reflects thoughts and emotions.	Very broad range of words used accurately. Expresses subtle differences in nuance.	Uses advanced grammar and complex sentences accurately.	Speaks smoothly in long, connected sentences, with few pauses.	Wide variety of strategies to control conversation, show involvement, and solve problems.
8 7	Slight accent, with some mispronunciations. Individual words clear and generally good intonation.	Broad range of vocabulary, with few inaccuracies. Can convey details and emotions.	Forms simple sentences accurately, and attempts more complex patterns. A few errors in tense and number agreement.	Mostly even tempo, with few noticeable pauses. Uses English hesitations and fillers.	Starts and maintains conversations, shows involvement, and helps partner or elicits help if problems arise.
6 5	Noticeable accent and some mispronunciations that could cause misunderstanding. Flat intonation.	Speaks with some detail using everyday vocabulary. Occasional errors in word choice or word form.	Uses set patterns accurately, and attempts original sentences. Some errors, but meaning is still clear.	Fairly even tempo, with some lengthy pauses for planning, and occasional repetitions and corrections.	Tends to follow partner's lead. Shows some involvement and makes an effort to solve problems.
4 3	Strong accent and frequent mispronunciations. Unnatural intonation.	Overreliance on basic vocabulary. Frequent errors in word choice or form. Strong L1 interference evident.	Memorized phrases or strings of isolated words only. Cannot compose original sentences.	Slow pace, with frequent repetitions and corrections. Cannot respond without noticeable pauses.	Answers questions, but has trouble maintaining a conversation and solving problems. Sometimes resorts to Japanese.
2 1	Sample too short or unclear to evaluate.	Simple words with no variation.	Single words, or errors in memorized patterns.	Mostly silent or speaking Japanese.	No strategies evident.
0	Complete silence or Japanese only				

	発音	語彙	構造	流暢さ	会話の方略
10 9	英語発音の方法がよく把握できている。自分の考えや気持ちをイントネーションで表わすことができる。	幅広い分野の語彙を正確に使い、微妙なニュアンスの違いも表現できる。	高度な文法と複雑な文章を正確に使える。	よどみなくスムーズに長く文章を続け、ほとんど間（ま）なく話せる。	多種多様な方略で会話をコントロールし、相手の話に関心を持つことを示したり問題を解決したりできる。
8 7	少しなまりや発音まちがいがある。一つずつの単語をはっきり発音されていておおむね良いイントネーションである。	ほぼ正確な幅広い分野の語彙で、詳細な事柄や感情も表現できる。	単文を正確に作れ、複文や重文も作ろうと試みる。時制や単複の一致などの基本的な文法に間違いが少しある。	ほとんど同じテンポで、明らかな間をおくこともなく話せ、英語で充填語やつなぎの言葉を使う。	相手の話に関心を持つことを示し、問題があれば相手を助けたり、助けを促しながら会話を始め、続けられる。
6 5	明らかになまりや誤解を招くような発音のまちがいがある。イントネーションは平たんである。	日常的な語彙を使い、ある程度詳しいことが表現できる。語彙や品詞の選択にときどき間違いがある。	決まった構文は正確に使え、自分で創造的に文章を作ろうとする。間違いはあるものの意味は通じる。	考えるための長い間や、ときどき訂正や言い直しがあるが、まあまあ自然なテンポで話す。	相手のリードに任せる傾向がある。相手の話に関心があることを示し、うまくコミュニケーションがとれないところを解決しようとする。
4 3	強いなまりや頻繁な発音まちがいがある。不自然なイントネーションである。	基本的な語彙の使用が目立つ。語彙や品詞の選択を頻繁に間違える。和製英語や日本語からの直訳が多い。	暗記しているフレーズや語句のみで、自分で創造的に文章を作れない。	頻繁に繰り返しや訂正をしながらゆっくりとしたペースで話す。明らかな間を取らずには返答できない。	質問には答えるものの、会話を継続させたり、問題を解決するのは難しい。ときどき日本語を使う。
2 1	発話が短すぎる、またははっきりせず評価できない。	シンプルな語彙のみでバリエーションがない。	一語文、または暗記しているパターンであるが間違いがある。	ほとんど発話が無い、または日本語を話す。	会話方略を使わない。
0	発話が全くないか日本語のみ				



Appendix 18: Examples of worksheets designed to address difficulties in the micro-aspects of conversation

WH Questions:

Who _____?
(e.g. Who is your friend?)

Why _____?
(e.g. Why did you buy that?)

What _____?
(e.g. What colour was it?)

Which _____?
(e.g. Which cinema did you go to?)

When _____?
(e.g. When did you do that?)

How many / much _____?
(e.g. How much did it cost?)

How _____?
(e.g. How was it?)

Yes / No Questions

Did you _____?
(e.g. Did you like it?)

Have you _____?
(Have you been there before?)

Was it _____?
(e.g. Was it tasty?)

1) Think of questions your teacher might ask you. Add details to the events on your storyboard. Some details you could add include:

PRICE (How much was it?)

COLOUR (e.g. What colour was it?)

2000 yen

Red
Blue
Stripes

TASTES (e.g. What did it taste like?)

It was tasty
Strawberry flavour

APPEARANCE (e.g. Was it cute?)

Cute
Ugly
Big

FEELINGS (e.g. How was it?)

It was fun
It was boring

2) Listen to the video. What questions were asked?

1. _____ good?

2. _____ long?

3. _____ far _____

_____?

4. _____ one?

5. What _____ ?

3) Listen to your friend's conversation. Ask the following question types. Check the box when you have asked:

Who?
What?
When?
How?

Did you?
Was it?
How much / many?
(your own question)

Conversation strategies:

Here is a list of conversation strategies you could use to keep a conversation moving:

1. **ASKING 5-W'S-AND-AN-H QUESTIONS** (*Basic question starters for asking questions and especially follow-up questions.*)

Who...? What...? Where...? When...? Why...? and How...?

2. **AGREEING OR DISAGREEING** (*It's okay to disagree when you do it in a fun, friendly atmosphere.*)

<agree> I agree. I think so too. That's right! That's what I think. Definitely! Uh-huh.

<disagree> I don't agree. I don't think so. I'm not sure. That's not right! No way! Un-un.

3. **ASKING FOR AN OPINION** (*To help develop our own opinions, it is useful to hear other people's opinions. But first we need to ask for them.*)

What do you think? What do you feel... In your opinion... From your point of view... For you... In your experience...

4. **ASKING FOR CLARIFICATION** (*What you do if you're not sure whether you understood or not.*)

Did you say, "_____" I thought you said, "_____"? You did *what*? You went *where*? You said, "_____"? Did you mean...

5. **ASKING FOR EXAMPLES** (*You can ask for examples to help understand what your partner is saying.*)

Could you give me an example? For example? Like what?

6. **ASKING FOR MEANING** (*You can ask for meaning when you want to understand a word or phrase.*)

What does "_____" mean? What does that mean? What do you mean? What's that?

7. **GIVING MEANING OR EXPLAINING MEANING** (*Telling the meaning of a word or phrase.*)

<give the meaning of a word> "Something" means... I think it means... That means...

8. **OFFERING CLARIFICATION OR CONFIRMATION** (*Saying more to help your partner understand what you mean.*)

Yeah, that's right. No, I said "something." I mean... I meant to say... What I mean is...

9. **ASKING FOR MORE INFORMATION** (*You can ask for more information when you want your partner to say more.*)

Could you tell me more about that? Could you say more? And then...? And so...? And...? I'd like to know... I'd be interested to know...

10. **ASKING FOR REPETITION** (*You can ask for repetition if you want to hear something again.*)

Could you say that again? Could you repeat that? Could you say that one more time? What was that? What did you say? What? Huh? One more time, please. Once more, please.

11. **ASKING YOUR PARTNER THE SAME QUESTION** (*You can ask the same question he or she just asked you.*)

How about you? And you? What do you think? Do you have any idea?

12. **AVOIDING SOMETHING TOO DIFFICULT** (*After making an effort, you can avoid saying something too hard for you.*)

Never mind. Forget about it. It's not a big deal. Oh, well. Whatever. Anyway...

13. **CORRECTING YOURSELF** (*You can correct yourself while you talk. Just say the correct word and continue talking.*)

No. Wait. Oops. I mean...

14. **GETTING TIME TO THINK** (*It's okay to take time to think. But let your partner know that you are thinking!*)

Just a moment. Hang on a second. Wait a sec. Let me think. Uh... Um... Well... Hmm...

15. **GIVING EXAMPLES OR EXPLANATION** (*Giving examples and explanation helps your partner understand.*)

<give an example> For example... For instance... <give an explanation> Let me explain...

16. **GIVING MORE INFORMATION** (*Telling your partner something besides what was asked; adding to your answer*)

<give extra information, for example, your feelings or adjectives>

17. **INTERJECTING** (*When you use words or short phrases while your partner is speaking.*)

Uh-huh. I see. Yeah. Really? Oh? Wow. That's great. Wonderful. Cool. Interesting. That's too bad. Oh, no! I can't believe it. You're kidding. Amazing. Brilliant.

18. **INVOLVING YOUR PARTNER** (*Helping your partner to do his or her part in the conversation.*)

<ask your partner lots of questions> What do you think? How about you? Don't you agree? How do you feel? What about you? Do you have any ideas? Do you know what I mean?

19. **MAKING EYE CONTACT** (*Showing your partner you are listening and care about what they are saying by looking them in the eye.*)

20. OFFERING A CORRECTION (*Noticing a partner's error and saying what you think is a correct form.*)

Did you say, <*mistake*>? You said <____>, but I think it's <____>. Is it <____>? Are you sure that's right? I think you mean <____>? *Don't you mean* <____>?

21. OFFERING NEW WORDS OR ALTERNATIVES (*Helping a partner by giving them new words or different words.*)

Is it <____>? Do you mean <____>? Maybe you mean <____>. Is the word you're looking for...

22. REPEATING (*Simply saying something again.*)

<say it again>

23. SHADOWING (*Repeating all or part of what your partner says aloud, softly, or in your mind.*)

<repeat part of what your partner said>

24. SHOWING INTEREST (*Showing your partner that you are interested in the conversation.*)

Uh-huh. Yeah. I know what you mean. Great. That's interesting. Unbelievable. You're kidding!

25. SHOWING YOU UNDERSTAND (*Showing your partner that you understand what he or she said.*)

I understand. I see what you mean. I see. I got it.

26. SUMMARIZING (*Retelling some of what your partner said in a shorter form.*) <say a shorter version of what your partner said>

In other words... You mean... It sounds like... So, you think...

27. USING GESTURES (*Helping your partner understand by using gesture.*)

<gesture>

28. USING NAMES (*Using one another's names can help us relax and learn more.*)

So, <name>. What do you think, <name>? You're <name>, right? Could you tell me your name again?

29. USING SIMPLER WORDS (*Helping your partner understand by using simpler words to explain something.*)

Appendix 19: Reflective questions for teachers

1. Why did choose to adopt Simon's lesson plans in your classes?
2. In what ways had you previously taught speaking that differed from this approach?
3. What teaching materials did you use from the lesson plans?
4. Which materials you find most effective?
5. Were there any materials you chose not to use?
6. Why did you choose not to use those materials?
7. In what ways did you assess students during the syllabus?
8. In what ways did this form of assessment assist the students?
9. In what ways did the materials assist the students in speaking?
10. Were there any specific elements of the syllabus that you think were particularly useful?
11. Were there any specific lesson plans or class activities that were particularly useful?

Appendix 20: Rubric for assessment during assessment exercise

Item	1	2	3	4	5
New text	Item not completed	Answers in one or two sentences	Answers in complete sentences, but many structural and grammatical items are missing, or misused	Constructs a new text mirroring all the items and structures of the recount text with few mistakes	Constructs a new text mirroring all the items and structures of the recount text with no mistakes
Answers Questions	Doesn't respond to questions and fluency is badly reduced	Answers in one or two words and fluency is reduced	Answers in very simple sentences without illustration. Fluency is slightly affected	Answers in full sentences with some illustration. Fluency is only slightly affected	Answers fully and illustrates new information. Fluency is not affected
Illustrations	Does not add any illustrations or attributes	Adds a simple one-word illustration one or two events	Usually adds illustration in a complete sentence	Illustrates each event of their speech with one or two attributes	Illustrates each event with complete sentences and numerous attributes
Orientation	No orientation; leaving questions as to WHO, WHERE, WHEN and HOW		Missing two or three pieces of key information, but orientation is clear		Students orient the listener with WHO, WHERE, WHEN AND HOW
Pronunciation	Knowledge of katakana or Japanese needed to understand many words		Majority of words are pronounced clearly, so as to be understood by a native English speaker with no Japanese background		All words are spoken clearly so as to be understood by a native speaker with no Japanese background
Fluency	Hesitates and clearly reciting a memorized script		Some hesitation during narrative, as though students have memorized some portions of the text		No hesitation in narrative or when answering questions
Temporal conjunctions	Never uses temporal conjunctions		Missing 2 or 3 temporal conjunctions and no variation		Use of varied temporal conjunctions between all events
Past Tense	Wrong tense used		Use of past tense for most verbs but miss 2 or 3		Students always use the past tense
Non-verbal	Students do not make eye contact, gesture or seem relaxed	Students make some eye contact but gestures and eye contact are unnatural	Students are nervous but use occasional eye contact and natural gestures	Students are relaxed and use some eye contact and gestures	Students are relaxed, use eye contact naturally, and gesture to illustrate their points
Clear Voice	Barely audible		Audible, but slightly muted or too loud		Voice is clear and of an appropriate volume
Sequence of events	Events are not time ordered		Some events are not logically time-ordered		All events follow a logical time ordered sequence
Vocabulary bank	Students can only use Japanese		Students use one or two English words		Students speak entirely in English
Dictation	Students say nothing		Students read a script, but it is difficult to understand		Students read a script, which the listener can understand

Appendix 21: Example of analytic rubric negotiated with students

	1	2	3	4	5
Greeting 挨拶	Did not respond to greeting 挨拶に反応しませんでした	Misunderstood the greeting and / or no reciprocation 挨拶を誤解と／または見返りに質問をしませんでした	Responded with a <u>one word</u> answer and no reciprocation 1 ワードの答えと見返りに質問をしませんでした	Responded and reciprocated 100% of the time 正確に答えたと100%を見返りに質問をした	Responded and reciprocated <u>with some originality</u> 正確に答えたと独創てきな見返りに質問をした
Orientation 要約文 Key information = WHO, WHAT, WHERE, WHEN 重要な情報＝誰、何、何処、何時	No key information 重要な情報なし	Included only 25% of key information and / or too specific 重要な情報の 25% と／または具体的すぎる	Included only 50% of key information and / or too specific 重要な情報の 50% と／または具体的すぎる	Included only 75% of key information, but not too specific 重要な情報の 75% でも具体的すぎない	Included 100% of key information and not too specific 重要な情報の 75% と具体的すぎない
Sequence of Events イベントのシーケンス	Only 1 event 1 つのイベントだけ	Only 2 or 3 events and very basic details 1 つまたは 2 つのイベントだけと非常に基本的な詳細	4 or 5 events with basic details 4 または 5 のイベントけど基本的な詳細	5 events or more, with some details 5 イベント以上といくつかの詳細	5 events or more with a lot of detail 5 イベント以上と多くの詳細
Past Tense Verbs 過去時制の動詞	Never used appropriate verb tense 正しい動詞の時制なし	Appropriate verb tense in only 25% of cases 25% 正しい動詞の時制	Appropriate verb tense in 50% of cases 50% 正しい動詞の時制	Appropriate verb form in 75% of cases 75% 正しい動詞の時制	Appropriate verb form in 100% of cases 100% 正しい動詞の時制
Temporal Conjunctions 接続詞	Never used temporal conjunctions 接続詞なし	Conjunctions used appropriately only 25% of the time and / or repeated 25% 接続詞は適切に使用と／または繰り返し	Conjunctions used appropriately only 50% of the time and / or repeated 50% 接続詞は適切に使用と／または繰り返し	Varied conjunctions used appropriately 75% of the time 75% 接続詞は適切に使用	Varied conjunctions used appropriately 100% of the time 100% 接続詞は適切に使用
Answering Questions 質問に答える	Could not answer any questions 答えることができませんでした	Only answered 25% of questions and / or one word answers 質問の 25% に答えることができると／または 1 ワードの答え	Answered 50% of questions and / or one word answers most of the time 質問の 50% に答えることができると／または 1 ワードの答え	Answered 75% of questions some using detail 質問の 75% に答えることができるのといくつかの明細	Answered 100% of questions and used details in many cases and sometimes reciprocates question 質問の 100% に正しい答えるのうえでたくさん明細があると時々質問を与える

Appendix 22: Self-assessment sheet completed by student

1.	Greeting	① ② ③ ④ ⑤
自己評価の理由を説明して下さい		
<p>とっさに Yo! と言ってしまったけれど、Hey と言い直すことができた。</p> <p>I have said "Yo" but after that I could say "Hey".</p>		
2.	Orientation	① ② ③ ④ ⑤
<p>あまり話すことができなかった。 I couldn't talk a lot.</p> <p>I was really nervous so I couldn't remember the words.</p> <p>言いたいことはたくさんあったけれど、緊張で単語が出てこなかった。</p>		
3.	Sequence of Events	① ② ③ ④ ⑤
<p>何個か言えたけど、上手く伝えることができなかった。</p> <p>I could tell about somethings but couldn't tell well.</p>		
4.	Past Tense Verbs	① ② ③ ④ ⑤
<p>時制がぐちゃぐちゃで自分でも何を言っているかわからなかった。</p> <p>I didn't know what I saying.</p>		
5.	Temporal Conjunctions	① ② ③ ④ ⑤
<p>緊張すぎて自分が何を言ったか覚えていない。</p> <p>I don't remember what I saying because I was really nervous.</p>		
6.	Answering Questions	① ② ③ ④ ⑤
<p>何とか答えることはできたと思う。</p> <p>I could answer something.</p>		

Appendix 23: Cycle Two: post syllabus student recount texts

Transcript A:

(Japanese words are included in *italics*)

A: Teacher

B: Student

1. A: Hello, how are you?
2. B: I'm fine. And you?
3. A: I'm good. What have you been doing recently?
4. B: Mmm, one day, during summer vacation, I went to the Nakamura University and I met my friends. Then I practiced Japanese archery.
5. A: Archery?
6. B: Yes, Japanese archery.
7. A: Are you good?
8. B: Yes.
9. A: Yes?
10. B: OK, but I'm beginner.
11. A: How long?
12. B: Two years. About two years. And . . . I practiced Japanese archery for two hours
13. A: Two hours?
14. B: Two hours.
15. A: You must have been tired?
16. B: Yes. After that I went to my friend's house on foot and I practiced the *Bon* dancing
17. A: *Bon* dancing?
18. B: Yes, *Bon* dancing.
19. A: How far is your friend's house from here?
20. B: Oh . . . about . . . two . . . twenty minutes.
21. A: Twenty minutes? So, you did archery, then walked to your friend's house, then dancing?
22. B: Yes. We practiced with my friend's grandmother and after that we took part in the summer festival.
23. A: Really?
24. B: Yes.
25. A: Which one?
26. B: Near here. My friend's local summer festival. And we dance, *Bon* dancing.
27. A: Really?
28. B: Yes.
29. A: What's *Bon* dancing?
30. B: and . . .
31. A: Is it, like, waving your arms?
32. B: Yes.
33. A: Nice, ok.
34. B: and . . . I stayed at my friend's house . . . that night.

Transcript B:

A: Student

B: Student

(“*Shirasu*” are juvenile sardines. “*Shirasu-don*” is a bowl of juvenile sardines served on rice)

1. A: Hi
2. B: Hi, hello, what's up?
3. A: I'm good, what about you?
4. B: Oh, I'm very sleepy today.
5. A: Why do you so sleepy?
6. B: Because I slept very late.
7. A: Oh, late.
8. B: I went part time job too many hours, so I was very, very sleepy.
9. A: You seems very sleepy
10. B: Oh, yeah ... what have you been doing recently?
11. A: Me?
12. B: Yeah!
13. A: This Monday I went to Enoshima by bicycle.
14. B: Oh, Enoshima?
15. A: Yeah, with my friend. And it took about three hours each way. So, I was very tired.
16. B: Oh, I think so!
17. A: But it was very fun!
18. B: Oh, yeah, yeah, yeah.
19. A: First, I uh, first in Enoshima, I went to the sea.
20. B: Oh yeah?
21. A: Yeah, but the sea was not so clean because of many, many trash. So dirty!
22. B: Yeah, I think so! I think so!
23. A: But, I played with water for a little bit. It was fun. Then, I ate *shirasu-don* in Enoshima. It was really tasty.
24. B: Yeah, that's very nice.
25. A: I want to go there and eat that again some day. You want to come?
26. B: Oh yes! Which do you like, boiled *shirasu*, or raw *shirasu*? I like raw *shirasu*.
27. A: Raw?
28. B: Yes, raw is very tasty. I think so.
29. A: Oh, I didn't like the raw *shirasu*.
30. B: Oh really?
31. A: Yeah, I tried raw a few years ago but I didn't like the taste, I like boiled *shirasu*. So, I ate boiled *shirasu*.
32. B: Oh, I very like it. I'm very hungry now.
33. A: Really?
34. B: Yeah.
35. A: Did you eat this Monday? This...
36. B: Morning?
37. A: Morning! Not Monday!
38. B: No time. Because no time. Can't eat morning. So, I'm very hungry. I want to eat. I can't wait!
39. A: OK, OK. What have you been doing recently?

40. B: I went Shibuya with my best friend. I went Shibuya, did many things. First, I went watch movie.
41. A: Oh!
42. B: I watched Pokemon.
43. A: Oh, that's nice.
44. B: And I saw *Ani ni Ai Saresugite Komatteiru*. Do you know this movie?
45. A: I don't know!
46. B: Oh, this movie is love story. My friend very, very like that movie's main character actor. He is very handsome.
47. A: Who?
48. B: Umm, Katayose Riyota.
49. A: Oh, I know, Exile!
50. B: Yeah, he's the Exile's member. She, my friend like he, but I like Shirahama Aran. He's Exile's leader.
51. A: Oh, he is so handsome!
52. B: Yes! I like, very like his face! After that I went shopping. In Marukyu. Do you know Marukyu?
53. A: Yeah, I know.
54. B: I bought many clothes, so my money fly away!
55. A: Fly away?
56. B: Yeah! So much! I like Adidas, I like very much sports maker, Adidas. I bought Adidas sneakers.
57. A: Sneakers?
58. B: Stan Smith, do you know?
59. A: Stan Smith, I know.
60. B: Stan Smith is very expensive. Uh, so my money fly away.
61. A: Your money flew away!
62. B: Yeah, and I went home.