THE EFFECTS OF RAISING LEARNERS' AWARENESS OF METAPHORICAL VOCABULARY ON WRITTEN PRODUCTION IN THE CONTENT-BASED CLASSROOM

by

PHILLIP JAMES BENNETT

A thesis submitted to
the University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of English Language and Applied Linguistics
School of English, Drama, and American & Canadian Studies
College of Arts and Law
The University of Birmingham
December 2015

ABSTRACT

It is widely recognised that language learners require extensive vocabulary knowledge to cope with the demands of studying content fields in English. As well as being rich in general academic and technical terms, academic discourse has been shown to make frequent use of metaphor to express abstract concepts and to achieve rhetorical goals. While research has shown the benefits of raising learners' awareness of the underlying motivation of metaphorical expressions, these findings have yet to be applied to authentic classrooms over longer periods of study.

This thesis examines the effects of raising Japanese learners' awareness of metaphorical expressions in a CLIL anthropology course. It examines the written work from two groups of learners: a control group whose language instruction focussed on academic and high frequency vocabulary and an experimental group who received instruction on course-specific metaphorical themes. Variation in metaphor production is compared for the two conditions and across learner abilities, and the interaction between the frequency, dispersion and salience of metaphors in classroom input and learner output is considered. The study then investigates the influences of word frequency, part of speech, phraseology and the L1 on learner metaphor production before concluding with recommendations for pedagogic practice and further study.

DEDICATION

This thesis is dedicated to my family, without whose support, patience and understanding I could never have got this far.

ACKNOWLEDGEMENTS

From the early planning stages of this study, through the long periods of data collection and transcription and then the writing up process, I have depended on advice and support from many people. I would like to offer my sincere gratitude to you all.

To my PhD advisors, Professor Jeannette Littlemore and Dr Nicholas Groom, for their expert guidance, for encouraging me to pursue my ideas, for pointing me in the right directions, and for giving me countless reading tips. The supervision meetings have always been something I looked forward to - they were both enjoyable and rewarding.

To my teaching partner and valued colleague Dr Debra Occhi, for constant support both inside and outside of our class. You've been just as passionate about my thesis topic as I've been myself, and you were much more on the ball about pointing out metaphors in class than I was!

To my wife Yumiko and our children Aidan and Emily for your encouragement, patience and understanding. I hope I haven't done too bad a job of balancing family life with the disappearing-off-to-spend-hours-in-front-of-the-computer!

To my friend and research collaborator Tim Stoeckel, for being a trusted sounding board and for your patience as I couldn't keep up with our own projects. I've learnt an awful lot from you, and I hope you didn't mind my constant updates on how many words I'd transcribed or written!

And finally to the students who participated in this study, thank you very much. It's your work that this thesis is built on, and I hope learning about metaphor in vocabulary has been as useful for you as it's been interesting for me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Abbreviations

Abbreviation	Refers to
BAWE	British Academic Written English Corpus
[BAWE####]	Notation for writing from BAWE sample
BNC	British National Corpus
CC	Closed-class (part of speech)
CEFR	Council of Europe Framework of Reference
CL	Cognitive Linguistics
CLIL	Content and Language Integrated Learning
CMT	Conceptual Metaphor Theory
COCA	Corpus of Contemporary American English
[CON#]	Notation for writing by participant from the control group
[CON L#]	Notation for sample from control input corpus (L = lesson number)
DFMA	Discarded for metaphor analysis (MIPVU term)
EFL	English as a foreign language
EMI	English-medium instruction
ETS	Educational Testing Service
[EXP#]	Notation for writing by participant from the experimental group
[EXP L#]	Notation for sample from experimental input corpus (L = lesson number)
ICLE	International Corpus of Learner English
IELTS	International English Language Testing System
L1	First language (native language)
L2	Second language
LOCNESS	The Louvain Corpus of Native English Essays
MI	Mutual information (collocation measure)
MICUSP	The Michigan Corpus of Upper-level Student Papers
MIP	Metaphor Identification Procedure
MIPVU	Metaphor Identification Procedure Vrije Universiteit
MIV	Metaphor Identification through Vehicle terms

Abbreviation	Refers to
MRW	Metaphor-related word (MIPVU term)
NICLE	Norwegian component of the International Corpus of Learner English
NICT JLE	National Institute of Information and Communications Technology Japanese Learner English Corpus
OC	Open-class (part of speech)
POS	Part of speech
SPICLE	Spanish component of the International Corpus of Learner English
TOEFL	Test of English as a Foreign Language
TOEIC	Test of English for International Communication
WIDLII	When in doubt, leave it in (MIPVU term)

Glossary of Japanese terms and cultural entities

Term	Refers to
banzuke	a ranking list for sumo wrestlers
bunka	the Japanese translation of the word *culture*
dohyou	the clay ring on which sumo bouts are held
ekiden	a long-distance relay race
freeter	a temporary worker, especially one who works short-term contracts and who prioritises their lifestyle and hobbies over a career
geta	traditional Japanese wooden footwear
Gothic Lolita	a Japanese fashion and social group in which young women wear frilled clothing that blends Victorian and anime styles
Harajuku	a district in Tokyo known for youth fashion
Hatsune Miku	a *vocaloid* character who has become a cultural phenomenon
idol	a Japanese entertainer who often appears as a pop singer, television show panellist or actor, and who is expected to act as a good role model
kanzashi	traditional Japanese hair ornaments
karuta	traditional Japanese playing cards
katakana	a Japanese writing system used mainly for loanwords
Kyary Pamyu Pamyu	a Japanese pop star and fashion model who has popularised the 'cute' fashion style associated with Harajuku, Tokyo.
maid	young women who work as waitresses in maid cafes, restaurants where the customers are treated as masters; maids usually wear costumes that blend French maid and anime styles
omuraisu	a Western-influenced Japanese dish (omelette with fried rice)
otaku	a Japanese social group consisting of young people who spend their time and money playing video games and reading comic books; more recently, the term has been extended to mean anyone who is passionate about their hobby.
purikura	a photo booth popular with teenagers (shortening of *print club*)
shinkansen	a bullet train
spaghetti neapolitan	a Western-influenced Japanese dish (spaghetti with ketchup)
taco rice	a Western-influenced Okinawan dish
udon	a Japanese thick wheat flour noodle

Term	Refers to
vocaloid	a computer voice software that allows users to have their lyrics sung by a voice actor; now a music genre.
wa	a Japanese cultural concept that relates to harmony and valuing the group over one's own interests
wasei-eigo	Japanese expressions that are derived from English words that do not exist in English
yokozuna	the highest rank for sumo wrestlers

CHAPTER 1. INTRODUCTION TO THE STUDY

As internationalisation progresses and global markets become ever more competitive, many national governments are looking to their education systems to cultivate a sense of global citizenship in their graduates and to attract more international students. This has driven the growth of English-medium instruction (EMI) programmes, and Japan is no exception to this trend (Brown & Iyobe, 2014; Dearden, 2014). The Japanese government in 2009 announced a plan to dramatically increase the number of university programs conducted entirely in English, and in 2014 launched a project aimed at making the nation's universities more internationally competitive (MEXT, n.d.; MEXT, 2014). These moves show laudable ambition, and in order to achieve successful outcomes, it is necessary to consider the learning demands of studying academic content in English and to assess how learners might best meet them.

Academic discourse, whether in publications or the classroom, is noted for having its own specialist terms and for the frequent use of multiword expressions and metaphorical language (Hyland, 2012; Low, Littlemore, & Koester, 2008; Steen, Dorst, Herrmann, Kaal, Krennmayr & Pasma, 2010). Consequently, the demands on non-native speakers' lexical knowledge are considerable, as they must process meaning on topics that are often abstract in nature and understand instructors' stance towards complex issues if they are to comprehend the course content. Yet learners require more than just receptive comprehension; as active learning and communicative approaches to teaching remain popular, instructors will expect

learners to voice their opinions, request clarification and participate fully in both spoken and written interactions.

One issue language instruction faces is that, at present, there are few principled approaches to breaking down the lexical challenge for learners. The most common method is to divide the lexicon according to frequency, with priority given to mastering highly frequent words before moving on to mid-frequency or specialist vocabulary (Nation, 2013, p. 14 - 32), and while this is an entirely logical, well-founded approach, it cannot meet all learning needs. One reason is that this method is primarily aimed at increasing learners' vocabulary *breadth*, or the number of words they know, rather than improving the quality, or *depth*, of their word knowledge. Depth of knowledge involves, among other things, an understanding of the range of meanings a word can possess and the nuance expressed with each, so it is crucial for fluent performance, especially in academic contexts where learners may be expected to express themselves with precision and subtlety.

However, depth of knowledge is not without its problems. It is much more difficult to measure vocabulary depth than vocabulary breadth (Batty, 2012; Stewart, Batty, & Bovee, 2012) and is likely to involve subjective decisions as to whether attested uses constitute separate senses or not (Hanks, 2013, p. 69 - 70). It is also more difficult to approach the development of vocabulary depth in a principled way; questions such as which words to prioritise and which aspects of word meaning to consider have no clear answers. But here the specificity of each academic discipline may work to instructors' advantage. It has been noted that particular fields draw on metaphorical themes to convey important abstract concepts. For example, economics makes use of the metaphor THE ECONOMY IS AN ORGANISM (Charteris-Black,

2000, p.155 - 156), which motivates the special senses that *health*, *growth* and *decay* take on in economic discourse.

This use of specific language patterns to express concepts that are central to individual disciplines offers another way to address vocabulary in a principled manner in academic settings. As well as enabling learners to extend their depth of vocabulary knowledge, this approach could be adapted to the discourse patterns of any given field, making it an ideal technique for EMI settings. As learners become more familiar with the ways that vocabulary is used in each discipline, they should also become more confident participants in academic dialogue, gaining the ability to apply metaphorical language to express themselves with accuracy and perhaps some individual flair. While there will always be debate over the extent to which language learners should be steered towards the norms of a given discourse community (Seidlhofer, 2005), the view taken here is that it is necessary to have an awareness of what is conventional in order to fully comprehend course content. In no way, however, should this lead to the curtailing of learners' linguistic creativity, for that is an important part of attaining a sense of language ownership.

This thesis describes an intervention study into the effect of using metaphor awareness-raising techniques on learners' written production in a Content and Language Integrated Learning (CLIL) anthropology course titled Japanese Popular Culture. Covering issues such as cultural borrowing, national identity and hybridization, the class required learners to engage with abstract topics and produce written reflections on important course themes. Language instruction for the control group centred on teaching high frequency and academic vocabulary, while learners in the experimental condition were provided with a metaphor workbook that was

used as both a reference material and a textbook for studying metaphors that appeared in the course.

As a longitudinal classroom study that draws on cognitive linguistic findings into the learning of motivated language forms, this research serves as a complement to experimental designs and corpus-based analyses. While classroom studies necessarily entail some loss of experimental rigour and cannot draw on the large samples of corpora investigations, they can offer insights into the subtleties of learner development – sudden bursts of improvement or regressions, avoidance and overuse, and the influence of context on performance. Their value lies in taking theoretical claims and experimental findings into everyday settings so that performance can be observed in a dynamic learning environment.

Chapter two of this thesis reviews the theories and experimental findings that underpin this study. Conceptual Metaphor Theory (Lakoff & Johnson, 1980/2003) was hugely influential in raising awareness of metaphor as a pervasive linguistic phenomenon. A brief overview of the theory is provided, and criticisms that have shaped and refined it are considered. The chapter then considers issues related to the identification of metaphor; in particular it will discuss the distinction between metaphor and metonymy, the merits of drawing on intuition or using authentic sources, and the criteria by which metaphor can be objectively defined. The chapter then shifts its focus towards metaphor and language learning. Insights from investigations of learner corpora lead into a discussion of the broader notion of metaphoric competence in language learners. Cognitive linguistic studies into teaching metaphor are then presented, along with the issues these have raised. How knowledge of metaphor is related to lexical development more generally is then

discussed, and the chapter concludes by reviewing the claims made for and against action research, and relating this study to such claims.

Chapter three describes important features of the context in which the study took place, including the course objectives and participants. The procedure that both conditions followed is detailed, including a description of how the language produced in the control condition was used to inform development of the workbook that acted as the primary method of instruction for metaphor awareness raising in the experimental condition. Finally, the sources of data that were obtained to present as detailed a picture as possible of metaphor in the class environment are explained.

Chapters four to seven each present the results of investigations into different elements of learners' metaphor production. The following paragraphs present a summary of the research questions that guided each investigation.

The fourth chapter examines the effect of the experimental treatment by considering how the metaphorical language produced in the two entire study conditions differed. The first research question considers the extent to which the chosen method of metaphor identification was able to detect a change in overall levels of metaphor production between the two groups of learners. The other research questions in this chapter each consider a different facet of the data to examine the ways in which output in the two conditions differed. These questions cover metaphorical output among the various parts of speech, language that was selected for targeted explicit instruction in the experimental condition, and the interaction between writing topic and metaphorical output.

Chapter five then introduces an important variable to the analysis - that of learner ability. First, the relationship between overall metaphor production and learner ability is considered. The next research question examines whether learner ability is related to how open- and closed-class parts of speech are used to produce metaphor. The third research question considers how learners differ in their use of target language, considering both quantity and variety of output. Finally in this chapter, the range of rhetorical functions that learners of different ability are able to use metaphor to perform is investigated. This analysis reveals trends both at the higher functional level and in terms of the types of lexis learners typically draw on.

Chapter six takes a step back to consider the classroom environment more broadly. The degree to which a sustained focus on metaphor affects the quality, frequency and regularity of metaphorical input is the first question to be considered. The chapter then considers the relationship between classroom input and learner output. The second research question asks whether changes in metaphorical input may have led to increased metaphor production. Finally, the chapter considers the longitudinal interaction between input and output and looks for signs of increased production over time.

The final investigations in chapter seven examine the effect of language factors that are external to the classroom. The first of these is the question of whether word frequency influences learners' metaphorical output. Frequency is recognised as a strong predictor of word recognition, but fewer studies have been carried out on productive use. Next, the chapter turns to the question of over- or underuse of particular parts of speech. It has been noted that learner writing bears some similarities to spoken discourse in terms of its distribution of parts of speech

(Granger & Rayson, 1998), but whether this tendency extends to metaphor use is a question that has received little attention. Another aspect of language that is closely related to metaphor is phraseology. The third research question in this chapter examines differences in phraseological patterns containing metaphors and asks whether experimental group learners showed greater production of such patterns. The final research question looks for evidence of learners' L1 influencing their production of metaphors in English. Both facilitative and debilitative effects are considered, and the influence of the L1 compared with that of the L2 is explored.

Chapter eight provides a summary of the main findings of this study as they relate to classroom practice and lexical development. The limitations of the study will be reviewed, and potential avenues for further investigation will be considered.

CHAPTER 2. BACKGROUND TO THE STUDY

Metaphor is intrinsic to human life. While commonly associated with fields of artistic endeavour, it can readily be put to more functional purposes in advertising, debate or journalism, as well as in seemingly prosaic matters such as gesture and casual conversation. At its extremes, metaphor can be so striking that it will be remembered for hundreds of years or so commonplace as to pass by unnoticed.

At a basic level, metaphor can be thought of as seeing one thing in terms of another. That is, the qualities of one particular entity are mapped onto another unrelated entity that is being described. Metaphorical language can have the effect of highlighting certain features of the described entity such that it becomes more concrete, vivid, or noteworthy. This can be seen in the citations from the Corpus of Contemporary American English (COCA) and British National Corpus (BNC) below[1].

(1) Nancy Pelosi was so upset with Robert Gibbs when he said that Democrats are going to be slaughtered in November. [COCA]

(2) … sometimes an affair that doesn't move on fizzles out as time passes by. [COCA]

(3) One night this little piece of fluff walked into Emergency with a broken fingernail. Jeff went kind of goo-goo, and six months later he was married. [BNC]

It is common to refer to the domain of experience from which the metaphorically used word is drawn as the *source*, and the domain to which it is applied as the *target*. In (1), the verb *slaughtered* is taken from the source domain of

[1] In this thesis, attested language data will be drawn from COCA, a 520 million-word balanced corpus of US English, the BNC, a 100 million-word balanced corpus of British English, the enTenTen13 corpus, a web-crawled, 19 billion-word corpus of English (Jakubíček, Kilgariff, Kovář, Rychlý, & Suchomel, 2013), and a sample of BAWE, a corpus of academic writing from British university students.

killing and applied to an election as a target domain, adding a sense of violence and cruelty to political defeat. The examples in (2) operate in a similar fashion, with the source domain of fire (*fizzle out*) describing the targeted passion of a romantic affair, and physical movement (*move on*, *passes by*) providing the source domain for both the development of a relationship and the progression of time. The expressions in (1) and (2) are relatively standard ways to refer to the concepts they describe, and in fact both the Macmillan English Dictionary for Advanced Learners and the Longman Dictionary of Contemporary English only list *fizzle out* in the sense of an emotion fading away[2]. The examples in (3) demonstrate how metaphorical language can be used for affective purposes. Both *little piece of fluff* and *goo-goo* come from source domains that imply a sense of criticism to their targeted entities. They also exemplify less conventional usage patterns; *fluff* in this sense is typically defined as something that is entertaining but lacking in seriousness or importance, but it is not usually applied to people, while *goo-goo* does not appear in either of the above dictionaries[3]. The Oxford Dictionaries website (https://en.oxforddictionaries.com) offers the definition of *goo-goo* as 'amorously adoring' or 'childish or meaningless'.

As these examples indicate, metaphor can perform various roles in language. Such roles include providing an understanding of abstract concepts in ways that are more immediately accessible, such as seeing time as space or physical movement, and allowing for the application of nuance to discourse. Metaphor can be used for

[2] According to the Oxford English Dictionary, the original meaning of *fizzle* was "to break wind without noise". This sense was recorded in the 16th century, around 300 hundred years before the 'hissing and sputtering' sense came into use. This serves as a graphic example of how conventional usage can shift over time. See section 3.4.2.2, page 68 for a further comment on why this expression would not be coded as a metaphor using the procedure adopted in this study.

[3] In this thesis, the Macmillan English Dictionary for Advanced Learners and the Longman Dictionary of Contemporary English are used to provide standard definitions. Exceptions to this will be stated in the text.

both poetic and rhetorical purposes, and can be either innovative and insightful or highly conventional (Mahon, 1999, p. 69 - 70).

Several studies have drawn attention to the ubiquity of metaphorical language across genres (Cameron, 2003; Deignan, Littlemore, & Semino, 2013; Steen et al., 2010), and as such, metaphor has begun to attract interest in applied linguistics research. Many studies (Boers, 2003; Holme, 2004; Littlemore, 2004) have made use of the principles of cognitive linguistics to show how the motivating forces behind everyday metaphorical language can be explained.

This research has also found support from broader trends across the many sub-domains of linguistics. There has been a growing awareness of the need for language learners to develop vocabularies that are broad (number of words known), deep (knowledge of those words), integrated (connected with other words), and automatic (easily accessible) (Hulstijn, 2001; Nation, 2013, Chapters 1 - 2; Read, 2004; Schmitt, 2010, Chapter 1). Corpus linguistics has also begun to provide evidence of the patterning that is frequently found in regular language between the traditionally separate lexical and grammatical language components (e.g., Hoey, 2005; Hunston & Francis, 1999; Sinclair, 1991). These views of language as a complex yet explicable construct incorporating both specific lexical information and more schematic structural patterns are also fully accommodated within a cognitive view of language (Langacker, 1987, 1991, 2008b).

The following sections of this introduction will outline key theoretical arguments regarding metaphor as a psychological, linguistic and social phenomenon as well as methodological considerations in the identification of metaphor in language. The emerging role of learner corpora as tools to investigate metaphorical language

production is discussed, and the concept of metaphoric competence in language learners is used to highlight the range of purposes for which knowledge of metaphor can aid learners. Findings pertinent to the teaching of metaphor will then be considered, and the relationship between the development of vocabulary knowledge and metaphor discussed. The chapter concludes with a discussion on the value and limitations of action research approaches, and positions the current study in relation to this paradigm.

2.1. Conceptual Metaphor Theory

The foundation of modern-day research into metaphor is Lakoff and Johnson's (1980/2003) *Metaphors we Live by*, which put forward the theory that metaphorical reasoning, grounded in human experience, is a principal mechanism by which we make sense of the many abstract phenomena in our lives. It holds that recurrent language features offer evidence of how we conceptualize our world and its social constructs as well as our own psychological states. The examples below, which were drawn from introspection, illustrate the conceptual metaphor IDEAS ARE PLANTS:

- His ideas have finally come to *fruition*.
- That idea died *on the vine*.
- That's a *budding* theory.
- It will take years for that idea to *come to full flower*.
- He views chemistry as a mere *offshoot* of physics.
- Mathematics has many *branches*.
- The *seeds* of his great ideas were *planted* in his youth.

- She has a *fertile* imagination.
- Here's an idea that I'd like to *plant* in your mind.
- He has a *barren* mind.

(Lakoff & Johnson, 2003, p. 47, emphasis in the original)

According to Conceptual Metaphor Theory (CMT), the italicized words in the examples above are not merely colourful language; they are reflections of how the concept of an idea is structured in the mind. This structuring is also claimed to be systematic since just as plants can begin life, grow successfully, and spread or wither away through unfavourable conditions, so ideas have initial stages, possibilities to adapt and become more ambitious or unfortunate endings. In other words, there is a consistent mapping of attributes from the source to the target domain. No mapping is ever perfect; however, CMT proposes that each mapping highlights certain aspects of the target domain at the expense of others. In this instance, the aspects of ideas that are brought into focus are their developmental nature and unknowable outcomes, while other features, such as the moral import or sense of ownership that is sometimes attached to ideas is obscured.

It is important to note that what is being argued here is not that the conceptual metaphor IDEAS ARE PLANTS resides in any particular linguistic expression. Rather, it is our ontological awareness of the principal players and epistemic knowledge of ideas and plants that allows understanding of language reflecting this metaphor. As a result, even if we encounter a novel expression that uses the metaphor, we are able to apply our understanding of plants to ideas in order to extract meaning. As Lakoff later argued (1990, p. 51 - 65) in his Invariance Hypothesis, metaphorical mappings

maintain the coherent referential relationships, or cognitive topology, between the source and target domains, allowing for knowledge of the source to lead to inferences about the target.

One criticism of CMT has been that the language used to exemplify conceptual metaphors did not come from authentic discourse, which leads to several concerns. One is that in some cases, these examples are either not attested in corpus data or occur at very low frequencies[4]. This calls into question their validity as evidence of conceptual structuring. While it is true that the conceptual mappings can easily be understood in such language, the fact that it occurs so rarely must cast some doubt on Lakoff and Johnson's claim to be describing "ordinary" language (Lakoff & Johnson, 2003, p. 5). As a result, recent years have seen a growth in corpus-based approaches to metaphor identification and analysis (Deignan, 2005; Stefanowitsch & Gries, 2006), and these have cast doubt on some of Lakoff and Johnson's claims.

Corpus-based approaches enable researchers to investigate metaphorical patterning in authentic language data in a much more systematic, replicable fashion. No longer reliant on intuition, researchers have been able to gather empirical evidence to support or reject the classifications of conceptual metaphors (Semino, 2006; Stefanowitsch, 2006b). However, corpus evidence of the phraseological aspects of metaphor has also cast doubt on some elements of CMT. Deignan (2005, p.145 - 167; 2006), Hoey (2005, p.38 - 62) and Sinclair (1991, p.44 - 51) have all noted that words used in differing senses commonly appear in their own individual syntactic patterns. In other words, metaphorical senses have a tendency towards

[4] For example, a search for *budding theory/theories* in COCA and the BNC found no hits. A search of the enTenTen13 corpus found 7 hits, at a frequency of 0.0004 per million words (April 2017).

certain forms or collocates that differ from those of their literal senses. For example, a search of 200 concordance lines each in COCA for *seed* and *seeds* suggests that although metaphorical usage is possible in both forms (4 - 6), it is more common in the plural form. 16 instances of *seed* (8%) were metaphorical, as opposed to 25 metaphorical instances of *seeds* (12.5%). Of the 25 metaphorical uses of *seeds*, 20 (80%) were in the *seeds of* + [abstract noun] pattern, while only 6 (37.5%) of the instances of *seed* fell into this pattern.

(4) Both Iranian and American officials hoped it would be a seed for US-Iran detente. [COCA]

(5) This is the seed of drama. [COCA]

(6) Within any idea or system lie the seeds of its own destruction. [COCA]

Moreover, Deignan (2006, p. 113) found that while the literal sense of *blossom* has a strong tendency towards being a noun, the metaphorical sense occurs overwhelmingly as a verb. Taken together, these findings suggest that the preservation of cognitive topology in mappings from source to target domain that is described in the Invariance Hypothesis can often be overruled. What Deignan (2006, p. 119 - 120) proposes is a more interactive system, akin to Fauconnier and Turner's Blending theory (2003), in which both source and target domain influence the metaphorical output.

Another concern is the overlapping of conceptual metaphors. Although *budding theory/theories* is extremely uncommon, other metaphorical uses of *budding* do appear, most notably to describe career aspirations (*budding entrepreneur/artist*) as well as other abstract nouns (*budding romance*). In fact, the expression *budding*

artist is included as an example of the PEOPLE ARE PLANTS conceptual metaphor (Lakoff, Espenson, & Schwartz, 1991, p. 191). The same can be said of *offshoot*, *branch*, and *fertile*, all of which appear in their metaphorical senses with concepts other than ideas.

Examples such as these, in which source domain-related terms co-occur with target terms from several domains, were used by Grady (1997) to suggest that conceptual metaphors ran the risk of being arbitrarily named. Grady instead proposed a hierarchical organization of metaphor, in which 'primary metaphors' appear at a superordinate level to conceptual metaphors. Primary metaphors are schematic descriptions that derive from our embodied experience of the world and could potentially subsume many specific conceptual metaphors. For example, the conceptual metaphors IDEAS ARE PLANTS, THEORIES ARE BUILDINGS, and INFLUENCE IS A FORCE would all be manifestations of the primary metaphor ABSTRACT ORGANIZATION IS PHYSICAL STRUCTURE.

CMT, it should be remembered, was developed to explain abstract thought rather than metaphorical language. It holds that consistent patterns of metaphorical language are evidence of how the mind structures abstract concepts. However, it has also been argued that language is a product of social interaction (Cameron, 2003, p. 267 - 268; Cameron & Deignan, 2006). Gibbs (2013, p. 59 - 60) lists a range of forces that may influence linguistic output, including cultural, cognitive, or linguistic resources, the local environment, physical and mental condition at that time and individual motivations. Certainly, it is well documented that metaphor can be used for pragmatic and rhetorical purposes (Charteris-Black, 2004; Deignan et al., 2013),

and so a place for social context appears warranted in any theory of metaphorical language.

Does metaphorical thought, driven largely by embodied experience and our interactions with our environments, lead to linguistic metaphor, or does our regular exposure to discourse, and with it metaphorical language, influence our thought processes? Recent studies have argued that both processes may have a part to play. In a paper seeking to establish common ground between relevance theory and cognitive linguistics, Wilson (2011, p. 198 - 202) argues that inferencing is required to understand virtually all words in context, but this process becomes easier over time as repeated encounters with conceptually linked language, such as plants and ideas, lead to the development of cross-domain mappings such as those conceived of in cognitive linguistics. Empirical evidence to support this claim is provided by Bowdle and Gentner (2005), who found that metaphorical language could be processed through either comparison between the source and target terms or simple categorization. While comparison was required to comprehend novel metaphors, as the language became increasingly conventional, the processing strategy also shifted towards categorization. This result was termed 'the career of metaphor.' This echoes Giora's (1997) earlier finding that the salience of an encountered expression for a particular speaker determines how it will be processed.

Other researchers in the cognitive vein have long allowed a place for both cognitive and social factors in language development. In his work on cognitive grammar, Langacker (1987, p. 381 - 386, 1991, p. 2, 2008b, p. 218 - 221) consistently maintains that it is abstraction from exposure to language and conceptual understanding of the world that drive the acquisition of language patterns.

2.2. Metaphor and metonymy

Metaphor is a powerful device in the creation of meaning, but non-literal meaning can also be created by other processes. Metonymy is a ubiquitous influence in structuring relationships at the lexical, grammatical, and discourse levels (Barcelona, 2007; Panther & Thornburg, 2004). Unlike the cross-domain relationships that characterise metaphor, metonymy is defined as a relationship between one aspect of an entity and either that entity as a whole or some other aspect of it. For example, in (7), the Japanese city of Tokyo is being used to refer to the government of Japan, which is based in the city. That is, one prominent aspect of Japan, its capital city, stands for another aspect, the government.

(7) *In 1998, Tokyo agreed to participate in joint research with the United States on a theater missile defense system intended to protect Japan and U.S. bases located there.* [COCA]

In this way, metonymy performs a referential or focussing function in the construction of meaning. Figure 2.1 shows Langacker's (2009, p. 46) depiction of the metonymic relationship. The metonym acts as a reference point (R) that helps to locate the target meaning (T) from within the broader dominion of possible referents for a given concept.

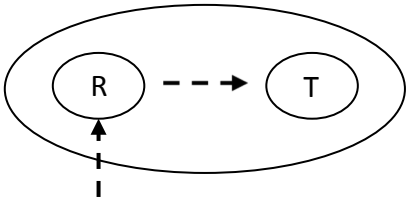


Figure 2.1. The metonymic relationship

The main focus of this thesis will be on metaphor, but the concept of metonymy is important to consider because the distinction between these tropes is often not as clear as the example above might suggest. Metaphors can be described as exhibiting a relationship of similarity in that a comparison can be drawn between domains that highlights their shared features. In this way, our understanding that seeds, despite their limited size, possess the potential to grow into larger plants enables a comparison with small events or actions that may act as stimuli for much more consequential outcomes, as in (4 - 6). Metonymy, meanwhile, can be described as a relationship of contiguity (Peirsman & Geeraerts, 2006, p. 273), where the entities involved are related in some fashion. Contiguity, however, may be conceived of in different ways as a result of human perceptions of the world (Littlemore, 2015, p. 23). The administrative apparatus of the Japanese government is based in Tokyo, providing a physical link between the city and the institution and a prototypical example of metonymy. In other cases, however, the link may be based more on a perception of co-occurrence. For example, in daily life, we frequently experience a relationship between quantity and verticality, and this motivates the use of words such as *high* (8) to describe non-physical entities. This may be seen as metonymy since concepts such as rates and prices are often depicted on charts where they are plotted on the vertical axis. If a chart is seen as belonging to the same domain as that which it depicts, we have metonymy. Conversely, we may claim that physical verticality is in a separate domain to amounts of people and determine that this is metaphor.

(8) They have extremely <u>high rates</u> of unemployment and discrimination, violence.

[COCA]

Likewise, our conception of the world as a planetary body orbiting the sun might be viewed as entirely distinct from how we see human societies in general, yet human experience leads us to see the two notions as somewhat entwined in a complementary relationship. In (9), the two uses of *world* demonstrate some of this range of meaning. The first use refers to the life experiences of a particular group of people, while the second use appears to relate to societies or the physical world in general. It might be claimed that there is stronger evidence for a crossing of domains and hence classification as metaphor in the first case, since the reference is only to individuals' lives. The second usage, however, alludes to an expanded, possibly global, scope, and we may see more of a metonymic relationship.

(9) When confined to a room in a hospital or other facility, seniors may feel that their world is growing more limited. As they mostly see medical staff and focus on their condition, they may lose sight of the broader world. [COCA]

Examples such as these suggest that metaphor and metonymy exist on a cline, rather than being completely distinct tropes (Deignan, 2005, p. 63 - 71; Radden, 2002, p. 409 - 413). An explanation of how metonymy was treated in this study is provided in section 3.4.2.2 (page 68).

2.3. Methods of metaphor identification

For many years, academic discussion centred on philosophical issues or the psychological processing of metaphor. More recently, there have been attempts to put studies of metaphor on a more secure, empirical footing by resolving the details

of how metaphorical data is obtained and the particular criteria by which it is identified.

2.3.1. Obtaining and working with data

The use of introspection in studies of metaphor has a long history, and some of the seminal texts that adopt a cognitive approach to language description have used only invented examples to argue their cases (Lakoff & Johnson, 1980/2003; Langacker, 1987, 1991, 2008b). Gradually though, as computing power has grown and allowed researchers to gather, sift and scrutinise data ever more easily, studies of attested language have begun to reveal the limitations of introspective approaches and the subtle complexity of language in use.

Using authentic data stored electronically has many advantages over introspection. At a very basic level, computers can now reliably store far more information than a human brain ever could. They also free researchers of the need to rely on the version of a language stored in the mind of a single person or a small group of people. The range of meanings and usage patterns a single word can take is subject to tremendous variation across cultures, societies and fields, and only by casting as wide a net as possible can we hope to capture this rich diversity. It has also been argued that humans are unreliable in their views on language (Deignan, 2005, p. 85 - 88; Sinclair, 1991, p. 39; Stefanowitsch, 2006b). To use Hanks' (2013, p. 5) terms, what is *cognitively salient* in our own minds may be less conventional, and therefore informative, than the *socially salient* language we use every day. Sinclair (1991, p. 39) argued that one's own impressions of language would overlook the usage patterns that are so important in communication and that could only be

revealed by looking at language in use. Sinclair did not dismiss introspection, stating that it would inevitably appear when evidence had to be weighed, but he stressed that it should play no part in creating the evidence itself.

If computers are used to store authentic language, there still remains the issue of how it is to be searched for figurative expressions. If the amount of language makes this feasible, one approach is to carefully read through the data and hand-code metaphorical expressions as they are found. This is essentially the approach taken in recent studies by Cameron (2003) and Steen et al. (2010). The obvious advantage of a hand-coding approach is that it is thorough. As long as there is a reliable, valid identification procedure, then all metaphors in a text ought to be recorded. The equally obvious disadvantage is that this approach cannot realistically be applied to large amounts of data.

An alternative to hand-coding metaphor is to use concordance software to search large corpora. Stefanowitsch (2006a, p. 1 - 6) describes possible ways this could be accomplished. These involve searching for terms from the lexical fields of the source and/or the target domain. This requires some insight into the terms that are likely to appear, which could be obtained or supplemented by inspecting a thesaurus or by building on the findings of previous research, which may include hand-coded smaller data samples (Deignan, 2005, p. 93). An alternative approach would be to make use of data that has been annotated for semantic fields. Programs such as Wmatrix (Rayson, 2008) are now able to automatically tag data into broad semantic fields which are then subdivided into narrower categories. As long as the tagging program is sufficiently accurate, researchers can then quickly identify language used to refer to particular concepts.

2.3.2. Metaphor identification criteria

The gradual move towards the use of authentic data has brought the issue of metaphor identification into the spotlight. Researchers who used intuition to create examples were able to select prototypical, inarguable cases of metaphor upon which to base their theories, but using authentic language allows for a flood of examples, probably the majority, which are far less clear cut (Cameron, 2003, p. 59).

Metaphor is found when a word or phrase appears marked or incongruous with its context but can be understood by some process of comparison with, or similarity to, the domain to which it has been applied (Cameron, 2003, p. 3 - 4). However, language is never static, and as new usage patterns drift into conventionality, their markedness can be eroded. Language is also a dynamic system, constructed online, influenced by social setting and understood through intuition and probability (Ellis, 2006a; 2006b; Gibbs, 2013). Thus, language resists categorisation, and the same word form can appear more or less metaphorical depending on its context.

Cameron's (2003, p. 58 - 66) metaphor identification through vehicle terms (MIV) approach requires investigators to identify incongruous language as it appears in contextualised discourse. It relies on repeated rounds of inter-rater checking to establish agreement, and is a valid process as the final decisions that are reached are detailed to allow for replication studies. This approach takes a dynamic view of language in that it allows for the effects of nearby context to influence a decision as to whether a particular word is coded as metaphorical or not. MIV also allows for the coding of metaphor at the phrase level. Cameron and Deignan (2006, p. 675) reject the notion that the linguistic form of a metaphor and its ideational content can be separated. Instead, they argue for a theory that accounts for the emergence of

language in a particular socio-cultural setting, taking into account the linguistic, cognitive and affective dynamics of the situation.

The Metaphor Identification Procedure Vrije Universiteit (MIPVU) (Steen et al., 2010) is a hand-coded approach that addresses the problem of metaphorical 'fuzziness' by using corpus-based dictionaries as reference sources and explicit coding rules to maintain consistency in the decision-making process. Unlike MIV, MIPVU mostly identifies metaphor at the level of individual words, although it does consider phrasal verbs, grammatical polywords (e.g., *of course*) and compound nouns as single items. The procedure has evolved through several formats[5] and has achieved good measures of reliability. Forms of MIPVU have also been used by many other researchers (Chapetón, 2010; Littlemore, Krennmayr, Turner & Turner, 2014; Nacey, 2013; Turner, 2014), allowing for external feedback to be provided. MIPVU separates language and thought, making no claims that language coded as metaphorical is actually processed as such.

2.4. Investigating metaphor in learner corpora

Early generation learner corpora began to appear in the 1970s and 80s, usually as part of investigations into error analysis. Since then, they have developed not only in size but also in design by adopting stricter selection criteria for learner ability, writing topic and sampling conditions. As such, they are now recognised as valuable

[5] Earlier formats were known as the Pragglejaz procedure and MIP. MIPVU has refined these formats by providing more precise operating procedures for identifying the unit of analysis and for using dictionaries to determine meaning, and by allowing for the identification of directly and implicitly expressed metaphors as well as indirect metaphors. MIPVU also provides a greater range of coding tags to cover issues that can arise when judging contextual meaning, such as unclear meanings and borderline cases.

research tools that can reveal aspects of learner development that had previously been overlooked (Granger, 1998; McEnery, Xiao & Tono, 2006, p. 101).

One of the most well-known learner corpora is the International Corpus of Learner English (ICLE), which was established in 1990 and has now grown to include 3.7 million words of writing from EFL learners in 16 distinct L1 settings (Université catholique de Louvain, 2011, Granger, 2003). It is supplemented with a comparison corpus of essays written by native English speakers of roughly comparable backgrounds to the learners in ICLE. This corpus, the Louvain Corpus of Native English Essays (LOCNESS), contains 324,000 words from British high school and British and American university students (Université catholique de Louvain, 2013). Major publishing companies have also developed their own corpora of learner writing. The Longman Learners' Corpus contains 10 million words of examination and essay writing from English learners of various nationalities (Pearson, n.d.), while the Cambridge Learner Corpus is a continually growing collection of exam scripts that currently stands at over 50 million words (Cambridge University Press, 2015).

Researchers have only recently turned to learner corpora to investigate developments in metaphorical knowledge. Philip (2005, 2006; 2007) has carried out studies of metaphor and phraseology based on a corpus of writing assignments by advanced level Italian learners of English. Chapetón (2010) took an approach that blended MIV and MIP to analyse metaphorical language in SPICLE, the Spanish component of ICLE. Nacey (2013) meanwhile, used a method that contains elements of MIP and MIPVU to search for metaphors in NICLE, the Norwegian section of ICLE, and LOCNESS. Finally, slight variations of MIPVU have been applied to the language of German and Greek (Littlemore et al., 2014) and French and Japanese learners

(Turner, 2014) stored in the Cambridge Learner Corpus. Taken together, the findings from these studies have confirmed that conventional metaphor appears in the written output of learners from various backgrounds and at a range of proficiencies. When the results of these studies are compared[6] (Figure 2.2), one interesting finding is that the amount of metaphor appears to increase with proficiency, to the point where some of the highest level learners produce more metaphor than the native speakers in LOCNESS. This may, however, be influenced by the writing prompt that learners were responding to. Certainly, in the case of the examinations that were the source of data in the Littlemore et al. and Turner studies, there is a progression in task demand from writing about more personal topics, such as short emails or postcards, at the lower levels to giving evaluations and hypotheses on abstract issues in an essay format in the higher levels (Cambridge English, 2017). Furthermore, it should be noted that the studies included in the table employed slight variations upon the MIPVU procedure as it was outlined in Steen et al., 2010. For example, the results from NICLE do not include direct or implicit MRWs (Nacey, 2013, p. 138 - 139), and the Littlemore et al. and Turner studies allowed for some decomposition of phrasal verbs and polywords into their component elements and permitted the comparison of meaning from lexical items that had crossed word class boundaries (Littlemore et al., 2014, p. 121 - 122; Turner, 2014, p. 69 - 77) (see section 3.4.2.2, page 68 for a description of the MIPVU procedure). It should also be remembered that the quantity of metaphor produced is not the same as any notion of metaphoric competence, a point which be discussed in section 2.5.

[6] The Chapetón (2010) results were not included in Figure 2.2 as the classification system for parts of speech included several phrase-level groups not present in the other studies, which could affect any comparison of results.

It is also possible that the kinds of metaphor produced vary with proficiency. One of the points raised in Littlemore et al. (2014, p. 127 - 128) was that the use of open-class metaphors appeared to overtake that of closed-class metaphors for learners of higher proficiency. These learners were responding to tasks that required them to construct arguments and give evaluations, which likely led to greater metaphor use, so it is difficult to separate the effect of the writing task from that of learner ability. When the data from Turner (2014) are adapted into the same format (Figure 2.3), we see that for German, French, and Japanese learners, open-class metaphors become more frequent than closed-class metaphors at the B2[7] level, while for Greek speakers this occurs at C1. The Norwegian learners in the NICLE analysis were intermediate to advanced level EFL learners (Nacey, 2013, p. 132 - 133), who also appear to use more open- than closed-class metaphors. However, an examination of the other levels quickly reveals that there is considerable fluctuation in this progression, with most groups exhibiting sudden increases in metaphor production in at least one of the levels and periods of decline also evident. This suggests a need for replication studies that may help to clarify the picture.

[7] The CEFR framework for describing proficiency is divided into three broad levels (A: basic user, B: independent user, C: proficient user), each subdivided into two sublevels to create a six-level scale (A1: breakthrough, A2: waystage, B1: threshold, B2: vantage, C1: effective operational mastery, C2: mastery) (Council of Europe, 2001, p. 22 - 23).

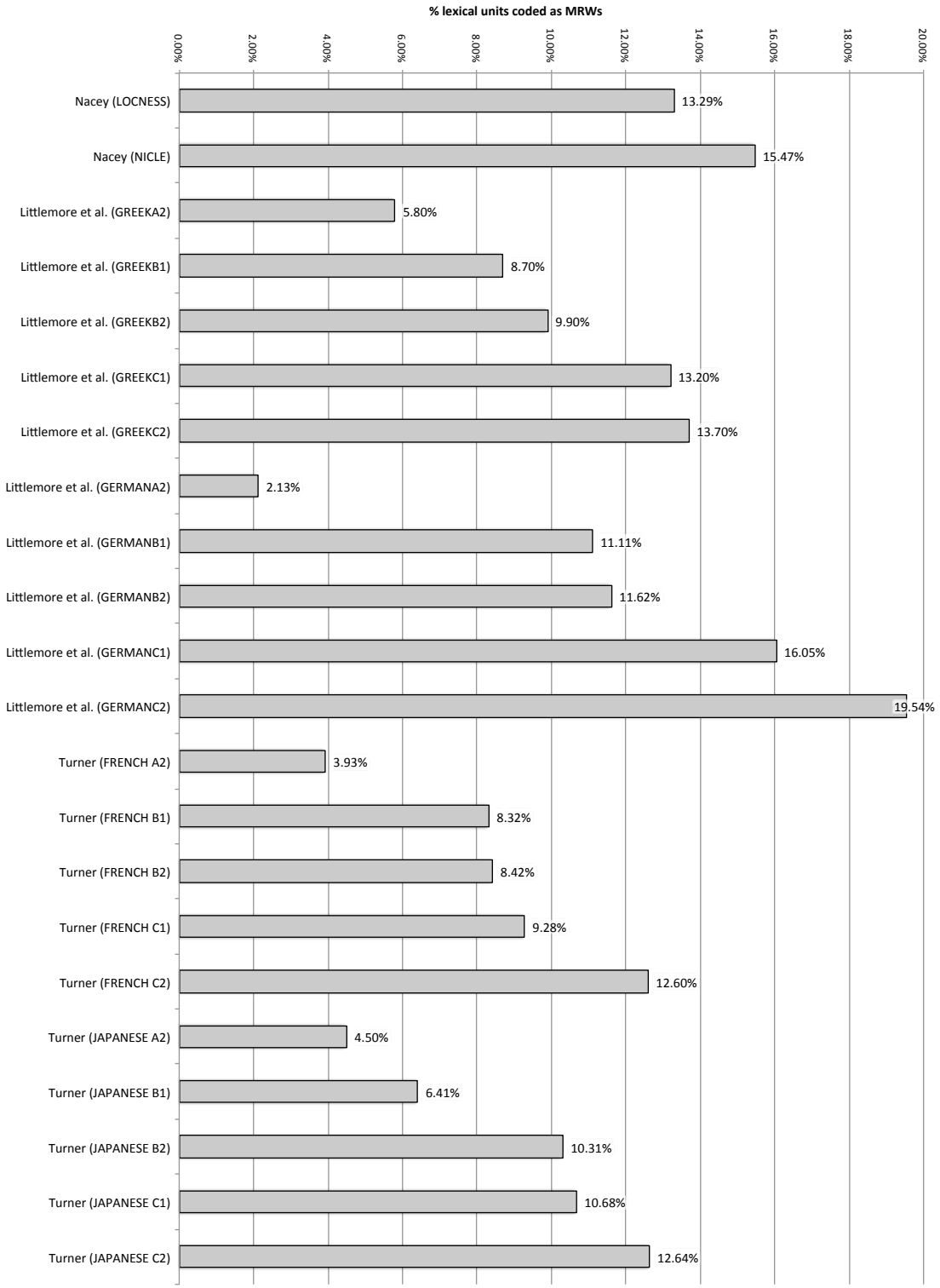


Figure 2.2. Percentage of lexical units coded as MRWs in MIPVU-related studies [8, 9, 10]

[8] Data adapted from: Littlemore et al., 2014, p. 125 - 126; Nacey, 2013, p. 136 - 138, 149; Turner, 2014, p. 90-91)

[9] The NICLE and LOCNESS studies could be considered comparable by topic. The A2, B1, B2, C1, and C2 studies could be considered comparable by ability.

[10] The NICLE data does not count uses of *of* or *for* as MRWs. This is to harmonise the coding with other studies (see Nacey, 2013, p. 137 - 139).

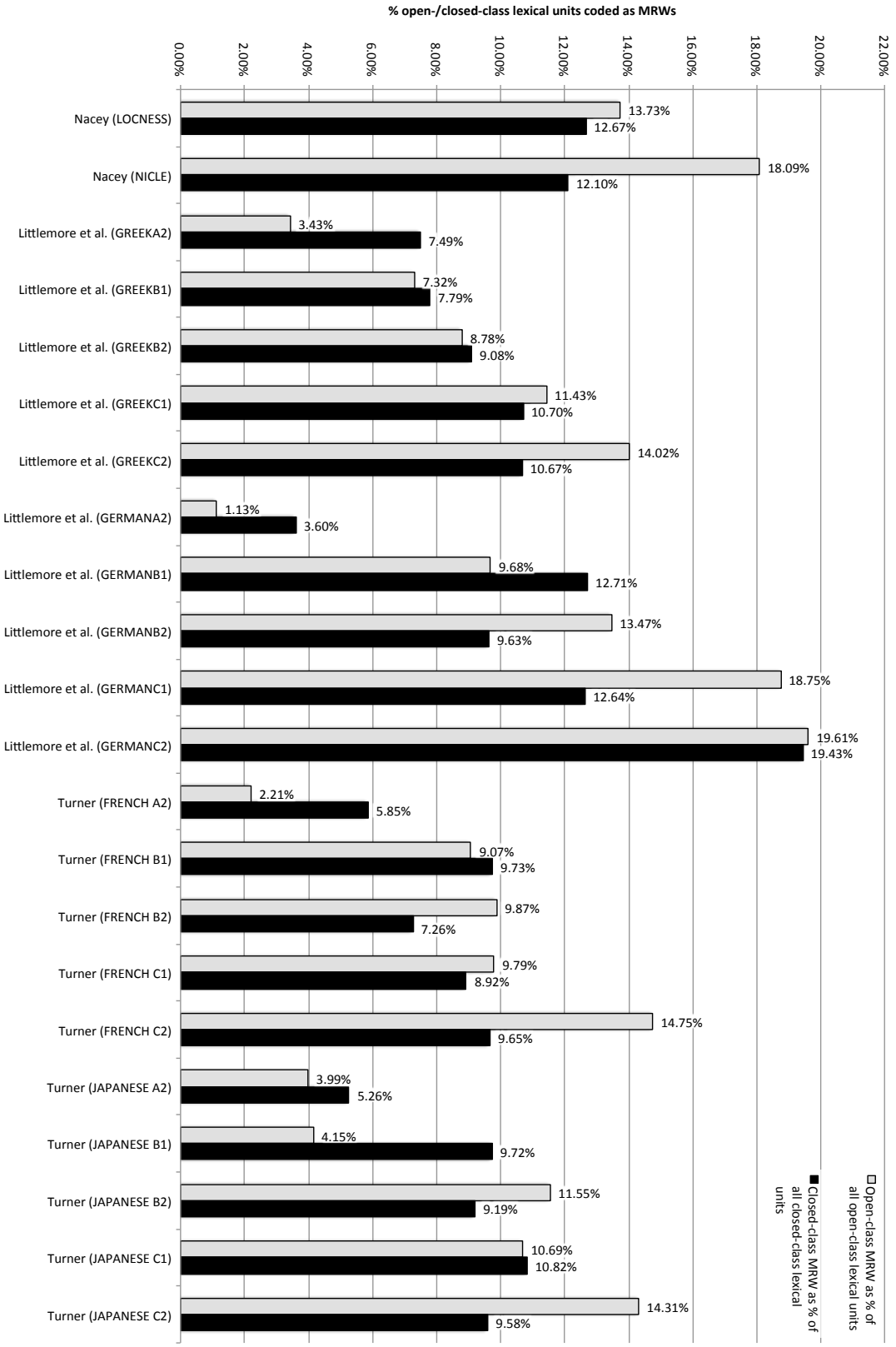


Figure 2.3. Percentage of open- and closed-class lexical units coded as MRWs in MIPVU-related studies[11]

[11] Data adapted from: J. M. Littlemore, personal communication, September, 2016; Nacey, 2013, p. 136 - 138, 149; S. Turner, personal communication, March, 2017

2.5. Metaphoric competence

Thus far, we have considered learner metaphor in terms of the quantities of metaphorical words produced in writing, but this is clearly only a partial view of what is a much broader skillset. Littlemore and Low (2006b) and Low (1988, p. 129 - 135) consider metaphoric competence to also include: being able to use metaphor in such a way that the intended message is understood, having an awareness of when metaphor goes beyond what is conventional into idiosyncratic or stylistic uses, recognition of how topic and vehicle terms are typically combined, ability to mark or hedge metaphors, sensitivity to the social impact of some metaphors, appreciation of how metaphors can encode multiple meanings and understanding how to use metaphor when interacting with others.

These competencies range from the basic to the more advanced, but it is not difficult to see how language learners may need most, if not all, of these forms of knowledge in academic settings. Certainly, learners will need to be able to convey their thoughts on abstract matters, which by necessity will entail using metaphor in ways that closely approximate the typical patterns of the L2. As they become more proficient, it might be expected that learners begin to make use of the rhetorical functions of metaphor as they seek to interact more purposefully with readers and listeners. At higher proficiencies still, we may see signs of rhetorical mastery with learners employing metaphor creatively and with appropriate signalling.

Another way of viewing metaphoric competence is through Bachman's (1990, p. 84 - 98) framework of language competencies, which places greater emphasis on accomplishing communicative goals. Littlemore and Low (2006b) use this framework

to stress the value of metaphor knowledge in performing evaluative, persuasive, heuristic, and imaginative functions as well as text organisation. When viewed this way, the ubiquity of metaphor becomes more apparent, and metaphoric competence can be seen as playing a role in many of the essential elements of language use.

Thus competence can be seen in terms of both knowledge of metaphor and awareness of how to use metaphor. But we should also be aware that metaphoric competence stretches beyond linguistic knowledge into the more general cognitive dispositions of learners. Individuals with particular learning styles may be more or less suited to drawing on metaphor as a vehicle for comprehending the world and for self-expression (Johnson & Rosano, 1993; Littlemore, 2001).

It is clear that the influence of metaphorical thinking is potentially widespread and is far from limited to matters of language learning. However, while this does emphasise the importance of metaphoric competence, sheer pervasiveness may render it a challenging research construct as it can be seen and defined in different ways depending on the context. At this point, many linguistic studies into acquisition of metaphor have focused on isolated elements of language such as phrasal verbs or idiomatic phrases, but there have been few attempts to examine learners' functional use of metaphor or studies that have looked at metaphor as an emergent feature of learners' growing lexical competence.

2.6. Cognitive linguistic approaches to teaching metaphor

Metaphor has long been of interest to cognitive linguists. One obvious reason for this is that figurative expressions might be considered somewhat arbitrary in their

form and meaning, and therefore can only be learned through memorisation (Boers & Lindstromberg, 2008, p. 18; Littlemore & Low, 2006a, p. 95). This contradicts one of the central tenets of cognitive linguistics (CL), which is that all language is explicable to some degree. Accordingly, many studies using CL-inspired methodology have been carried out to raise awareness of the motivating factors that underlie figurative expressions (e.g., Beréndi, Csábi, & Kövecses, 2008; Boers, Lindstromberg, Littlemore, Stengers, & Eyckmans, 2008; MacArthur & Littlemore, 2008).

From a language learner's point of view, conventional metaphor presents an unusual challenge. It is not a topic commonly dealt with in published textbooks, and may be overlooked in favour of more traditional aspects of a language curriculum such as grammar structures, topic-themed vocabulary and four-skills based activities. One reason for this may be that due to time constraints in language programmes, there is often pressure to cover new material. Learning about conventional metaphor, however, often involves learning more about words that have already been covered to some degree. Metaphorical extension is among the principal forces that drive polysemy (Littlemore & Low, 2006a, p. 23 - 24), and metaphorical expressions composed of relatively simple words have been shown to present serious comprehension difficulties for language learners in academic environments (Littlemore, Trautman Chen, Koester, & Barnden, 2011).

In order to address this gap in EFL pedagogy, Boers (2000b, p. 566 - 568) suggests five aspects of metaphoric awareness that should be brought to learners' attention.

- Metaphor appears frequently in everyday language.

- Conceptual metaphor underlies many metaphoric expressions.
- The meaning of many metaphoric expressions can be explained – they are non-arbitrary.
- Different cultures may use different conceptual metaphors, or if the same conceptual metaphor exists, it may have far greater prominence in one culture than another.
- Even if the same conceptual metaphor exists in two cultures, it may be realised by different linguistic expressions.

As well as awareness of metaphor in language, learners will also require specific knowledge of metaphors in order to understand and use them appropriately. Cameron and Deignan (2006, p. 687) list three forms of metaphor knowledge: linguistic, conceptual and pragmatic. Linguistic knowledge is an understanding of the form or forms that the metaphor takes. Conceptual knowledge would require awareness of the underlying conceptual metaphor as well as its particular real world referents, which may be more limited than the conceptual metaphor itself implies. Finally, the ability to use metaphors appropriately in discourse requires pragmatic knowledge.

Another tenet of CL is that language learning does not occur in any separate faculty of the mind, but rather it involves the entire range of human cognitive processes (Langacker, 2008a, p. 5 - 9). Accordingly, CL methodology makes use of several theories drawn from educational and cognitive psychology (Boers & Lindstromberg, 2008, p. 11 - 13). The first of these is 'dual coding theory', which holds that generating mental images to accompany verbal information will aid

recollection (Paivio, 1969). The Keyword method of vocabulary learning makes use of this theory. 'Trace theory' (Baddeley, 1997) meanwhile, states that encountering a learning opportunity many times will strengthen the trace that is left in memory. This is the theory underlying the 'spaced retrieval' approach to memorization that is used in some online vocabulary learning systems. Lastly, 'levels of processing theory' (Craik & Lockhart, 1972) posits that greater mental involvement, or elaboration, with a learning task will increase the likelihood of the learning moving into long-term memory.

In general, classroom-based CL studies tend to find greater gains for experimental groups over control groups taught using more traditional methods (e.g., Beréndi, Csábi, & Kövecses, 2008). However, there is a lack of evidence for the long-term benefits of CL methods, as the results of delayed tests have sometimes been disappointing (e.g., Boers, 2004, p. 215 - 217), or because learners in experimental groups have been found to perform better in controlled studies than in experiments that require them to transfer knowledge of metaphor so that they can become more autonomous (Condon, 2008, p. 150 - 151; Skoufaki, 2008, p. 118). Boers (2004, p. 216) concludes that one-off treatments using CL methods are insufficient to yield long-term benefits, but there are suggestions that if these techniques are repeatedly brought up in classes, then retention might be improved. However, few longitudinal studies have been conducted at present, and there is a clear need to continue this line of research.

Taking a cognitive view of language learning has also raised awareness of several other issues, such as language transfer, learner ability, individual differences, and

pedagogy, that should be considered in order to maximise the effectiveness of CL methods in the classroom.

2.6.1. Issues of transfer

Although the theory of contrastive analysis (Lado, 1957) has endured changing fortunes, it is still widely accepted that a learner's L1 has a great influence on the learning of subsequent languages. This is especially true for metaphor because, as a vehicle for the expression of abstract thought, it is particularly likely to be affected by differences in conceptual norms and their realisations in language. Fortunately, since cognitive linguistics allows a place for construal and 'fuzzy' prototype categories over the relatively fixed classifications of more traditional views of language, it is well placed to offer explanations of how and why L1 influences affect L2 acquisition (Boers, 2003).

One way that transfer can affect metaphorical understanding is through differences in conceptual structures. Cultures vary in the prominence they give to particular concepts in language. For instance, it has been reported that compared with other languages, English contains a relatively high proportion of figurative expressions related to sailing and to hats, French uses a large number of food-related expressions, and Spanish makes frequent use of religion or superstition in its idioms (Boers & Demecheleer, 2001; Boers & Stengers, 2008). Deignan (2003, p. 259) has suggested that this may be due to geographic and cultural differences. Another observation is that the same source domain may be drawn into different conceptual mappings. Deignan and Potter (2004, p. 1248) found that the word *mouth* was used in a greater variety of expressions related to eating in Italian than in English.

These variations may affect both receptive and productive language skills. Littlemore (2003) found that Bangladeshi students attending university in the UK were prone to interpreting metaphors in lectures according to the values of their own culture. This led to misunderstandings of both basic content and the evaluative stance of the lecturer towards the material being presented. For language production, it has been noted in several studies that even if two cultures share a conceptual mapping, there is no guarantee that the actual linguistic realisations of that mapping will correspond (Beréndi et al., 2008, p. 88 - 89; Chen & Lai, 2013, p. 18; Deignan & Potter, 2004, p. 1248). Awareness of the existence of a conceptual metaphor does not entail understanding conventional usage (Boers, 2004, p. 217).

It has also been suggested that the distance between L1 and L2 may affect learners' interpretation and production of metaphorical language. Learners whose L1 is closely related to English have been found to perform well on comprehension tests of English metaphors (Boers, 2000b). However, there may be disadvantages to linguistic similarity for language production. Boers (*ibid*, p.563) also points out that there is a tendency for learners who perceive their L1 to be close to the L2 to overuse phrases that they are confident are acceptable in both languages. This has been observed even in advanced learners, such as the Norwegian university students investigated by Hasselgren (1994), who coined the term 'lexical teddy bears' to describe their habit of playing safe in their language production. Speakers of relatively unrelated languages or those who are from more distant cultures, however, such as Taiwanese or Japanese learners of English, may suffer from less conceptual overlap with their L1 (Chen & Lai, 2013, p. 17 - 18), leading to difficulties in both understanding and producing metaphors in a conventional fashion.

A further complication in the case of Japanese learners is that their own language now contains an enormous number of loanwords, most of which come from English[12]. Among these loanwords are many expressions which are metaphors in English, such as ホイッスルブロウワー (*hoissuruburouwa-*, whistle blower) and ダークホース (*da-kuho-su*, dark horse). These expressions carry the same meanings as in English, but metaphorical borrowings can also undergo semantic change, as in ボディーブロー (*bodi-buro-*, body blow) which only retains the meaning of a punch to the body, and blending with original Japanese words, as in かおアップ (*kao appu*, literally 'face up', meaning a close-up photo of a face). Certain phrases known as *wasei-eigo*, or Japanese-derived English, may appear to use English words but are in fact Japanese creations. These include ゴーイングマイウェイ (*go-ingumaiuei*, 'going my way', or being an independent thinker) and ヴァージンロード (*va-jinro-do*, 'virgin road', or the aisle in a wedding ceremony). The effect of these expressions on learning English has yet to be explored in research.

2.6.2. Issues of learner ability

Few studies have directly addressed the question of what level of proficiency learners should reach before metaphor is introduced in the classroom. This issue could be approached in several ways.

Firstly, if we consider the language that learners are exposed to, then it could be argued that raising awareness of metaphor should begin in the elementary stages of

[12] Daulton (2008, p. 11-13) has estimated that potentially more than 50,000 words have been borrowed from Western languages into Japanese, around 10% of the entire lexicon. The proportion of imported words that come from English grew from 51.9% between 1912 and 1925 to about 90% by 1975 (Shibatani, 1990; Yazaki, 1975; cited in Daulton, 2008, p. 12).

learner development. The concept of time and language such as prepositions and phrasal verbs typically appear in introductory courses, and as Boers (2004, p. 221) points out, “Whenever genuine communication takes place, abstract ideas may be expressed, and thus figurative language may be needed.”

A second consideration is the language that might be required by learners to accomplish tasks aimed at raising metaphor awareness. If learners are required to interpret metaphors by themselves, then clearly they will need to be familiar with the basic sense of any words they encounter (Boers, 2004, p. 221 - 222). If however, learners are presented with metaphors using unknown language, then an explanation of the basic sense of the word may help to aid retention. Once learners have core vocabularies, then they may be able to classify metaphorical expressions or to speculate about the expressions that realise a given conceptual metaphor.

Finally, learner corpora (section 2.3) could be examined to provide evidence of the metaphorical language that learners at a range of proficiency levels are capable of producing. Although this approach requires overcoming numerous practical issues, not least the collection of a large sample of suitable learner texts and identifying metaphors therein, it does have the potential to offer insights that other methods cannot. Littlemore et al. (2014) identified metaphor produced by native German and Greek speakers learning English. For both language groups, they sampled 100 essays drawn equally from five test-taking levels in the Cambridge Learner Corpus. Their study found a gradual increase in metaphorical language as ability improved (Figures 2.1 and 2.2). Lower level learners (CEFR level: A2, TOEIC: 225[13]) typically only used

[13] Minimum TOEIC scores that correspond to each CEFR level were produced by ETS (Tannenbaum & Wylie, 2006, 2008). They are included here to aid comparison, since TOEIC scores are the most

metaphor in prepositions and fixed expressions. Students of low-intermediate to intermediate ability (CEFR: B1, TOEIC: 550) were beginning to use metaphorical senses of basic vocabulary to express their own perspective. While the total amount of metaphor produced at the higher-intermediate level (CEFR: B2, TOEIC: 785) increased only a little, both language groups began to use more metaphors with open-class items than closed-class items, and some learners appeared able to use metaphor creatively and to perform discourse-organising functions. More advanced learners (CEFR: C1 and C2, TOEIC: 945 and above) showed an ability to use phraseologically correct metaphor, and to apply metaphor for persuasive or rhetorical purposes. The rich detail provided by studies based on learners' actual written output is invaluable for understanding how language skills develop over time. Language and language learning are both complex systems, and if we focus on broad results such as the total amount of metaphor produced, we may miss more subtle patterns in the data, as has been shown by this study. Groom (2009, p. 32 - 33) observed a similar tendency for learner production of recognised collocations to actually decrease as they gained in proficiency, simply because the learners overcame their habit of repeating phrases and broadened their range of expression.

In general, these studies support the use of metaphor awareness-raising activities in the classroom from relatively early stages in learners' development. During the early stages of learning, when core vocabularies are still forming, learners can be introduced to some of the extended senses that high frequency words are able to take on. Drawing attention to the metaphorical origins of certain

---

commonly used measure of English ability in Japan and were used as a measure of proficiency for the participants in this study (see section 3.6.1, page 100).

grammatical features, such as using *going to* to express future meaning (Radden & Dirven, 2007, p. 224) may also help to ground abstract structural patterns into concrete experience for beginners. By focussing on language use as it truly is or how it has evolved to be, these awareness-raising activities could help to prime learners to be more attentive to metaphor in the future.

2.6.3. Issues with learner differences

The differences that learners bring with them to the classroom have been widely recognised as having an impact on learning outcomes. Dörnyei (2005) describes several factors that may affect language learning, including personality, aptitude, motivation, cognitive style and learning strategy use. Of these, the factor that has been most studied in relation to metaphorical competence is cognitive style. A learner's cognitive style reflects their tendency to approach learning tasks in a particular manner. While many different style dimensions have been identified, this has led to debate over the overlapping of categories and the validity of a distinct style construct (Dörnyei, 2005, p. 124 - 129). Consequently, much research in this area now conflates cognitive style into two broad dimensions: wholist/analytic and verbaliser-imager. Riding & Cheema (1991, p. 197 - 205) detail the wholist/analytic dimension as shown in Table 2.1.

Generally, analytic cognitive styles have been seen as preferable traits for language learners, but because metaphorical competence can encompass drawing links between separate entities to create blended meanings, it may also lend itself to holistic learning styles (Littlemore, 2001). Analytical thinkers, meanwhile, may prove to be more successful at tasks that require them to identify correspondences and

Table 2.1. The wholist/analytic dimension of cognitive style[14]

Wholists	Analytics
Field dependent – focus on the whole situation rather than the details	*Field independent* – able to separate details from their backgrounds
Impulsive – make quick responses	*Reflective* – slow, thoughtful responses
Levellers – assimilate information rapidly, possibly losing detail	*Sharpeners* – perceive individual details discretely
Divergers – original, associational thinking; solve open-ended problems	*Convergers* – logical, narrow thinking; solve problems with clear answers
Holists – scan large amounts of data searching for patterns	*Serialists* – search small amounts of data in a step-by-step fashion

relationships between source and target domains (Boers & Littlemore, 2000, p. 183 - 184).

The verbaliser/imager dimension relates to how individuals tend to process information and interact with their environment. Verbalising involves processing information as words and seeking stimulating external environments, whereas imaging involves thinking in pictures and is more inwardly focussed (Dörnyei, 2005, p. 129). Most people demonstrate aspects of both verbal and imaging behaviour as the situation warrants. In metaphor research, Boers and Littlemore (2000, p. 184) found that learners identified as strong imagers were more likely to describe conceptual metaphors in terms of stereotypical scenes drawn from the target domain.

There is still much to discover about the relationship between cognitive style and metaphorical competence. Although Littlemore's (2001) study found that wholists were able to process metaphors more quickly than analytics, there were only weak correlations between different aspects of metaphorical competence. This suggests that various learning styles may lend themselves to different aspects of metaphorical understanding, and that competing theories of metaphor may apply

[14] Adapted from: Dörnyei, 2005, p. 127 - 128; Riding & Cheema, 1991, p. 197 - 205

better to some learners than others (Boers & Littlemore, 2000, p. 184). Perhaps the only firm conclusion that can be drawn for classroom approaches at present is that a variety of approaches may be needed to suit different learners.

2.6.4. Issues of pedagogy

A wide variety of methods are available for raising learners' awareness of metaphor. In the light of the previously mentioned findings suggesting that cognitive style can greatly influence learner performance, it seems pedagogically sound and ethically appropriate to utilise a variety of teaching methods so as to cater to the full range of learning preferences that may exist in a classroom. Consideration should also be given to learner ability. Although there is little empirical data on which decisions could be based in this regard, the authors of some studies have recommended suitable levels of proficiency for the methods they investigated. Further research along the lines of that carried out by Littlemore et al. (2014) would be of great benefit in helping to clarify the specific purposes that learners of varying abilities are able to use metaphorical language to achieve.

The following section will summarise a variety of learning experiences that can be used to bring metaphor into the classroom and the findings of studies that have used these techniques. The activities have been grouped to reflect the basic cognitive skill that learners would be required to use or would be exposed to. However, inevitably there will be some overlap between groups. At heart, all of the activities are aimed at raising awareness of the systematicity and motivation that underlies lexical knowledge.

Explanation / elicitation

Probably the simplest way that metaphor can be introduced is for teachers to devote class time to explaining or eliciting the motivation underlying non-literal language. Alternatively, written worksheets could be presented that explain this motivation. While this approach is largely teacher-centred and does not make use of deep processing strategies, its great advantage is that it is flexible. It can be applied whenever suitable language appears in the lesson, and if used in combination with other techniques, would allow for regular review. Several studies have pointed to the advantages of this simple technique over methods that present language only referring to its target meaning, without mentioning the basic or original sense (Beréndi et al, 2008; Boers, Demecheleer, & Eyckmans, 2004; Verspoor & Lowie, 2003).

Visualisation

It has been held that activities which make use of dual-coding will facilitate recall (Boers & Lindstromberg, 2008, p. 27 - 37). Images that illustrate the basic meaning of a figuratively-used word have been shown to provide a mnemonic boost to learners who tend towards a verbal, rather than an imaging, cognitive style when they are used alongside text (Boers et al., 2008, p. 204). However, the benefit of using images appears to be far greater for recall of meaning than for the form of an expression (Boers, Piquer Píriz, Stengers, & Eyckmans, 2009). In fact, images may impede the recollection of form for multi-word expressions and more difficult vocabulary. The general conclusion is that while images can aid recall, if learner production is the

goal, then having an activity stage in which learners can attend to the form of an expression without the distraction of an image may be preferable (Boers et al., 2009).

Classification

Another simple procedure that could be used with lower-ability learners is classifying expressions according the conceptual metaphors that they instantiate. This is a technique that is ideal for situations in which course content dictates the language that is likely to appear, such as in academic settings or content-based instruction programs. Fields with specialist terminology that is derived metaphorically, such as economics, are particularly likely to benefit from such an approach (Boers, 2000a; Charteris-Black, 2000).

Comparison

One flexible strategy that can be used with learners of all abilities is comparing metaphor between the L1 and L2. At its simplest, the technique could just involve asking students whether similar language is used to express the same concepts in their L1. This would be useful for raising awareness of which conceptual metaphors are equivalent between the two languages. A more challenging task was used by Deignan, Gabryś and Solska (1997), who asked Polish students to translate metaphorical sentences in English into their mother tongue, and then to discuss the similarities and differences between the expressions in the two languages. The students found instances where the same metaphor and linguistic expression were used, where the same conceptual metaphor was realised by a different linguistic expression, where the conceptual metaphor differed, and where similar linguistic expressions realised different metaphors.

Investigation

One possible technique that appears to have only recently attracted interest (Philip, 2006) is the use of reference materials to raise awareness of metaphor. Conventionalised metaphors will appear as separate senses in standard dictionaries, and dictionaries of phrasal verbs and other figurative language also exist. Although this method does not involve deep processing, it could be used as a confirmation exercise in combination with other strategies, and would also serve to raise awareness of dictionary use, particularly for low-level learners.

Application

In order to facilitate transfer from receptive to productive knowledge, teachers can provide open-ended tasks in which students are encouraged to experiment with metaphorical language they have just learned. Chen and Lai (2013) report that Taiwanese learners of English were able to use various idiomatic expressions related to anger in expository essay writing after having been shown the underlying conceptual metaphors ANGER IS FIRE and ANGER IS A HOT FLUID IN A CONTAINER. Furthermore, their efforts provided feedback for the instructor, who could gain insight into students' understanding of particular conceptual metaphors and the degree to which their L1 influenced their output. It is also likely that phraseological issues will appear in written output, which can then be taken up by instructors. Many other activities that promote production of metaphorical language can be found in Lazar (2003).

Deduction

Providing learners with contextualised metaphorical language and requiring them to deduce the underlying conceptual metaphors is a challenging task, but one that promotes deeper processing and that can be applied in a variety of ways. One is that learners could be provided with concordance lines containing highlighted metaphors. This was the approach taken by MacArthur and Littlemore (2008), who found that presenting both transparent and opaque uses of the same metaphors together may help learners to appreciate the systematicity operating within language. Overall, the study raised many questions that require further investigation, particularly relating to learners' cognitive styles, the influence of form on understanding, and the presentation format of the concordance lines. In order to provide more helpful contexts from which to deduce meaning, an alternative method would be to produce a contrived text (e.g., Boers, 2000a, p. 141; Lazar, 2003, p. 41, p. 83).

Another form of activity that requires deductive reasoning is when learners are first provided with the etymological background to an expression (i.e., its original literal meaning) and are then required to deduce the metaphorical meaning (Boers, Eyckmans, & Stengers, 2007).

Analysis / evaluation

Higher-level aspects of metaphorical competence, such as rhetorical or persuasive use, may be more applicable for teaching to high-intermediate or advanced learners. This skill would involve the ability to critically analyse metaphor in writing or other media, to consider the reasons behind its use, its success in achieving the desired

effect, and possible alternative metaphors. Lazar (2003) presents several activities that involve the analysis of metaphor in advertising, journalism and poetry.

Attention to form

One of the key findings of corpus linguistics has been that language exhibits far more phraseological patterning than had previously been thought (Hanks, 2013, p. 140; Hoey, 2005, p. 5 - 14; Sinclair, 1991, p. 102 - 104). Metaphor is no exception to this, and it has been shown that in naturally occurring discourse, figurative and literal senses are marked by the restricted and distinct collocations and colligations to which they are bound, and which in turn have clear semantic and prosodic effects (Cameron & Deignan, 2006, p. 678 - 686; Deignan, 2005, p. 157 - 167). Since metaphorical competence includes awareness of the form and pragmatic functions of figurative language, it follows that phraseological patterning should be brought to learners' attention – to teach metaphor is to teach phraseological patterns. However, perhaps due to the complexity of presenting language in such detail, aspects of figurative patterning have yet to gain prominence in published teaching materials. From a pedagogical point of view, consideration also needs to be given to learner avoidance of language that is perceived to be difficult and the calquing of L1 patterns into the L2 (Philip, 2007).

Thus it might be said that while CL studies have shown promising results in raising learners' awareness of metaphor, they have also drawn attention to a large number of other influences that are present in the classroom. But it should also be remembered that the issues that were described in this section are concerns that apply to other aspects of language learning as well. Vocabulary instruction, in

particular, is likely to face issues with L1 influences and variations in learning styles, and the awareness-raising approach favoured by Boers (2004) for metaphor learning might be effectively integrated into vocabulary instruction more broadly.

2.7. Vocabulary knowledge and metaphor

Vocabulary itself is often described in metaphors that emphasise certain facets of lexical knowledge and obscure others. The breadth/depth distinction sees vocabulary as a multidimensional construct that is expected to grow in different ways as learners develop. It also implies that vocabulary is amenable to measurement. Metaphor can be considered part of the depth aspect since this reflects the range of meanings a word form can take. However, this view of vocabulary has drawn criticism for failing to explain how the two aspects are related and indeed if they are separate at all. Vermeer (2001) has argued that there is essentially no difference between breadth and depth, while other studies have found that the degree of interrelation increased with proficiency (Nurweni & Read, 1999, p. 170 - 171). This suggests that lower ability learners with smaller vocabulary sizes also lack deep knowledge about the words they do know. As they gain in proficiency, however, they add not only new words, but also extend their knowledge of previously known words. In this view, metaphorical meanings typically might not be known at first, but successive encounters with different senses of words will stretch the boundaries of word knowledge to encompass figurative meanings.

Another view of vocabulary posits a network of connected word knowledge (Meara, 2009). New words appear as nodes in the network and gradually acquire more connections with other words (such as collocates, colligates, and synonyms)

through repeated exposure, becoming more integrated into the overall system each time. This is a more complex image but one that allows for the dynamic and non-linear nature of vocabularies to come more to the fore. It also accommodates some aspects of metaphor more easily, as the different phraseological patterns of each sense of a word could be expressed with separate sets of connections. However, it is not easy to depict a range of meanings being immanent in a single node in a network, which is obviously a crucial feature of metaphor.

The cognitive linguistic concept of radial 'fuzzy' categories (Littlemore, 2004, p. 41 - 57) may be of use here. Rather than seeing a word as a particular unit, radial categories propose a basic sense of a word surrounded by less prototypical (and often figurative) senses, as is shown for the word *barren* in Figure 2.3. This view of lexical items allows for the nodes in a system to contain a potential range of meanings – a more realistic proposition for describing vocabulary.

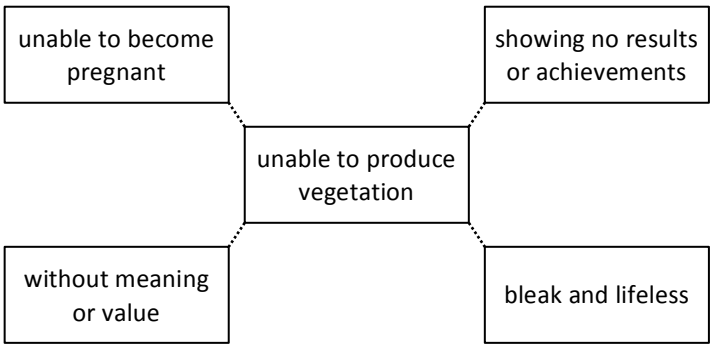


Figure 2.4. Radial category for *barren*

Taking these final two views of vocabulary in combination, we have a system of nodes within nodes, just as atomic molecules are themselves composed of smaller elements. The various senses of lexical items are each linked with their own network of syntactic and paradigmatic relations. Lexical prowess would thus be measured not

only in the size of this network, but also in its degree of connectedness. Many researchers have explored elements of this view of the lexicon with word association tests and attempts to measure lexical organisation (Fitzpatrick, 2009; Higginbotham, Munby, & Racine, 2015; Meara & Wolter, 2004), but it is probably fair to say that such instruments have yet to reach widespread use. The more plausible depiction they offer of the lexicon may be somewhat counteracted by a preference among teachers for a neat 'vocabulary score', which may explain the enduring appeal of the breadth/depth view of vocabulary.

In truth, there is still a dearth of evidence for how the mental lexicon develops as learners gain proficiency. Studies on vocabulary size, or breadth, are beginning to appear more frequently, but only a few longitudinal findings have been published (Cobb & Horst, 1999; Dóczi & Kormos, 2016, Chapter 2; Ozturk, 2016; Webb & Chang, 2012), and interpretation of results can be a challenge given the variability of learning contexts. Investigations into other aspects of lexical knowledge have been hampered by the challenge of developing reliable instruments, and the branch of research into metaphorical vocabulary growth is just beginning to form.

2.8. Action research and other approaches

As a classroom-based investigation in which the author took on the roles of both teacher and researcher, this study exhibits several of the hallmarks of action research, an approach that has attracted considerable attention in language education. Action research is typically described as a cyclical process instigated by the noticing of an issue that is impeding learning or preventing classes from proceeding in a satisfactory manner. There then follows a preliminary stage in which

evidence is gathered, which leads to the development of a hypothesis or action plan intended to address the initial concern. This plan is then implemented and evidence is again collected to assess the impact of the classroom treatment. Depending on the efficacy of the intervention, further rounds of modification and intervention may be required, or the outcome may be deemed successful, and the results may be shared with colleagues or the broader community (Edge, 2001, p. 3; Nunan, 1991, p. 3 - 14).

While action research commonly features this cyclical pattern, it is in fact a broad term that encompasses a spirit of reflective teaching and a concern for practitioner welfare. This aim to produce better qualities of life reflects the fact that action research was first employed in social psychology and community activism in the early twentieth century (McTaggart, 1994, p. 316). Action research is sometimes described as having a dichotomous relationship with traditional research, a view that sees practice and theory as separate endeavours (McDonough & McDonough, 1997, p. 23). Indeed, this is something of a long-running debate; the educational reformer John Dewey wrote passionately on the value of drawing on both experience and rationality to guide educational change:

> *Experience is no longer a mere summarizing of what has been done in a more or less chance way in the past; it is a deliberate control of what is done with reference to making what happens to us and what we do to things as fertile as possible of suggestions (of suggested meanings) and a means for trying out the validity of the suggestions. When trying, or experimenting, ceases to be blinded by impulse or custom, when it is guided by an aim and conducted by measure and method, it becomes reasonable – rational. When what we suffer from things, what we undergo at their hands, ceases to be a*

> *matter of chance circumstance, when it is transformed into a consequence of our own prior purposive endeavors, it becomes rationally significant—enlightening and instructive.* (Dewey, 1916, p. 319)

Edge (2001, p. 4) picks up this theme and stresses that the findings of action research must be linked to rational argument to further our understanding of human education through insights into the dynamics of a particular context. The question is whether this should be seen as a reversal of the *outside-in* orientation of traditional research (Prawat, 1991, p. 740 - 743) or as establishing a more *reflexive* relationship between theory and practice (McDonough & McDonough, 1997, p. 23). Elliot (2004, p. 20) argues that the traditional separation of theory, the quest for knowledge for its own sake, and action, a motivation to improve circumstances, is invalid, and that since human knowledge is filtered by cultural influences, research needs to take account of the conditions in which education is taking place. Action researchers should therefore draw on theory to inform their work, but such understanding needs to be tempered by the unique demands of each situation rather than slavishly adhered to.

The value of contextual awareness is often cited as one of the benefits that action research can offer to practitioners. This line of thinking suggests that particular contexts themselves can offer the most profitable solutions to issues that emerge in the classroom (Edge, 2001, p. 3). If theory is taken to be the most generalizable description of a situation, then this may come at the cost of a loss of consideration for learners' L1 backgrounds, their individual learning styles, or the particularities of the course or curriculum in which data were collected.

Another prominent feature of action research is its goal to strive for improvement in practitioners' lives. It should be noted that much traditional research also aims to address educational issues or refine methodology and practice, although this is sometimes left implicit in the authors' message. Where action research might be seen to differ more clearly is in its claimed goal to empower participants in order for them to lead more fulfilling lives, whether that be through greater professional engagement on the part of teachers or by helping learners to achieve their objectives (Edge, 2001, p. 4). This view describes action research as an emancipatory exercise in which teachers are encouraged to reflect on and potentially challenge the assumptions of national, institutional, or even personal educational policy.

Allwright (2005) took the emphasis on improving lives further with his Exploratory Practice paradigm, which regards action research as a placing an unfair burden on teachers to acquire essentially the same investigatory skills as academic researchers. He further charges both traditional and action research as having a parasitic relationship with the classroom environment, by taking up valuable time, resources, and energy that might be more profitably spent on the more regular demands of teaching and learning a language (2005, p. 355).

One criticism that could be levelled against action research is that in addressing an issue relevant to a certain context, it loses relevance to practitioners in other situations. While the techniques it employs may not differ noticeably from those of traditional research, the findings it obtains are valued for their validity in that particular context rather than being widely generalizable (Crookes, 1991, p. 77 - 80). This view may be rooted in the positivist philosophical position which sees theory as

being derived from logical principles linked to observed verifiable outcomes. However, it has been argued more recently that neither observations nor theory can be considered entirely free from practitioners' worldviews, and that knowledge developed amidst the complexity of a social environment is as legitimate as any other kind (Johnson, 2006, p. 239; Messick, 1988, p. 4 - 5). As the Dewey quote above claimed, when research in the classroom is conducted in a principled fashion, with data collected, interpreted, and reported in an objective and transparent manner, then meaningful insights may be gained into the process of acquiring a language in an ecologically valid context. This study takes the view that applying the methods of cognitive linguistic studies on metaphor learning to regular classroom situations may reveal insights into the acquisition process which in turn can be further examined in experimental settings. To be more specific, natural classroom settings such as the CLIL course in this study provide supportive linguistic and conceptual contexts as well as continuity of exposure to target language features that may not be easily replicated in experimental research designs. Thus this study can be seen as taking an action research approach, one that advocates a reflexive relationship between theory and practice, taking insights from each to feed back to the other.

2.9. Concluding thoughts

This chapter has reviewed the principle works and issues that have informed and guided this study. Conceptual Metaphor Theory has received criticism, but the essence of the theory serves as a straightforward model for introducing language learners to the metaphorical patterning that frequently appears in classroom input.

Metaphor studies have increasingly turned to corpora as sources of socially salient language use, and this has enabled more objective, replicable techniques for metaphor identification to be developed that will benefit the research community as they continue to be refined. Now that scholars have identified some of the distinguishing features of metaphor use in various genres, including that produced by language learners, we are in a better position to begin exploring this fascinating aspect of language and language learning. To do so, it will be necessary to consider the various components of metaphorical competence and plan studies that examine learner metaphor awareness from multiple perspectives. This in turn has huge potential to augment our understanding of learners' lexical growth by providing a principle from which to analyse the thorny issue of vocabulary depth. Finally, the study has been defined as an action research investigation that seeks to test out the principles of cognitive linguistics in an ecologically valid context and to identify issues that might be investigated further in expanded experimental studies. It has been argued that there is a need for investigations on metaphor awareness raising to be carried out in classroom contexts so as to fully appreciate the issues and challenges involved.

CHAPTER 3. THE RESEARCH CONTEXT AND AN OVERVIEW OF PROCEDURES

3.1. Introduction

As is widely recognised, the circumstances in which language production occurs may affect learner output. Any analysis should therefore take note of the context in which such production took place. This chapter provides a description of the teaching context, the physical surroundings, the participants themselves and the sources of data obtained in the study.

3.2. The context for the study

The study was carried out in a small university with a liberal arts curriculum in Japan. For the first three semesters of instruction, the university uses team-teaching to blend English language education with the content of other disciplines (i.e., a CLIL approach). In each class, there are two instructors, one a professor of the content field and the other a language instructor, who have shared responsibilities for developing language and content goals. This provides a form of sheltered immersion in which students are introduced to academic discourse norms and learn to complete progressively more challenging tasks. In the fourth semester, learners complete a study abroad program in an English-speaking country. For the final two years, students enrol in solo-taught English-medium content courses in which they are expected to fully participate in discussions, lectures and projects. They also carry out a research project that will culminate in a graduation thesis written in English.

The students who enrolled for the anthropology course on Japanese popular culture in the 2013 and 2014 academic years were involved in the study. This was a third-semester course in which learners were expected to already be familiar with the format and tone of academic essay writing and to be gradually working towards acquiring an understanding of high frequency and academic vocabulary. In the same semester, learners would have been enrolled in language program courses aimed at developing their knowledge of grammar and lexis as well as improving their oral and written communicative competence. Learners would also have been enrolled in team-taught courses on other liberal arts subjects.

In the Japanese Popular Culture course, the content instructor was a professor of anthropology from the United States of America. The language instructor, the author of this thesis, was an assistant professor from the United Kingdom. These two instructors had been in a team-teaching partnership for the same course for three years prior to the commencement of the study. The course taught in the 2013 academic year was designated as a control group, for which instruction followed the pattern that had been adopted in previous years, while the 2014 course acted as the experimental group.

The general goal of the Japanese Popular Culture course was to have students consider the origins and influences of popular culture in Japan and how these were related to Japan's interactions with the rest of the world. It sought to challenge stereotypical notions of culture as a discrete expression of national identity and instead to view it as a result of many years of transnational interaction, both influenced by and influencing other nations' cultures. Table 3.1 provides an overview of the topics and issues covered in the course which will be referred to in

Table 3.1. Materials and topics covered in the course

Title of lesson material	Topics / issues covered
Music in Japan (documentary)	Genres of music that are popular in Japan; Western influence on Japanese music; manufactured bands; the popularity of Japanese musicians overseas
Hatsune Miku: The World's Virtual Diva (promotional video)	Virtual pop stars; the use of technology to create human voices; non-musicians becoming able to produce music; the Internet as a device for sharing music
Fashion in Japan (documentary)	Fashion as a form of rebellion; fashion groups in Japan; the influence of Japanese fashion designers
Japanese Tribes (text)	Sub-cultures in Japan: Gothic Lolitas (fashion group), Otaku (computer and animation obsessives), Freeters (temporary workers), NEETs (people not in employment, education or training)
Hybridity and Hybridism (text)	Hybridity in this context refers to how cultural entities are adapted and influenced through interaction with other cultures. Hybridism refers to the belief that in spite of the cultural borrowing that has clearly taken place in modern Japan, the country has retained a unique essence.
Makiko's New World (documentary)	Life in Meiji era Japan; women's role in Meiji society; Meiji influences on modern Japan; Japan's interactions with the West
Home Alone (newspaper article)	Trends in living arrangements in modern Japan; the decline of traditional three-generation families; increases in solo lifestyles
Sumo Story (documentary)	Traditions of sumo; the pressures of westernisation on sumo; sumo in the modern world; challenges to remain relevant
The Japanese Version (documentary)	Japanese borrowing and adaptation of western culture; low culture; Japan's relationship with the west

subsequent chapters. Many of the language examples drawn from the output corpora will also make reference to the topics outlined here.

Class activities were often based around video materials and written texts as sources of content input. Students typically watched short sections of these videos

or read parts of the texts, and then worked together to deal with language difficulties or to complete comprehension and response questions. Towards the end of each semester, students began a small research project to investigate a chosen element of Japanese culture and to describe it with reference to the concepts they had learned in class. The culmination of this project was a short presentation students gave to the whole class in English without using notes.

3.3. Participants

All of the participants were informed that the purpose of the study was to gain insights into their language production while taking into account the instruction that they received. They gave informed consent for the study to use their written work as language output data, their scores on the Test of English for International Communication (TOEIC) as a measure of language proficiency, and for the investigator to obtain feedback on the sources of particular expressions they used in their writing. They also agreed for the classes to be recorded so that data could be obtained on the input they received. The informed consent form was bilingual (Japanese and English) in format, and it explained that participants could choose to withdraw from the study at any time without consequence. It also explained that personal information would be protected by the investigator. All participants' names were replaced with pseudonyms in the input corpora (see section 3.6.2, page 106) and learners are referred to with a numeric code in this thesis. The study received ethical approval from both the institution where it was carried out and the University of Birmingham.

There were 23 participants in the control group. 13 were male, and 10 were female. All 23 of the participants spoke Japanese as their first language, and one was also highly proficient in Korean. The participants were all second-year university students and were in their eighth year of formal English study. Their mean age at the commencement of the study was 19.17 years (SD = 0.49). In the experimental group, there were 23 participants. 12 were female, and 11 were male. All were native speakers of Japanese. Like the control group, they were all in their second year of university education and their eighth year of formal English study. Their average age at the beginning of the course was 19.09 years (SD = 0.29). A table listing participant data is provided in Appendix A (page 424).

3.4. Procedure for the study

This section will first discuss the instruction provided to the control and experimental groups and how this differed between the conditions, including the principles that guided vocabulary instruction and the treatment of metaphorical language. It will then detail the process by which metaphorical language was identified in learner writing, including deviations from the published MIPVU process and how linguistic metaphors were linked to broader conceptual metaphors. The section then explains how data obtained during the control condition was used to inform the development of materials to raise awareness of metaphor in the experimental condition. These materials were compiled into a metaphor workbook, which was the principle method by which metaphor was introduced in class. Finally, the section details the various forms of data gathered to provide evidence of learning outcomes.

3.4.1. Classroom instruction

As was mentioned in 3.2, instruction for the control group followed the pattern that had been established in the previous years' teaching. The course followed a CLIL approach that discussed issues related to anthropology through language development activities.

To address learners' vocabulary building needs, the transcripts from the video and text materials used in the course were analysed using the Vocabprofile function on the Lextutor website (www.lextutor.ca), and useful high frequency and academic vocabulary was selected for instruction. Regular institutional vocabulary testing had revealed that while third-semester learners typically had reasonable comprehension of the most frequent 1,000 words of English, many still had gaps in understanding of the second 1,000 words, and academic vocabulary was extremely limited. As such, words selected for explicit instruction were usually drawn from the second 1,000-word band of the General Service List (West, 1953) or the Academic Word List (Coxhead, 2000).

Often, the targeted words appeared in metaphorical senses, and the instructors would sometimes draw attention to how a particular figurative meaning had been derived from a more literal sense, but there was no consistent treatment of metaphor in class nor was there any attempt to highlight common metaphorical themes in the course input. The input corpus for the control condition revealed that the word *metaphor* or its related forms appeared only 18 times in the entire semester, 16 of which were in a single lesson activity examining a section of text that was particularly dense with metaphors.

In the experimental group condition, efforts were made to cover the course material at the same pace as the control group. While it was not possible to precisely control for time on task as in a purely experimental study, the two courses covered the same course content at largely the same pace, never falling more than one hour of teaching time out of sync with each other. Learners received the same written homework assignments at the same stage of each course in order to control as much as possible for developmental effects over time. Since the course in which the experimental condition took place had two lessons more than the control condition, the two courses matched each other for content coverage until that point, and the experimental group received different material for the additional lessons. No extra writing homework was assigned during this period. A table showing material covered in each lesson and writing topics assigned is provided in Appendix B (page 425).

The experimental group differed from the control group in that explicit metaphor awareness-raising activities were included throughout this condition. These replaced some of the regular activities that focussed on high frequency or academic vocabulary. As Appendix B shows, the metaphor activities appeared in around two-thirds of the lesson periods and included activities from a metaphor workbook which had been created based on an analysis of the language used in the control condition (see section 3.5, page 94 and supplementary material) as well as periodic review activities of previously studied language. Outside of explicit metaphor-focussed activities, learners' attention was also drawn to incidences of metaphorical language appearing in course materials by the instructors. The intention was to raise learners' awareness of the prevalence of metaphor in academic discourse so as to help them to realise that it was not merely a decorative

device. It was also intended that learners would come to appreciate the need to develop their vocabulary depth as well as breadth by learning to make use of the different senses of already known words. This increased emphasis on metaphor could be seen in the input corpus for the experimental condition, which revealed that *metaphor* and its inflected and derived forms appeared 384 times during the semester, and were dispersed widely - with only one class period not featuring some mention of metaphor.

Each week in both conditions, participants were asked to produce a piece of reflective writing in response to a given prompt that addressed a recent course theme. A record was kept of the specific instructions given for written homework assignments to the control group learners, and these were repeated in the experimental condition. Both groups of learners were given regular reminders that the purpose of the weekly writing homework was to experiment with new language and to express their thoughts regarding the course topics. These reminders were given as instructions to the whole class when the assignments were introduced and also when the learners' notebooks were returned to them. Although they were given some class time to review errors in their writing, learners were told that accuracy was not the focus of the exercise and that the intention was for them to engage with course concepts. The following examples illustrate how this message was reinforced to learners:

(10) <TEACHER 1[15]> So what I would like you to do in your notebooks, on one page, I want you to write in your notebooks. […] Try to use new language. In this

[15] Teacher 1 refers to the author of this study. Teacher 2 is the professor of anthropology.

course we will teach you lots of vocabulary and we will do some metaphor work, so I want you to be creative. I want you to push to use new language. I will not give you bad marks if you make mistakes. I will give you good marks if you try to use new language, try to be creative, all right? The topic is, what Japanese popular culture means to me, or to you. So what is your personal image, if you like, of Japanese popular culture. [EXP L1][16]

(11) *<TEACHER 1> So what we'd like to do is keep doing these notebooks every week, giving you topics to practice the writing. But what is important is that you learn to express yourselves using new language and the ideas that we talk about in the class, the content of the class, and also learn from your language mistakes and try and keep, [...] developing your writing skills.* [CON L3]

(12) *<TEACHER 1> In the future, Japanese weddings in the future, how do you think they will continue to evolve? Okay, and you learnt in Makiko's world, weddings were like this. We looked at modern, or fairly modern weddings. Think about the future, how will they evolve? What trends will emerge? What cultural influences will appear? Okay? So you're kind of speculating, you're kind of guessing a little bit here. [...] Use your own ideas. Is that okay?* [CON L23]

Learners submitted their writing in the following class period, and the texts were scanned and transcribed for data analysis. Each piece of written work was marked with a simple correction code that covered basic grammatical errors and aspects of vocabulary selection. While the feedback did address inaccuracies in learners' writing, the course instructors repeatedly emphasised that such corrections were

[16] This code refers to lesson 1 of the experimental condition.

only for their benefit and did not mean that their writing was of poor quality. The language instructor also added a simple comment that responded to the content of the learners' writing. These steps were intended to provide participants with supportive feedback, and to give them a sense of an audience for their writing. The correction code used in the study and a sample piece of writing with teacher feedback can be seen in Appendix C (page 431).

A second form of feedback was provided periodically to the whole class. This served as both a warm up activity and a way to review language from previous lessons. One such instance took place in the fifth lesson of the semester, after learners had studied language from the MORE OF AN ABSTRACT THING IS AN INCREASE IN SIZE OR HEIGHT metaphor theme. Two learners had produced the sentences shown in (13 - 14), in which verbs which presumably had been intended to activate metaphorical meanings had been used to describe physical subjects, both Japanese musicians. This had created unintended literal interpretations, which was felt to be an important point to bring up for the whole class.

(13) She was born in Japan and exploded in Japan. [EXP22][17]

(14) Japanese manufactured bands has boomed and progressed to Asian countries. [EXP7]

A simple worksheet was created in which learners were asked to see if they could find and correct errors based on those produced by some members of the class. There then followed a whole-class review containing the following extract, which illustrates how these particular errors were dealt with by the instructors:

[17] Examples of learner output have been given tags to link each with the participant data in Appendix A (page 424). [EXP22] refers to participant number 22 in the experimental group.

(15) *<TEACHER 2> ...don't feel bad that we are now correcting [your errors] [...] Cos you've got to try it to see what works and what doesn't work.*

<TEACHER 1> [...] We did, do you remember, things like rocketed and mushroomed and snowballed, and they mean, like, increases, yeah? Okay, the thing is, [the written example says], she was born in Japan and exploded in America. [...] That really means she really did explode.

<LAUGHTER>

<TEACHER 2> Like her body blew up into pieces. She doesn't live any more. Poor thing.

<TEACHER 1> So, because of that, when you use it as a metaphor, the subject must be something abstract like her career exploded, okay?

<TEACHER 2> Yeah, her popularity.

<TEACHER 1> [...] She was born in Japan and her career or her popularity exploded, no problem. That's okay, because careers are abstract things, but if you say she exploded, that's not abstract, that's a real thing, and it kind of means that she really did explode. Okay?

[...]

<TEACHER 1> One of the ways, things with metaphors is, we use the same words, but they often have, sort of different words that they work together with. So things like bombs, if a bomb explodes, it really does explode, but [if] abstract nouns explode in English, it means they become more popular. [...] Metaphors, [...], metaphors need abstract subjects. Popularity explodes or your career explodes, the demand explodes, okay, the number of people, the number of people who are studying overseas exploded. [...] And [...] the same

sort of thing for boomed. This band has boomed. Maybe this band's career has boomed. [EXP L5]

It was hoped that such feedback and simple activities that required learners to consider the operation of metaphor in context would encourage learners to experiment with language without feeling the pressure not to make mistakes.

After the transcriptions of participants' written output had been completed, they were compiled into two output corpora, one for each study condition. These transcriptions became the primary data source for addressing the research questions of this study (see section 3.6.3, page 118).

In order to provide information about the sources that participants were drawing on in their decisions to use metaphorical language, selected class members were asked to complete a feedback form when their writing was returned to them (see section 3.6.4, page 119). The forms were completed during class time, and were returned to the instructors within a few minutes.

3.4.2. Identification of linguistic metaphors

Once instruction in the control group course had been completed, the output corpus based on the writing of control condition participants and the lesson materials were examined. Before materials could be developed to help raise experimental group participants' awareness of the utility of metaphor in describing course topics, it was necessary to analyse the language that was already present in the input and to identify prominent themes that would be taught in the experimental condition. The MIPVU procedure was adopted for this purpose. MIPVU is the culmination of years of development work by several prominent metaphor researchers (Steen et al., 2010.

It has been shown to reliably identify metaphor at the level of individual lexical items across many genres of discourse. The following sections detail the application of MIPVU that was followed in this study, cases where the methodology differed from the original procedure, and issues regarding the application of MIPVU to output produced by language learners. An example of the application of MIPVU to a sample of participant writing will also be provided.

3.4.2.1. Initial application of MIPVU

For the initial analysis of course materials, the MIPVU procedure was carried out as described by Steen et al. (2010). During this process and the later linking of linguistic to conceptual metaphors, however, certain cases were noted which suggested that a strict application of MIPVU would not be optimal for pedagogic reasons. In many instances, these were lexical items which had undergone a form of metaphorical development at some point in their etymological history, but which would not be considered as metaphors under MIPVU, either due to their original senses having become obsolete or because grammatical shifts in usage meant that dictionaries now considered them as separate lexical entities in their own right (see section 3.4.2.2). While items such as *grassroots*, *deep-rooted*, and *forefront* do not meet the MIPVU criteria to be counted as metaphors, they are potentially useful devices for encouraging learners to consider meaning extension through metaphorical processes since they are composed of relatively frequent morphemes that are thematically related to other metaphors that appeared in course materials.

Another issue was that metaphor was not the only factor influencing semantic extension; many lexical items were also taking on extended meanings through

metonymic derivation. Although the MIPVU procedure does not address metonymy, it is often found operating alongside or within metaphor. Therefore, while the pedagogical focus of this study was on raising awareness of metaphor in order to promote vocabulary development, some consideration of how metonymy would be dealt with was necessary. These cases will be described in further detail in the sections that follow.

A final issue related to the treatment of multi-word units in the data. Analysis of the classroom materials and participants' writing revealed several chunks of language that were listed as phrases in dictionaries, but were not considered polywords under MIPVU. As will be explained in section 3.4.2.2, this risked distorting the counts of metaphors being used for text organizing purposes, and a variation on the established MIPVU procedure was therefore adopted.

3.4.2.2. Application of MIPVU in this study

Identifying metaphors with MIPVU requires following several steps that are designed to provide an objective, reliable, and valid decision-making process for researchers. In general terms, the steps are intended to ensure a clear understanding of the text as a whole, to establish the lexical units that comprise the text, to consider whether the contextual meaning of each lexical unit is different from, yet can be understood in comparison with a somehow more basic sense, and to decide on the appropriate coding for each unit. The steps, including deviations from the MIPVU process, will be outlined below.

i. Understanding texts

The first step in the MIPVU process is to read through the whole text to ensure comprehension. The classroom materials were read carefully during class preparation and teaching, as well as in carrying out metaphor identification. The learner corpora texts were also read several times: first during transcription, then in order to provide feedback to learners, and finally during the data analysis.

One issue with applying MIPVU to language learners' writing is that some stretches of language may not be comprehensible (see also Nacey (2013, p. 117 - 120) for a discussion on dealing with errors and issues of coherence). In such cases, a code of DFMA (discarded for metaphor analysis) was assigned to that section of text, and the lexical items within were not recorded in any of the frequency counts. In total, 720 lexical items were coded as DFMA in the output corpora from the two conditions, 0.71% of the entire data set.

ii. Establishing lexical units

The second step is to divide the text into lexical units. Most individual words are lexical units by themselves, but MIPVU treats phrasal verbs, polywords, and some compound nouns as single units of meaning. These multi-word lexical units are analysed as wholes, rather than by considering each word separately.

For phrasal verbs, the identification steps differed slightly from the original methodology since MIPVU bases its decisions on the BNC tagset system (C5), which has since been updated. Instead, the output corpora were tagged using both the Wmatrix tagging function and Qtag, a corpus tagging program developed by Oliver Mason. Both of these programs use the C7 tagset, which adds an RP tag to adverbial

particles. A list was then compiled of potential phrasal verbs, which were checked with the criteria provided in the MIPVU procedure[18] to distinguish them from verbs with bound prepositions. Once a list of correctly tagged phrasal verbs had been compiled, AntConc (Anthony, 2015) was used to check all instances of the 25 verbs that most commonly form phrasal compounds[19] in the output corpora, and instances of overlooked phrasal verbs were added to the list. A small number of phrasal verbs were added as they were discovered during the coding steps that follow.

The BNC list of multi-word expressions (http://www.natcorp.ox.ac.uk/docs/multiwd.htm) was used to identify polywords, as in the MIPVU procedure. However, there do not seem to be any specified criteria by which polywords were selected for inclusion on this list. This created an issue because some language chunks that appeared in participants' output and were described as phrases in dictionaries, such as *from now on*, appeared on the list, while others, such as *on the other hand*, did not. MIPVU treats polywords as single lexical units to be analysed as wholes. As a result, the individual words in *on the other hand* would be analysed separately for metaphor, but *from now on* would be analysed as a single unit. One of the goals of this study (described in chapter five) was to analyse learner use of metaphor for rhetorical purposes, including the use of phrases such as these to provide textual coherence, but analysing phrases in a differing fashion would have distorted the results. Therefore, in order to allow for both a comparison between the results of

[18] In phrasal verbs, adverbial particles can be placed before or after the noun object, and pronouns must be placed in front of the particle. For prepositional verbs, the prepositional phrase can be moved to either the front of the sentence or in front of a wh-word. It can also be modified with an adverb (Steen et al., 2010, p. 30).

[19] These were: bring, carry, come, cut, find, get, give, go, hang, hold, look, make, pick, point, put, run, set, shut, sit, sort, stand, take, turn, work, and write (Steen et al., 2010, p. 169).

this study and other studies of learner metaphor as well as an investigation of rhetorical metaphor use, a dual system of coding was adopted. In chapter 4, as well as sections 5.3 - 5.5 of chapter five, the MIPVU procedure for polywords was followed, with only phrases that appeared on the BNC polywords list treated as such. In section 5.6, which deals with rhetorical use of metaphor, and in chapters six and seven, language that was listed by the Macmillan or Longman dictionaries as a phrase was counted as a single lexical unit, but the component words were analysed for metaphor separately. This was because although MIPVU treats polywords as single units to be analysed in their entirety, there is evidence that language learners may process the component words of multi-word expressions individually, rather than as wholes (Cooper, 1999; Martinez & Murphy, 2011; Siyanova-Chanturia, Conklin & Schmitt, 2011).

For compound forms, the MIPVU procedure was followed for solid and hyphenated compounds (e.g., *outstretched*, *t-shirt*), with such words being recorded as a single lexical item if they could be found in the dictionary. In addition, hyphens are often used in forms that, while not found in dictionaries, conform to norms of word formation. These included the addition of affixes to modify nouns or create compound adjectives (e.g., *co-author*, *mid-1980s*, *western-style*) and present or past participles being used as modifiers (e.g., *cake-cutting*, *rubber-tired*). These were analysed as separate lexical units, as in MIPVU. For spaced compound nouns, Nacey's (2013, p. 91 - 92) approach was preferred to that of MIPVU. That is, the listing of a spaced compound noun as a dictionary entry was taken to be evidence of its existence as a single entity, rather than following stress patterns to determine single referents. As Nacey notes, the conventions for spelling compounds as spaced,

hyphenated or solid are far from universal, and reliance on stress patterns to determine whether an entity has a single referent can sometimes produce odd or inconsistent decisions. In the data coding for this study, it was noted that *fast food* had different stress patterns in online dictionaries[20], although it clearly has a single referent. Thus, it was decided that a dictionary listing would be taken as evidence that a compound referred to a single concept.

In the learner input data, there were also many instances of compound formulations that were unattested in dictionaries and did not conform to regular word-formation procedure such as those outlined above. Nacey (2013, p. 92 - 95), having encountered the same situation in applying MIPVU to learner English, developed categories and procedures to deal with this, and these were adopted for this study. The first broad category of erroneous compound formation is that which is based on a standard compound form, but either excluded a hyphen (e.g., *good looking*), fused the words into a single form (e.g., *hiphop*), added an unnecessary hyphen (e.g., *air-raid*), or split a single word into two (e.g., *battle field*). Such formations were analysed as single units, since that is how the apparently intended meaning appears in the dictionary. A second category comprised errors that created non-standard compounds through either fusion (e.g., *rawfish*) or unnecessary hyphenation (e.g., *music-sale*)[21]. These were analysed as separate units for metaphor, since despite the erroneous written form, there is no evidence for them

[20] In the Macmillan online dictionary (http://www.macmillandictionary.com), the stress appears on the first word, while in the Cambridge dictionary (http://dictionary.cambridge.org), it is clearly on the second word.

[21] Nacey's final category, non-standard compound formations including a hyphen and a space, was not included in this study. Since the data were handwritten, such errors could not be identified.

Table 3.2. Compound forms in the output corpora

Condition	Non-errors				Errors					
	Solid (in dictionary)	Hyphenated (in dictionary)	Hyphenated (not in dictionary)	Spaced (in dictionary)	Standard (no hyphen)	Standard (fusion)	Standard (unnecessary hyphen)	Standard (split)	Non-standard (fusion)	Non-standard (unnecessary hyphen)
Control	59	22	60	105	6	9	5	20	2	21
Experimental	80	26	59	173	11	11	14	22	3	21

existing as single concepts. Table 3.2 presents the total number of compound forms in each category for the two conditions.

iii. Considering meanings of lexical units

Contextualised meanings

The third stage of the MIPVU process involves a consideration of the meaning of each lexical unit. The first part of this stage is to determine the meaning of the lexical unit in that particular context. As in the MIPVU procedure, the Macmillan dictionary was adopted as the primary source for distinguishing senses, and the Longman dictionary was used as a back-up for any instances where this was not possible with the primary source. In the majority of cases, establishing the contextual meaning of each unit was unproblematic; however, there were some exceptions. Sections of incomprehensible text had already been assigned the DFMA code, but there were some instances, such as (16) below, in which the intended meaning appears to be clear, but lexical choice creates an unusual image.

(16) *Japan could made original Japanese culture because Japan was not contact to other countries, and also Japan is a quarantined continent for a long time.* [CON22]

The source dictionaries define *quarantine* only in relation to people or animals, and neither offers an adjectival definition. Interestingly, however, the Merriam-Webster dictionary of US English (https://www.merriam-webster.com) defines one sense of verbal *quarantine* as, 'to isolate from normal relations or communication.' COCA also contains several instances of adjectival *quarantined* being used to modify spatial nouns such as *area*, *zone*, *wilderness*, *site*, *sector*, and *region*, though in no cases was it clear that the agent who performed the quarantining was also the patient of such an action, as was the case with Japan's self-exclusion from foreign influence described in (16). A search of BNC concordance lines found no instances of *quarantined* appearing as an adjective. This suggests some variation in regional usage patterns and implies that future studies may do well to draw on dictionaries representing other varieties of English depending on the circumstances of the data being analysed.

Accordingly, it was felt that since the broader meaning had been preserved and there was evidence of some degree of conventionality, rather than discard such cases as (16), they should be analysed as potentially metaphorical. In this and similar cases where words were used in senses that were unattested in the main source dictionaries, the code 'WIDLII' (When In Doubt, Leave It In) was assigned. This is consistent with the approach to non-standard lexical choices taken by Nacey (2013, p. 96 - 97), and is in line with the MIPVU position that there is no way to truly know

the writer's intentions - lexical items are coded for their *potential* metaphorical status.

Basic meanings

For each contextualised lexical item, the next step was to consider whether a distinct, somehow more basic, sense also existed. As in MIPVU, basic senses were regarded as more concrete, specific, or human-oriented. Again, decisions were primarily based on entries in the Macmillan Dictionary, with a separate numbered entry from that of the contextual sense being regarded as evidence of sufficient distinctness from the more basic sense[22]. In cases that could not easily be resolved in Macmillan, the Longman Dictionary of Contemporary English was used as a secondary source. Inevitably, however, there were cases where literal or metaphorical readings were equally plausible, as in (17), which describes a social group with an unfortunate reputation for both unsavoury interests and an unkempt appearance. In these instances, the code WIDLII was applied, since the word was potentially metaphorical.

(17) In the past, Otaku was despised and almost people was saying "I dislike Otaku, it is dirty-looking." [EXP22]

In addition, a small number of lexical units that appear in the dictionary in their own individual section, rather than being a numbered entry under a particular headword, were included in the counts of metaphors. For example, the adjectival expression *deep-rooted* is listed separately in Macmillan from the adjective *rooted*

[22] Since the phrasal verb *fizzle out* (see chapter two, page 8) only has a single sense in the Macmillan Dictionary, there is no more basic sense in modern use, and it is not coded as metaphorical in MIPVU.

and both the verbal and nominal *root*. While both *root* definitions contain several numbered entries that are either literal or more abstract, the adjective *rooted* is defined as 'If one thing is rooted in another, it is based on it, has developed from it, or is influenced by it.' This seems to combine literal and metaphorical interpretations, and the entry for *deep-rooted* takes this further into purely abstract interpretations by linking its use only to feelings, beliefs, or ideas in a single entry. The Oxford English Dictionary confirms that this sequence of definitions tracks the etymological development of the term *root*, which appears to have begun at least as far back as the 12th century. MIPVU takes the position that while historically older senses are more likely to be concrete, specific, and human-oriented, the etymological origins of a term are unsuitable as a criterion for determining whether it is a metaphor since typical language users cannot reasonably be expected to have knowledge of a word's etymology. Thus, under MIPVU, only separate numbered entries in the same grammatical category can be taken as evidence of distinct senses being in contemporary use. Words that have been assigned their own individual listing or have shifted part of speech cannot be compared to the forms from which they were derived (Steen et al., 2010, p. 35 - 37). Words such as *fervent* and *ardent*, which no longer carry their temperature-related original senses in common parlance, have been cited as justification for this stance (Steen et al., 2010, p. 6 - 7).

From a pedagogic point of view, however, it could be argued that the metaphoric origins of some words (though perhaps not *fervent* or *ardent*) can easily be reactivated, and that this has the potential to ease the learning burden in explicit vocabulary acquisition (Schmitt & Zimmerman, 2002, p. 163). Words such as *deep-rooted*, *free-spirited*, *grassroots*, and *forefront* all contain morphemes that are likely

to be recognised by learners. Likewise, the verbs *snowball* and *mushroom* and the adjective phrase *absorbed in* are derived from other forms that, while not necessarily high-frequency, might readily be explained to learners and lend themselves to the creation of mental images, which are believed to assist vocabulary retention (Boers et al., 2008; Ellis & Beaton, 1993). Since other terms that drew on the same metaphorical themes were present in the course materials, it was desirable to include these words in classroom instruction and to consider them as exceptions to MIPVU. Finally, a group of words (*static*, *undamaged*, *enormous*, and *immense*[23]) were also taught as metaphors despite their single dictionary entries conflating physical and abstract senses. A complete list of the lexical units that were treated as metaphorical despite their not fully conforming to MIPVU is given in Appendix D (page 433).

Comparison of meanings

If a distinct, more basic sense was found for a lexical item, the relationship between the two senses was considered. Where a correspondence existed such that the contextualised item could be seen as a semantic extension of the basic sense, and one that required a shift in referential domain from that basic sense to the domain referred to in the context, then the lexical item was potentially metaphorical.

Frequently, such extensions were drawn from physical, concrete experience of the world and mapped onto abstract experience; for example, the word *bound*,

[23] Entries in the Oxford English Dictionary confirm *static* to have an original physical sense, *enormous* to have originally meant 'abnormal' (and hence to have been incorrectly considered metaphorical in this study), and *immense* to have meant 'unmeasured'. *Undamaged* is clearly a derivation of 'damaged', which could be considered metaphorical under MIPVU. In fact, only *static* appeared in any of the MIPVU-coded output data (Appendix D, page 433).

which has a basic meaning referring to being tied in order to limit movement, was used to refer to workers' freedoms being defined by the terms of their contracts. Semantic extensions also linked human (or at least animate) concepts to abstract fields, as in the use of *alive* to refer to a tradition[24]. In some cases, the basic sense of a word was a more specific, narrowly defined concept, and the metaphoric senses appeared to have been broadened to apply to a range of situations. The word *mixture* provides an example of this. Its most basic sense in the Macmillan dictionary is the second entry, "a substance such as food that is the result of mixing different things," whereas several metaphorical uses drew on the first entry, "a combination of two or more different things, people, qualities etc.," to describe cultural entities that combined various influences.

One consideration in judging the relationship between two meanings lies in whether a given semantic extension crosses two separate domains, as in metaphor, or remains within a single domain, as occurs with metonymy. As was noted in section 2.2 (page 17), metonymy is also a common feature of language, one that can often serve as a convenient linguistic shorthand form by allowing a prominent feature to represent a whole entity. Instances of metonymy were noted in the learner writing, such as examples (18), in which the word *western* is used to refer to people of European or North American Christian heritage, rather than a simple geographic direction, and (19), where the number of people in a soon-to-be marriage denotes the people themselves in *couple*.

[24] The coding for this mapping was based on the Longman dictionary, rather than Macmillan, which conflates literal and metaphoric senses.

(18) The reason why Japanese people prefer the style [Christian wedding ceremonies] might be admiration for the Western people. [EXP1]

(19) ...in Japanese [wedding] style, the couple wears a kimono. [CON21]

Cases in which both metaphor and metonymy appeared to be present in semantic extension were also noted. One example was forms of the verb *see*, as in (20). Seeing literally refers to our ability to perceive things visually, but since this is also one of the principle ways in which we take in information and understand our world, then it could be argued that there is some element of metonymy at work in uses that denote the creation of opinions. There was also one instance of metonymy being found within a phrase with metaphorical meaning. In (21), we can see an example of a PART FOR WHOLE metonymic relationship, with *headlines* standing for what were presumably whole articles about a rock group. The entire expression *hit the headlines* adds to this the impression that the group also became famous, with the word *hit* being used metaphorically to suggest some form of impact upon a music scene.

(20) ...the fear that Sumo was seen by foreign countries as a barbarous sport. [EXP1]

(21) In 1985, this group hit the headlines. [CON17]

While the goal of this study was to raise learners' awareness of the potential for semantic extension within a specific content area, it was decided to limit explicit instruction to the area of metaphor, rather than to include metonymy as well. This was partly due to the relatively low level of proficiency of most learners; providing instruction on different tropes carried the risk of an excessive learning burden and

the potential for confusion, especially since the differences between metaphor and metonymy can often be unclear. Similarly, instruction on metonymy would have required further class time in the experimental condition, increasing the challenge of matching course activities between the two groups. Accordingly, it was decided that when both metaphor and metonymy were identified in the comparison of meanings, then the lexical item would be coded as potentially metaphorical. Instances where only metonymy was present were not coded, and do not appear in the results. In some cases, this led to a distinction being drawn between different uses of the same word. In (22), *world* is used as a WHOLE FOR PART conceptual linking to describe societies in general. Since this usage retains the sense of referring to the entire globe, albeit those parts that pertain to human society rather than the physical planet, no crossing of domains has occurred, and the expression was not coded as metaphor. In (23), however, the global aspect of *world* is no longer present; instead, the focus is on a particular field of human endeavour. A comparison can be drawn between the conceptualisation of our planet as a closed-off system with its own laws and ways of being and the cultural norms of those involved in sumo and how they differ from other human activities. Since two separate domains can be identified in this relationship, this usage was coded as metaphorical.

(22) Japan has some famous musician in the world. [EXP18]

(23) ...recently the sumo world accepted foreign wrestlers because popularity of sumo is decreasing. [EXP23]

However, as Deignan (2005, p. 59) notes, and as examples (20 - 23) show, distinctions between metaphor and metonymy are not always clear, as one trope

can easily blur into another. This was particularly the case with the language selected for the metaphor workbook on the body as a source of metaphorical expressions. In fact, on later reflection, it was noted that some of the expressions included were metonyms rather than combinations of metaphor and metonymy (24 - 25). The analysis of classroom output and native speaker writing (sections 7.3 - 7.4, pages 305, 314) revealed several cases that covered the range from pure metonyms through a combination of metaphor and metonymy to pure metaphors. A list of the counts for these words and some examples have been provided in Appendix D (page 433).

(24) The meal cost ¥5,000 a head. [Metaphor workbook]

(25) She has a real ear for music. [Metaphor workbook]

iv. Coding decisions

In cases in which lexical units were found to be used in senses that were distinct from a more basic meaning, yet could be understood in comparison with that basic meaning through having similar features across domains, then the lexical items were coded as indirect metaphor-related words (MRWs).

Coding direct metaphors

Direct metaphors were identified following the MIPVU procedure. Lexical units which shifted the topic of reference away from that of the immediate context were identified. If these units could only be understood by way of nonliteral comparison with the context, and were perceived as offering comment on some aspect of the context, then they were marked as potential direct metaphors. Language devices

that served to flag the presence of direct metaphor, such as *like* or *as if* were also coded as MFlags. An example of this can be seen in (26), where there is a clear shift in reference from a human to a fanciful image, marked by *like* as a signalling device. In line with the MIPVU procedure, only content words, in this case, *dragonfly* and *heaven*, that were part of the phrase identified as a direct metaphor were coded as direct MRWs (Steen et al., 2010 p. 57).

(26) What I feel good about Makiko's life is that she was <u>like</u> <u>dragonfly</u> in <u>heaven</u> because I felt even if she was busy about housekeeping, she had free time for her loving things. [EXP9]

Coding implicit metaphor

Implicit metaphor, or the use of substitution or ellipsis to refer to metaphorical terms, was also coded following the MIPVU guidelines. In the vast majority of cases, this was achieved with pronouns used for anaphoric reference, as in (27), in which *them* was coded as an implicit MRW.

(27) ...it is good to get other culture's good <u>thing</u> and combine <u>them</u> to your own culture... [EXP8]

3.4.3. Example of MIPVU applied to data from the study

As the above steps might suggest, MIPVU is a detailed process that requires considerable attention to detail and extensive use of reference materials. In line with other studies that have adopted this or related procedures (Nacey, 2013, p. 81 - 85; Pragglejaz Group, 2007, p. 3 - 13; Steen et al., 2010), an analysis of a short sample of

data from this study will now be provided. The sentence in (28) opens a piece of writing discussing sumo in the past, present, and future.

(28) Over many years, sumo has faced many obstacles along its history. [CON8]

In the subsequent text, the writer explains how sumo resisted pressure to shed what Meiji era reformers regarded as its barbaric image, how it now faces increased competition from imported sports such as basketball and football, and how in the future it will need to regain the youthful audience it once had. In the second step of MIPVU, each of the words was recorded as a separate lexical unit, so the following analysis concentrates on the remaining steps in the procedure. The contextual meaning of each lexical unit is considered, and whether a more basic sense exists is then determined. If so, these senses are compared and a decision is reached on whether to code the item as MRW or non-MRW.

OVER (preposition)

Contextual meaning: This use of *over* matches sense 15 in the Macmillan Dictionary: 'during.'

Basic meaning: The most basic sense is number one: 'above someone/something.'

Comparison of meanings: Lindstromberg (2010, p. 111 - 112; p. 129) explains that the basic spatial sense of *over* is commonly used for paths of movement that pass from one side of a landmark to another (29), and that this image can be extended to view time as a distance over which some entity may pass (30). By viewing time as a space, a comparison can thus be drawn between the basic and contextual senses.

(29) *....a cloud of enormous fruit bats rose up en masse from a nearby bird sanctuary and soared over our heads...* [COCA]

(30) *[New York] is the biggest city by far in the United States, and over the last 20 years, it had two very strong mayors with Giuliani and Bloomberg.* [COCA]

Coding decision: Metaphor-related word (MRW).

MANY (determiner)

Contextual meaning: The meaning provided in Macmillan is 'a large number of people, things, places etc.'

Basic meaning: Despite having numerous grammatical functions and appearing in several phrases, *many* only has a single sense.

Comparison of meanings: Both the contextual and basic senses are the same.

Coding decision: Non-metaphor-related word (non-MRW).

YEARS (noun)

Contextual meaning: The meaning employed here appears to be the first sense listed in Macmillan: 'a period of 365 days, or 366 in a leap year, divided into 12 months.'

Basic meaning: The most basic sense of year(s) could be either the first sense or sense three: 'the amount of time that a planet takes to travel round the sun.'

Comparison of meanings: Even if the more specific third sense were identified as the most basic, there does not appear to be the crossing of domains that would suggest a metaphorical mapping. Instead, the contextual sense would be more like a generalization of the basic meaning.

Coding decision: Non-MRW.

SUMO (noun)

Contextual meaning: The contextual meaning is, ‘a Japanese sport in which two very large men wrestle.’

Basic meaning: There is only one sense for *sumo*.

Comparison of meanings: Both the contextual and the basic senses are the same.

Coding decision: Non-MRW.

HAS (auxiliary verb)

Contextual meaning: As an auxiliary verb, *has* denotes aspectual rather than semantic content. Macmillan lists one use of the ‘have’ verb as ‘used for forming the perfect tenses of verbs’ (sense 1a).

Basic meaning: There is no more basic sense of *have* as an auxiliary verb.

Comparison of meanings: Both the contextual and basic senses are the same.

Coding decision: Non-MRW.

FACED (verb)

Contextual meaning: The closest sense listed by Macmillan appears to be sense two: ‘if you face a problem, or if it faces you, it is likely or certain to happen and you have to deal with it.’

Basic meaning: The basic meaning listed in Macmillan is the first: ‘to be opposite someone or something so that your face or front is towards them.’

Comparison of meanings: The contextual meaning is distinct from the basic sense, and can be understood in comparison with it. Physically facing something in the basic sense has been extended to confronting a more abstract situation.

Coding decision: MRW.

MANY (determiner)

As above.

Coding decision: Non-MRW.

OBSTACLES (noun)

Contextual meaning: The meaning in this context is sense one in Macmillan: 'a difficulty or problem that prevents you from achieving something.'

Basic meaning: Sense two in Macmillan is, 'an object that you must remove or go around in order to move forward.'

Comparison of meanings: The basic sense is clearly more concrete than the contextual meaning, which can be considered similar by comparison. There is a crossing of domains, from a physical object impeding movement to a more abstract problem preventing something being achieved.

Coding decision: MRW.

ALONG (preposition)

Contextual meaning: This particular word is interesting for coding as it appears to be a borderline conventional/novel usage. The Macmillan dictionary does not record any senses of *along* as a preposition that refer to temporal concepts. However, the Longman dictionary includes a sense for the phrase *along the way/line* that is defined as, 'during a process or experience, or during someone's life' (sense 4). Furthermore, an inspection of COCA concordance lines reveals very occasional use of *history* as an object of the preposition *along*, as in (31).

(31) Yet, how much more striking is the discordance between the factual beliefs of men, their religions, their para- or non-religious outlooks, not to speak of their

dominant individual and collective interests, than between their moral beliefs all over the world and along its history! [COCA]

Basic meaning: The basic meaning of *along* is sense one in Macmillan: 'moving on or beside a line.'

Comparison of meanings: According to Lindstromberg (2010, p. 81 - 86), the basic sense of *along* can be described as elongated movement or presence related to two entities in a parallel plane. There is also a connotation of progress without hindrance that can extend to metaphorical uses that express togetherness (e.g., *along with*, and phrasal verbs such as *come along* and *get along with*), but Lindstromberg makes no mention of temporal uses. While there is some doubt as to how conventional the use of *along* + temporal expression may be, the contextual meaning employed here is clearly distinct from the basic meaning and, if we consider it as a continuation of the TIME IS SPACE metaphor, it can be understood through comparison.

Coding decision: MRW.

ITS (pronoun)

Contextual meaning: Sense one in Macmillan is: 'belonging or relating to a thing, idea, place, animal etc when it has already been mentioned or when it is obvious which one you are referring to.'

Basic meaning: There is no more basic meaning.

Comparison of meanings: The contextual and basic meanings are the same.

Coding decision: Non-MRW.

HISTORY (noun)

Contextual meaning: The Macmillan Dictionary lists several senses for history, but the closest to that being used in this context is number four: 'the length of time that something has existed.'

Basic meaning: The most basic sense is number one: 'the whole of time before the present, and all things that happened in that time.'

Comparison of meanings: There is some distinction in the two meanings here, but instead of there being any crossing of domains, it appears that the contextual meaning is a simply a narrowing of the focus of the basic meaning.

Coding decision: Non-MRW.

3.4.4. Establishing reliability

Once the data had been coded, a second rater, also a researcher of metaphor trained in the use of the MIPVU procedure, analysed 6,956 lexical units of the corpus (6.60% of the total) so that inter-rater reliability could be checked. After an initial round of coding, agreement between the raters was 95.16% of decisions for indirect, direct and implicit metaphor. The Cohen's kappa measure adjusts this value to account for agreement that would have occurred by chance, in this case producing a value of 0.75 agreement. This shows a good level of agreement, but cases of disagreement were discussed in order to improve the accuracy of the data coding. In order to reduce the number of different cases of disagreement that had to be discussed at once, the initial focus of discussion was on prepositions, determiners, and pronouns, since these three parts of speech accounted for over 50% of all disagreement. In fact, the 169 instances of disagreement came from just 26 unique word forms. Once the issues with these parts of speech had been resolved,

agreement was 97.57%, and kappa had risen to 0.88. Cases of disagreement for open-class word forms were then examined. As might be expected, there was a greater variety of open-class forms, with 92 word types coded differently between the raters, but a notable proportion (35.71%) of these were the extremely high frequency forms *end*, *feeling*, *great*, *have*, *kind*, *thing*, and *way*. Most of these cases could be resolved after further dictionary consultation, and the final agreement reached was 99.93%, with a kappa value of 0.997. The results of these coding decisions were then re-applied to the whole data set in order to maintain consistency and improve accuracy.

3.4.5. Linking linguistic to conceptual metaphors

MIPVU is a procedure that identifies linguistic metaphors, but in order to help learners appreciate the lexical patterning in prominent course themes, it was necessary to link metaphor at the linguistic level to that of broader concepts. To achieve this, Steen's (1999) five-step approach was adopted. Steen developed each step in a logical progression that would explain the assumptions in connecting metaphoric language with presumed metaphoric concepts. These steps will be outlined with the use of *absorbed* in example (32), which, it will be shown, can be linked to the CULTURES ARE SUBSTANCES conceptual metaphor.

(32) ...Japanese society has absorbed other countries' culture... [CON11]

i. Metaphor focus identification

The first step that Steen outlines is to identify the elements in an expression that combine to create the metaphor. The first of these is the *focus*, which is simply the

linguistic unit, or to remain consistent with MIPVU terms, the lexical unit, that is used in a non-literal sense in the expression. In this case, MIPVU has already identified *absorbed* as a linguistic metaphor. The complementary element, which in some cases may not actually be present, is the *frame*, which is defined as the proximate linguistic environment of the focus. Here, the noun phrase *other countries' culture*[*s*] was selected as the frame, since culture was a prominent course topic and one that it was hoped learners would engage more widely with. Taken together, these two elements allow the identification of a complete metaphor since they cannot be related to each other in a literal fashion.

ii. Metaphor idea identification

Having identified the key elements of a metaphor, it is then necessary to fully state the propositions that form the expression under analysis so as to clarify the concepts and references it conveys. Bovair and Kieras (1981) produced a manual to guide the identification of semantic content in text. Their method requires listing each proposition in an expression as a predicate and one or more arguments. In the case of (32), this would be performed as follows:

P1. (ABSORB SOCIETY CULTURE)

P2. (MOD SOCIETY JAPANESE)

P3. (POSSESS COUNTRY CULTURE)

P1 identifies the verb *absorb* as having two arguments, the subject *society* and the object *culture*. P2 applies a modifying function to *society* to give it the attribute

of being Japanese. P3 then uses a POSSESS predicate to indicate that the culture is thought of as belonging to those countries.

The propositions above deconstruct (32) into a minimalist listing of concepts and their relationships, thus providing the preliminary link between linguistic form and conceptual meaning. The metaphorical idea can be found in P1, in which we see a nonliteral relationship between the predicate and arguments.

iii. Nonliteral comparison identification

The metaphorical idea that was identified in step ii now must be fully explicated by examining the cross-domain mapping that exists between the literally and non-literally used components. To achieve this, Steen (1999) drew on Miller's (1993) work using comparison statements to construct plausible interpretations of how metaphors are understood. Miller presents comparison statements to investigate various linguistic structures, but all are variants of the basic form SIM[F(x), G(y)], in which it is claimed that there is a relationship of similarity between a concept F with its argument x and another concept G with its argument y. With regard to proposition P1, the comparison statement would be written as:

ABSORB (society, culture) → (∃F) (∃y,y') {SIM[F(society, culture), ABSORB (y, y')]}

This statement posits that there is some activity F performed by a society on a culture which is similar in nature to the absorbing that proposed entity y does to proposed entity y'. This is the framework of a nonliteral comparison.

iv. Nonliteral analogy identification

Once the framework of a comparison statement has been constructed, the missing elements F, y, and y' must be inferred. As Steen (1999, p. 68 - 69) and Miller (1993, p. 384 - 385) note, no single combination of words can be claimed to definitively complete such gaps in the comparison statement. However, corpus data and codification in dictionaries at least provide some evidence of regular usage and conceptual reference. In the case of *absorb*, the literal sense as listed in the Macmillan dictionary is 'to take in a gas, liquid, or other substance.' COCA data suggest that the most common noun to appear as the object of *absorb* in its literal sense is *water*, and in this case, the subject position is typically occupied by physical entities such as *soil(s)*, *roots*, *fiber*, or *beans*. The most frequent literal verb forms that take *culture* as an object and reflect the same relationship as one substance absorbing another are *change* and *influence*. It might therefore be claimed that a plausible completion of the comparison statement constructed in step iii would be:

ABSORB (society, culture) → {SIM[CHANGE (society, culture), ABSORB (substance1, substance2)]}

Or more prototypically:

ABSORB (society, culture) → {SIM[CHANGE (society, culture), ABSORB (soil, water)]}

v. Nonliteral mapping identification

The final step of the process to link linguistic forms to conceptual metaphors is to flesh out the analogy that is being drawn by detailing the correspondences that are

implied between the various entities and processes in the two domains. In this case, the set of correspondences might be projected as follows:

- The way that a society can be changed by another culture corresponds to the way that soil absorbs water.
- The society corresponds to the soil.
- The cultural influence corresponds to the water.
- The way that an affected society can retain its original features alongside the external influences corresponds to the way the soil retains its physical composition yet contains another substance within itself.

This process was used to link linguistic metaphors identified by MIPVU from the control condition data to the conceptual metaphors that were used for instruction in the experimental condition and in the later analysis of the experiment group output. Metaphors representing common themes that appeared in the course were collected and used as the basis for a set of materials for raising learners' awareness of conceptual metaphor. In some ways, this is a challenging process to apply; the ability to use comparison statements and propositions to reduce statements to their semantic foundations is not one that is readily acquired, and repeated rereading of examples is often necessary. That being said, there is a clear logic to the process, and in step iv, where perhaps subjectivity is most likely to creep in, it was possible to make use of corpus data and dictionary entries to support the identification process.

One issue that did arise with regard to the linking of linguistic and conceptual metaphors was that during the examination of learner output, it was noted that several metaphorically-used words were drawn from the source domains of plants

and the human body, although their uses were too varied to be linked under a single conceptual metaphor. High-frequency adjectives were also commonly employed in metaphorical senses, again without falling under the umbrella of a particular conceptual metaphor. Accordingly, it was decided that, rather than limit instruction solely to conceptual metaphors, these other groupings would also be included in the awareness-raising activities of the experimental condition. This is consistent with the approach taken by published materials for the teaching of metaphor (Lazar, 2003). Additionally, since instruction would encompass groupings other than conceptual metaphors, the broader term *metaphorical themes* would be used in teaching materials. This term will be used in describing the groupings of metaphors used in this study, while *conceptual metaphor* will occasionally be used with reference to CMT or to Steen's procedure for linking linguistic and conceptual metaphors.

3.5. Creating the metaphor workbook

It was decided to compile instructional materials into a single workbook for two reasons. Firstly, the majority of metaphorical themes in the control condition occurred at intervals across the whole semester, rather than being concentrated into one particular unit. It was felt that having a booklet would enable the instructors to refer to metaphor as and when it appeared in class. Secondly, presenting the materials in a workbook provided a sense of unity, and it was hoped that learners would appreciate its importance if they received something substantial.

The metaphor workbook (see supplemental material) was used in the experimental condition as the primary method of providing explicit metaphor instruction. Language that appeared in lesson materials for the control condition was

used as the basis for identifying metaphorical themes to be included in the workbook. The concepts and issues raised in each set of course materials usually meant that particular metaphorical themes could be highlighted as they appeared in the course, although some general themes appeared in several of the materials. Once a theme had been noted, dictionaries and corpora were used to identify other examples of language that were part of the same theme. Analysis of the control group's output corpus also helped to identify language that might already be known by learners and provided some examples of common errors that could be addressed through awareness raising of appropriate metaphors.

This can be seen in language used to express the concept of culture itself, which appeared frequently in the course. The examples below suggest that some individuals were aware of language comparing cultures to substances (33) or bounded spaces (34), and that individual cultural entities could be seen as physical parts of the broader culture (35). At the same time, however, examples (36 - 39) are some of the many cases that suggested several learners were failing to distinguish between cultures in their entirety and individual cultural elements.

(33) Therefore, Japanese society has absorbed other countries' culture and imported products... [CON11]

(34) So, taking in good points of other countries' culture is good. [CON23]

(35) Sumo is one of the importance elements for Japanese culture. [CON23]

(36) ....there are a lot of new culture at that time [Meiji era]. [CON6]

(37) My impression of Japanese popular cultures is "wa" and "Japanese anime." [CON5]

(38) Japan has many popular culture... [CON2]

(39) For example, cherry blossom is one of the Japanese culture. [CON23]

By examining both the metaphors used to express key concepts in course materials and learners' own usage patterns, target language was selected that would be appropriate for learners' needs. In some cases, target forms were selected to fill apparent gaps in learners' productive capacity. In others, it was noted that while learners were already able to use particular metaphors, there was a tendency to rely on the same word to express a notion. For example, the various inflected forms of either *evolve*, *hybrid*, or *born* accounted for 86% of the cases where the CULTURES ARE LIVING THINGS metaphor was used (lexical variety will be discussed in detail in 4.5 and 5.5). Therefore, target forms were sometimes selected to draw attention to metaphors that learners either avoided or were unaware of, and on other occasions to encourage variety in lexical choice. The chosen metaphor themes and some examples from the course materials are provided in Table 3.3.

As can be seen in the table, metaphor themes were presented either as conceptual groupings linking source and target domains, groups of metaphors related to a particular source domain, or metaphorical senses of a particular part of speech. Grouping metaphors by source domain (e.g., plants or body parts) is a convenient arrangement as it creates a natural question in that learners can be asked to consider how meanings can be extended from their basic senses. This approach is sometimes taken in published materials for teaching metaphor (Lazar, 2003, p. 8 - 9). However, not all source domains are as fertile for metaphorical

language, and in content-based programs, it may often be more suitable to arrange metaphors by target domains that frequently appear in course content.

In CLIL environments, it is also likely that metaphors reflecting prominent themes will appear throughout the course, rather than in isolated clusters. For this reason, learners may need to refer back to previously studied material, which makes the use of a workbook rather than individual worksheets more practical.

For the workbook activities, a variety of approaches were adopted in order to accommodate different learning styles and to encourage deeper processing. Most metaphor themes were introduced through explanation or elicitation, since this allows information to be presented quickly and instructors to adapt input to suit learners of differing ability. As can be seen in Table 3.4, the teaching for most metaphors was then reinforced through other methods. Application was the most commonly used approach, since this was seen as a useful intermediary step before learners attempted to produce metaphors in their weekly writing tasks. Other common approaches were visualisation and deduction. These learning experiences are relatively easy to bring into classroom activities, and each caters to a different learning style, holistic and analytic respectively. Activities involving analysis or evaluation, in which learners analyse the effectiveness of metaphors in text, were not included in the workbook. This was partly due to the ability range of the participants, and also because the course materials did not lend themselves to such analysis.

Table 3.3. Metaphorical themes selected for the workbook used in the experimental condition

Metaphor themes	Examples
Introduction to metaphor	
PRECIOUS MATERIALS ARE VALUABLE, WONDERFUL OR BEAUTIFUL ENTITIES	*Kyoto is a jewel of a city.*
Sources of metaphor: plants	*The popularity of Japanese anime stems from its interesting storylines and beautiful artwork.*
Sources of metaphor: movement	
THE FUTURE IS IN FRONT, THE PAST IS BEHIND*	*looking back, going forward*
MOVING FORWARDS IS IMPROVEMENT	*For years, youth fashion in Japan was held back because of strict rules about school uniforms.*
MOVEMENT IS FREEDOM	*I enjoyed my vacation in Okinawa. It was a good escape for me.*
Sources of metaphor: the body	*The university stands in the heart of the city.*
Metaphorical adjectives	*There is a big difference between living in the city and living in the country.* *They provide high quality service.* *He has a dark side to his character.*
Metaphors for cultures	
CULTURES AND CULTURAL ENTITIES ARE LIVING THINGS	*Hip-hop is one of the fastest growing areas in both Japanese music and fashion.*
CULTURES ARE SPACES	*Western culture entered Japanese culture during the Meiji period.*
CULTURES ARE SUBSTANCES	*Japanese culture has absorbed many things from overseas.*
CULTURES ARE CONSTRUCTIONS	*Japan has constructed a culture that preserves a strong national identity.*
CULTURES HAVE MANY PARTS	*Purikura is a part of Japanese youth culture.*
CULTURES ARE POSSESSIONS	*Many Asian countries share the custom of using chopsticks to eat.*
Metaphors for describing increases	
MORE OF AN ABSTRACT THING IS AN INCREASE IN SIZE OR HEIGHT	*Since its first release in 2004, the popularity of vocaloid music has soared.*
Metaphors for describing success	
ACHIEVING SUCCESS IS BREAKING SOMETHING*	*Japanese fashion is breaking down barriers.*
MORE IS UP AND LESS IS DOWN	*The number of Japanese people who study overseas has dropped recently.*
Metaphors for passions, interests and emotion	
STRONG EMOTIONS ARE SICKNESS OR INSANITY	*Evelyn was crazy about her grandchildren.*
BEING INTERESTED IS ENTERING SOMETHING	*How did you get into vocaloid music?*
STRONG EMOTION IS PHYSICAL FORCE	*He was struck by sudden shame.*
Metaphors for time	
TIME IS MONEY*	*I spent five hours cleaning my apartment last weekend.*
TIME IS MOVEMENT OR SPACE*	*The summer vacation is approaching.*
ABSTRACT ENTITIES ARE PHYSICAL STRUCTURES	*Why has people's confidence in politicians become eroded recently?*
ABSTRACT FORCES ARE PHYSICAL BURDENS*	*The captain of a national sports team carries great responsibility.*

Note: Metaphor themes marked with an asterisk were included in the metaphor workbook, but could not be taught due to time constraints.

Table 3.4. Pedagogical approaches for metaphor themes in workbook

Topic	Metaphor themes	Pedagogical approach								
		Explanation / elicitation	Visualisation	Classification	Comparison	Investigation	Application	Deduction	Analysis / evaluation	Attention to form
Introduction to metaphor	PRECIOUS MATERIALS ARE VALUABLE, WONDERFUL OR BEAUTIFUL ENTITIES	✓			✓	✓	✓			✓
Sources: plants	/		✓					✓		
Sources: movement	THE FUTURE IS IN FRONT, THE PAST IS BEHIND			✓						
	MOVING FORWARDS IS IMPROVEMENT		✓				✓			
	MOVEMENT IS FREEDOM							✓		
Sources: the body	/		✓					✓		
Metaphorical adjectives	SIZE IS AMOUNT		✓					✓		
	SIZE IS IMPORTANCE		✓					✓		
Metaphors for cultures	CULTURES AND CULTURAL ENTITIES ARE LIVING THINGS	✓								
	CULTURES ARE SPACES	✓					✓			
	CULTURES ARE SUBSTANCES	✓					✓			
	CULTURES ARE CONSTRUCTIONS	✓					✓			
	CULTURES HAVE MANY PARTS	✓					✓			
	CULTURES ARE POSSESSIONS	✓					✓			
Metaphors for describing increases	MORE OF AN ABSTRACT THING IS AN INCREASE IN SIZE OR HEIGHT	✓	✓							✓
Metaphors for describing success	ACHIEVING SUCCESS IS BREAKING SOMETHING	✓				✓				
	MORE IS UP AND LESS IS DOWN	✓		✓			✓	✓		✓
Metaphors for passions, interests and emotion	STRONG EMOTIONS ARE SICKNESS OR INSANITY							✓		
	BEING INTERESTED IS ENTERING SOMETHING							✓		
	STRONG EMOTION IS PHYSICAL FORCE							✓		
Metaphors for time	TIME IS MONEY	✓					✓			
	TIME IS MOVEMENT OR SPACE		✓				✓			
	ABSTRACT ENTITIES ARE PHYSICAL STRUCTURES						✓			
	ABSTRACT FORCES ARE PHYSICAL BURDENS									✓

3.6. Sources of data

In order to present as detailed a picture as possible of the learners themselves, classroom exposure, learners' written output, and the influences for their choice of metaphorical expressions, several forms of data were collected during the study. The following sections detail the collection procedures, limitations of the data sources, and justifications for their use.

3.6.1. Language proficiency

Participants' scores on the TOEIC test were used as a measure of general language proficiency. The TOEIC test is a standardized instrument developed over 30 years ago by the Educational Testing Service (ETS), the body also responsible for the widely-used TOEFL test (ETS, 2016). The test was originally created for a Japanese audience, and is now used in 150 different countries (McCrostie, 2010, p. 2; Murai, 2016). It is widely administered in Japanese universities, with 50% of institutions using TOEIC as part of their entrance examinations and 49% offering university credit based on TOEIC scores (TOEIC, 2016). Such is the case in the institution in which this study was conducted; students are able to take the test at least once every semester, and scores are used as a benchmark to determine whether they are able to take more advanced courses. As such, a high degree of familiarity with the instrument by the participants of this study can be assumed.

While TOEIC has been widely used in the Japanese context for some considerable time, it does receive criticism. One concern raised is that the test lacks the ability to comment on examinees' productive capabilities since the most commonly used form contains only listening and reading questions (Hirai, 2002). A

study carried out by Woodford (1982) on 306 examinees found a relatively high correlation of 0.83 between TOEIC reading scores and a separate test of writing skills, but Hirai's (2002) study found a more moderate correlation of 0.66 between TOEIC reading scores and a separate test of business writing skills. While a Fisher r to z transformation reveals that there is a significant difference in the strengths of these relationships (r_a= 0.83, n_a=306; r_b=0.66, n_b=102; z = 3.41; *p* (two-tailed) = 0.0006), both values indicate moderate to strong relationships between TOEIC scores and writing.

A further concern relates to the purposes for which TOEIC results are sometimes used. As a norm-referenced test, the TOEIC is intended to provide an estimate of an examinee's proficiency set against that of a population of other test takers. However, the test is often used for relatively high stakes purposes in Japan, such as for job hiring and promotions or academic credit. This situation can lead to negative washback that promotes teaching test strategies or content at the expense of general language development or the use of the test as a measure of short-term gains, neither of which facilitate reliable measurement (Chapman, 2006, p. 1; Childs, 1995, p. 73). As with all multiple-choice test formats, we cannot be certain how much the scores of participants in this study were influenced by test strategies such as eliminating answer choices and simply 'studying for the test', rather than by general language development. The learning context may have helped to mitigate against this, however, since the curriculum itself is focused on liberal arts education rather than test-taking strategies, and the multiple testing opportunities lessen the significance of each single test outcome.

A final consideration relates to the variety of English that TOEIC purports to assess. While it is sometimes thought of as a test of business English, the test is described by its developers as a measure of "assessing English proficiency for business," that can "enable universities to better prepare students for the international workplace," (ETS, 2016) rather than as a test of business English per se. Indeed, it has been claimed that the test is more a measure of grammar and vocabulary knowledge than of the language required in a business context (Chapman, 2006, p. 6 - 7).

As the above examples show, the TOEIC is not without its flaws; however, alternatives to this instrument would raise further issues. One important consideration before administering a test is the degree to which the difficulty of the instrument matches the ability of the persons taking it. This is because the reliability of a test can be affected by the range of knowledge possessed by its intended audience; examinees who lack the required knowledge level for which the test is intended may still answer multiple-choice questions correctly simply by luck, and this degrades measurement (Burton, 2001, p. 43). Since most standardized tests, such as TOEFL, IELTS and the Cambridge Exams, as well as TOEIC, make some use of multiple-choice or similar item formats, this is an important point to consider. Although these standardized tests are claimed to cover an approximately similar range of difficulty (ETS, 2017; Tannenbaum & Wylie, 2008, p. 29), independent research has suggested the vocabulary demands posed by TOEIC are somewhat lower than that of English for Academic Purposes tests such as TOEFL (Chujo & Nishigaki, 2003, p. 79; Chujo & Oghigian, 2009).

Practical constraints were a further issue to consider. Instruments other than TOEIC would have had to be administered either during class time, which would have disrupted the naturalistic nature of the study, or at a time outside of class hours, posing logistic problems as well as increasing the demands on students' study time. It has been argued that learner motivation can suffer when tests or other activities are imposed for external reasons that lack relevance to personal goals (Dörnyei, 1994, p. 277; Ryan & Deci, 2000, p. 73), and this may have been the case here since standardized tests require considerable time to complete. In the case of TOEIC test scores, a clear motivation existed for participants to make maximum effort as there are direct benefits both in terms of academic outcomes and future career prospects. As such, it was felt that these scores represented participants' best attempts at demonstrating their ability.

Perhaps the most important reason for selecting TOEIC over other instruments was because of test familiarity. As Bachman (1990, p. 111 - 159; 223 - 226) notes, features of testing instruments, such as instructions, timing, and item format, have a large impact on examinee performance. Variations in these features are changes in the conditions under which examinees are assessed and thus are a source of random error in measurement. Because TOEIC was the only instrument that the participants had had consistent experience of, it was felt to be the test that would produce results least susceptible to this form of measurement error.

In order to aid comparison with other studies of learner production of metaphor, participants' TOEIC scores were converted into bands on the Council of Europe Framework of Reference for Languages (CEFR) using the scale produced by Tannenbaum and Wylie (2006, 2008). The scale was produced following a rating

procedure in which 22 panellists, experts in language instruction, learning, testing, and the CEFR, analysed TOEIC test items and individually assigned probabilities of learners with the ability level defined at each CEFR band answering each item correctly. The sum probabilities provided by individual panellists in each section of the test were then collated to allow for comparison and variability in judgments to be revealed. The discussions that followed were informed by the sharing of rationales and actual performance data on the TOEIC items drawn from over 100,000 test takers of varying ability levels (Tannenbaum & Wylie, 2008, p. 8 - 12). By following iterations of this process, variations in panellists' judgement were reduced and, for each section of the TOEIC test, cut scores were determined for learners with just enough ability to be placed at each level of the CEFR. Ultimately, the shared view of the panellists was that the TOEIC cut scores were suitably placed onto the CEFR scale in the A2 to B2 range, although there was some concern about the ability of the instrument to measure performance at the extremes (A1 and C1) of the CEFR scale (Tannenbaum & Wylie, 2008, p. 27).

Based on these cut scores, the CEFR levels of the participants in this study could be determined. The distribution of participants across CEFR levels is provided in Table 3.5, and the level of each individual participant is given in Appendix A (page 424). As can be seen in the table, the majority of participants were in the A2 band, with TOEIC scores less than 550. Based on the summaries in Tannenbaum and Wylie (2008, p. 46 - 54), learners at this level are capable of writing about familiar topics using relatively basic grammar and high-frequency vocabulary. They can also express themselves orally in short exchanges and are intelligible with some effort on the listener's part. For receptive skills, they are able to comprehend short texts or

recordings on topics with which they are familiar and to extract important information or make basic inferences. This limited range of ability necessitates the sheltered environment provided by the team-teaching model, and requires frequent repetition of input and active learning techniques such as group work to allow learners to support each other. In each condition, a small number of learners fell into the B1 and B2 bands, with TOEIC scores between 550 and 945. These learners are capable of writing independently on a greater range of topics, and are more able to use a variety of rhetorical styles. They can express themselves at greater length in spoken interactions and can generally cope with more complex information that is presented aurally. These individuals typically required less support from teachers and were often more willing to speak in front of the whole class.

The control group's mean maximum score on the TOEIC was 475.00 (SD = 137.76), while that of the experimental group was 498.48 (SD = 146.32). An independent-samples Mann-Whitney Test showed that this difference in overall ability was not significant (Con. *Md* = 470, *n* = 23, Exp. *Md* = 455, *n* = 23, *U* = 245.000, *z* = -0.429, *p* = .668, *r* = 0.06).

Table 3.5. Participants grouped into approximated CEFR bands

CEFR band	TOEIC cut score	Control group	Experimental group
A1	120	0	0
A2	225	18	17
B1	550	4	4
B2	785	1	2
C1	945	0	0
C2	/	0	0

Note: TOEIC cut scores are combined values for listening and reading sections of the TOEIC.

3.6.2. Input corpora

For both the control and experimental conditions, all of the lessons were recorded on an iPod camera and then uploaded to a computer. The camera was positioned at the side of the classroom with a view across the front of the room where the instructors would usually stand.

The video recordings were then transcribed to create an input corpus for each group. The purpose of creating these corpora was to develop a record of the input that all learners received. As such, only instances of teacher talk which were directed at the entire class were transcribed. Clearly, each learner's actual input would also contain unique interactions with instructors and other classmates, but attempting to include this in the data collection would have required providing microphones for individuals or for each group of students and would have greatly added to the burden of data collection. It would also have presented a challenge for data analysis, as each learner's exposure would be unique.

Much of the course content was presented in the form of documentary videos or written texts and worksheets. The transcripts of these materials were inserted into the input corpora at the points in which they appeared in class. The corpora therefore contain a mixture of spoken and written input. Annotations were also added to the corpora to facilitate understanding of what took place in class. These annotations included references to class activities, paralinguistic communication, and actions by class instructors such as writing on a whiteboard. Instances of learner speech were given simple annotations, but learners' actual words were not transcribed. The names of learners, the course instructors, and other individuals

working in the institution were all replaced with pseudonyms. A complete list of annotations used is provided in Appendix E (page 439).

The input corpus for the control group, containing transcripts from 28 lessons and the final exam, came to a total of 245,363 tokens. Because of changes to the academic calendar in the year the experimental group took the course, there were 30 lessons and the final exam that semester. The corpus for the experimental group is therefore slightly longer at 297,507 tokens. Appendix F (page 440) contains the word counts provided by different sources of input in each lesson across both conditions.

One limitation of the input corpora data is that they do not take into account exposure to other learners' speech. As was mentioned above, transcribing learner-to-learner interactions would have been impractical, but it would be possible to record learner utterances that were directed to the whole class. It was not expected, however, that such incidences would comprise a large proportion of classroom input. It has been noted in the literature that Japanese learners tend to experience anxiety more than other groups when speaking English, possibly due to the emphasis in secondary education on acquiring language to pass entrance examinations rather than communicative competence (Woodrow, 2006, p. 321; Yashima, 2002, p. 54). Based on prior experience of teaching the course, it was expected that most learner speech to the whole class would consist of short responses to the teachers' questions.

In order to estimate exposure from learners speaking to the whole class, 10% of all such instances were sampled and transcribed. In around 4% of cases, the learners' words were inaudible. In these cases, the mean length of all audible utterances was

used as a replacement value (mean utterance length in control condition = 2.51 words; experimental condition = 1.94). The majority of these short utterances were instances of learners responding to questions or prompts from teachers. Even in the audible cases, students' voices were often rather quiet, and the teachers frequently repeated or paraphrased responses so that the whole class could hear. It is likely that the learners were projecting their voices only as much as necessary for the teachers to hear, rather than directing their responses to the whole room. The values in Table 3.6 indicate that participants in the two conditions were likely exposed to a similar number of words of learner speech, but that much of this occurred during student presentations during the final class periods and could therefore only have influenced the final writing assignments. Overall, it seems that the total contribution to the input corpora from learner speech to the whole class would be relatively small, with the exception of the final lessons. There are also concerns over how audible this input would have been for all participants, and since many responses were often repeated by the teachers, it was felt that the input corpora should only contain teacher talk and classroom materials. This shortcoming is one that could be addressed in a future study, possibly with data collected from a smaller number of participants to allow for input from student-to-student interactions to also be recorded.

Table 3.6. Estimated classroom input from course participants speaking to whole class

Condition	Mean utterance length in words (SD)	In-class speaking	In-class speaking as % of input corpus	Presentations	Total	Total as % of input corpus
Control	2.51 (2.59)	5296	2.16%	12310	17606	7.18%
Experimental	1.94 (1.81)	3837	1.29%	13460	17297	5.81%

Note: All values are estimates based on 10% samples of all instances of learner talk to the whole class.

A further issue to address is that of the teacher-researcher relationship. Maintaining objectivity is a concern that must be addressed in classroom-based research, as to some degree, such studies involve the evaluation of oneself as a teacher. Furthermore, while researchers strive to maintain as impersonal a relationship to their subjects as possible, a teacher has a duty of care to their students and a clear interest in maximising their success and well-being. Since the author of this study was one of the teachers of the Japanese Popular Culture course, it was important to consider how to treat the control and experimental groups as similarly as possible within the set up of the study and how to limit the influence of subjective decisions on the participants' behaviour and data collection.

As was mentioned in section 3.4.1 (page 60), the regular course content material was taught using the same materials in both conditions, while the treatment of vocabulary differed by focusing on high frequency and academic vocabulary for the control group and metaphorical vocabulary in the experimental condition. During the periods of class time in which there was an explicit focus on vocabulary learning, classroom input would therefore have been somewhat different between the two conditions. These differences would have appeared in both aural input from the teachers, and in written input in the form of class worksheets.

Another factor that potentially created differences in exposure was the teachers' use of metaphorical vocabulary in incidental input, that is, the use of metaphor when discussing the course's anthropological content. One of the tenets of conceptual metaphor theory, supported by the findings of Steen et al. (2010), is that metaphorical language is indispensable when discussing abstract matters, and it would therefore not be surprising that the teachers of this course also drew upon

such linguistic resources in both conditions. However, the act of providing explicit instruction in all likelihood raised the salience of the target forms in the minds of instructors and increased the likelihood that such forms would also appear in incidental input. While this is a noteworthy point, it was also an expected outcome; the study investigated the effects on learner output of a sustained effort to raise awareness of metaphorical vocabulary through both explicit and implicit instruction.

In order to examine potential differences in input, the compositions of the two input corpora were examined. This revealed that teacher 1 (the author of this study) provided the greatest proportion of input (measured by word count) in both conditions, and that the two instructors in combination provided approximately 79% of the input in both conditions (Table 3.7 and Appendix F, page 440). It should be noted that due to unavoidable circumstances (illness and conference participation), teacher 2 was unable to join two of the experimental condition regular class periods and the examination period.

Table 3.7. Log-likelihood tests for composition of input corpora

Condition	Teacher 1 (% of total)	Teacher 2 (% of total)	Input type Audio materials (% of total)	Written materials (% of total)	Total
Control	114,381 (46.62%)	80,339 (32.74%)	35,536 (14.48%)	15,107 (6.16%)	245,363
Experimental	142,395 (47.86%)	92,432 (31.07%)	36,935 (12.41%)	25,745 (8.65%)	297,507
Log-likelihood	-44.14****	118.22****	429.15****	-1131.08****	
Bayes factor	30.94+++	105.02+++	415.95+++	1117.88+++	

**** = Significant at the 0.0001 level
Bayes factor effect size: +++ = Very strong evidence against the null hypothesis

Log-likelihood tests on the four input types revealed highly significant differences across the two conditions with very strong effect sizes in all cases.

Negative log-likelihood values for the teacher 1 and written materials word counts indicated that these categories were underrepresented in the control group relative to the experimental group, while positive values for teacher 2 and audio materials suggested that these types were underrepresented in the experimental corpus.

To examine the degree to which teacher 2's absences accounted for these differences, the data were recalculated with adjusted values for teacher input for the two regular classes and the examination period in which teacher 2 was not present[25]. With these approximated values, the calculations were repeated, and this time, the polarities of the log-likelihood values for the two teacher input categories were reversed, suggesting now that teacher 1 contributed more to the control condition corpus and teacher 2 to the experimental group corpus (Table 3.8). This implies that while there were differences in the quantity of teacher input that the two groups received, this may have been due to external factors rather than a bias caused by the experimental treatment. Differences in input will be discussed in section 6.3, page 240).

Table 3.8. Log-likelihood tests for composition of input corpora with approximated teacher input values for classes when teacher 2 was absent

Condition	Input type				Total
	Teacher 1 (% of total)	Teacher 2 (% of total)	Audio materials (% of total)	Written materials (% of total)	
Control	114,381 (46.62%)	80,339 (32.74%)	35,536 (14.48%)	15,107 (6.16%)	245,363
Experimental	134,886 (45.41%)	99,465 (33.49%)	36,935 (12.43%)	25,745 (8.67%)	297,031
Log-likelihood	42.45****	-22.40****	420.29****	-1141.86****	
Bayes factor	29.65+++	9.20++	407.09+++	1128.65+++	

**** = Significant at the 0.0001 level
Bayes factor effect size: ++ = Strong evidence against the null hypothesis; +++ = Very strong evidence

[25] The details of how the adjusted values were calculated can be found in Appendix F, page 440.

3.6.2.1. Taught and untaught target metaphors

In addition to providing a measure of classroom exposure to language, the input corpora were used to help make an important distinction between target metaphorical language that was explicitly taught in class and that which was not explicitly taught. As was described in sections 3.4.1 and 3.5, the experimental group received explicit instruction on target language grouped into metaphorical themes that were appropriate for describing course content. There was clearly the possibility that this instruction could have led to increased usage of such language in learner writing. In this case, however, it would be unclear whether learner production had been influenced by a greater awareness of metaphor in language or by the increased emphasis on those particular word forms. In order to investigate this, a distinction was drawn between forms that appeared in targeted metaphorical themes and were explicitly taught as metaphors (taught target metaphors) and forms that could be identified as belonging to the same metaphorical themes using Steen's procedure but which were not taught in class (untaught target metaphors). If experimental group learners made increased use of both taught and untaught target metaphors, this might indicate a greater awareness of metaphor as well as individual word forms.

Table 3.3 listed the target metaphor themes that were included in the metaphor workbook. Most of these themes were introduced in class, and learners completed activities that required them to consider literal and metaphorical meaning. For example, in lesson four of the experimental condition, learners were studying plant metaphors and had already seen that the roots of a plant are its connection to the earth, are important for its development, and are typically hidden from sight. They

were then presented with the following workbook example (40), and the teachers elicited the metaphorical meaning of *roots*.

(40) *<WORKBOOK> Although tempura is often thought of as a typical Japanese dish, it has its roots in Portuguese food that was introduced to Japan in the sixteenth century.*

<TEACHER 2> So you can see this is metaphorical because tempura is not a plant itself, right?

<TEACHER 1> It's not real roots. It's metaphor roots. So what does it mean here?

<TEACHER 2> What would the roots of tempura be?

<TEACHER 1> Let's just do this together. Let's do number one together. Tempura has its roots in Portuguese food. What does roots mean there?

<STUDENT> Origins.

<TEACHER 1> Origins, yeah, yeah. Okay, so this is why we're doing this, making you think of the [...], the core meanings, all right? [EXP L4]

In addition, some words that did not appear in the workbook were also taught explicitly. This often happened because workbook activities were designed to be somewhat open-ended so as to encourage learners to draw on their own knowledge and experiment with new language. In the plant metaphors exercise, learners were asked to label pictures of plants, and one picture, which was intended to elicit *wilt*, also attracted the response *wither*. This was listed on the class whiteboard with other responses and addressed in class by the teachers (41). Terms such as these, which could be linked to the same conceptual metaphor, were also recorded as taught target metaphors since learners received explicit instruction on them.

(41) *<TEACHER 1> ...if you see a plant that is doing this, what do you think about that plant? How do you feel about it? It is...?*

<STUDENT> Needs water.

<TEACHER 1> Needs water.

<TEACHER 2> Uh-huh.

<TEACHER 1> [...] So in terms of health? Is this a healthy plant?

<STUDENT> No.

<TEACHER 1> No, it's an unhealthy plant. Fine.

<TEACHER 2> Maybe it's too cold outside.

<TEACHER 1> Yeah, yeah, it's not happy. So things that are withering or sagging are not happy, not healthy. [EXP L4]

To identify untaught target metaphors in the output corpora, the data were semantically tagged with the Wmatrix program. This allowed for lists of other MRWs with the same semantic tags, which might also belong to the target metaphor themes, to be draw up. MRWs with the same semantic tags as taught target metaphors which could be grouped into these target metaphors themes using Steen's procedure for linking linguistic to conceptual metaphors were listed as untaught target metaphors.

Examples of untaught target metaphors can be found for the metaphor theme MORE OF AN ABSTRACT THING IS AN INCREASE IN SIZE OR HEIGHT. Workbook activities explained how words like *massive* or *soar* can also be related to abstract notions such as popularity. As well as these forms, some learners writing produced terms such as *expand* and *extend* (42 - 43), which can also be related to the same metaphor theme

using Steen's procedure. While these terms did appear in classroom input (44 - 45), there was no attempt to explain the relationship between literal and metaphorical senses. Accordingly, terms such as these, as well as those that learners produced but did not appear at all in input, were coded as untaught target metaphors. A complete list of taught and untaught target metaphors is included in Appendix G (page 443) and Appendix H (page 450).

(42) *I think it is okay to increase foreign wrestlers because sumo is expanding around the world.* [EXP5]

(43) *I think that very good extend the culture of Japan.* [EXP16]

(44) *<TEACHER 2> The weaving trade. So the people in the profession, in the trade, are weaving. Kyoto's mayors supported them, and expanded?*
<TEACHER 1> I think so.
<TEACHER 2> Expanded the weaving trade... [EXP L10]

(45) *<WORKBOOK> There are two ways we will work to build vocabulary in this course: 1. Through learning new words and expressions 2. Through extending your knowledge of words that you may already know* [EXP L1]

One issue that arose with the coding of taught and untaught metaphors involved metaphorical adjectives. The metaphor workbook contained one section that encouraged learners to consider how the meanings of adjectives could be extended through metaphorical thinking. However, while words that had been taught in class could be categorised as taught target metaphors, the list of untaught target metaphors was potentially very large as learners produced several hundred metaphorical adjectives. Unlike other themes, these were not related to each other

by semantic links but by grammatical function. Accordingly, rather than code all metaphorical adjectives as taught or untaught target metaphors, it was decided to limit untaught target metaphors to those from within the first 500 word families on the BNC-COCA word list (http://www.lextutor.ca/vp/comp/). Since instruction had focussed on the extended meanings of frequent adjectives, it was felt that this decision reflected the goal of the metaphor instruction learners received.

One issue that arose in the process of identifying metaphors in input was that the input corpora data were more nuanced than the learner writing. It was immediately apparent that word forms were often used literally while providing explicit instruction of that same word's metaphorical meaning, as in (46).

(46) *<TEACHER 1> What is happening up here?*

<STUDENTS> <RESPOND>

<TEACHER 1> Evolution, yeah. So, living things gradually change over time, yeah. Cultures also gradually change over time, yeah. Japanese culture now is very different from Japanese culture one hundred years ago or two hundred years ago. So the point is, you can use the word evolve to describe a culture. You could say Japanese culture has evolved, cos it's done the same thing. It has changed. [EXP L1]

In this case, *evolution* is being used in its literal sense as it refers to a picture prompt showing the evolution of human beings as a species. Later, the word *evolve* is used neither literally nor metaphorically; it performs a self-referential function. Finally, the teacher's explanation leads to the metaphorical form *evolved*.

Such cases presented a dilemma for the coding of data. However, the pedagogical view taken in this study is that metaphor is a primary driver in the creation of polysemy, and that improvements in lexical knowledge require learners to consider how literal and metaphorical meanings are related. Accordingly, it was decided to count literal and self-referential uses of a word form as part of the frequency counts for each metaphor if they were being used as part of the explanation for a metaphorical term.

Another issue was that some metaphorical word forms were used to describe referents other than cultures or cultural entities. Of course, the term *culture* itself is very broad and resistant to precise definition. After consulting several dictionaries, a definition was settled on that limited cultural reference to the ideas, beliefs and behaviours of the people of a particular area. Metaphors that referred to concepts or entities outside this definition, such as economies, businesses or populations (47), were not included in the frequency counts.

(47) *<TEACHER 1> Okay, cos the young people have to pay the money to support the old people, right, with their pensions, with the taxes and things. So this system would be fine, no problem. You've got lots and lots of young people to support the older people, no problem. But more recently, you've got a bigger problem, right, because there's less and less and less younger people to support the elderly, and the elderly, the numbers of elderly are growing, relatively.* [EXP L19]

Another feature of classroom discourse is the repetition of key words. Since the class was team taught, this occasionally occurred as instructors shadowed what each

other had said and offered summaries to facilitate comprehension by learners (48). These cases are somewhat problematic for coding since repetitions frequently offered no further context than what was originally given. However, since each utterance did provide another encounter with a target form, it was decided to include repetitions in the frequency counts.

(48) *<TEACHER 2> Okay, err, "Her small idea has…"?*

<STUDENT> <RESPONDS>

<TEACHER 1> Evolved.

<TEACHER 2> Evolved.

<TEACHER 1> Yeah.

<TEACHER 2> Evolved into a larger project. [CON L18]

Once metaphors that described culture had been identified in the input corpora, they were recorded along with the output corpora data in Excel files. The date on which they appeared in class was included for each metaphor. This combination of data allowed for frequency, recency and timing to be included in the analysis, providing a rich diachronic description of classroom exposure and learner output.

3.6.3. Output corpora

As a regular homework activity, learners in both conditions were assigned a weekly writing task that required them to reflect on course themes and to express their opinions about an element of the course content that had recently been introduced. Students were asked to write in full paragraphs and to produce a minimum of one page of double-spaced writing in a B5 notebook that they had been given.

The written homework was submitted in the following lesson and then scanned and transcribed onto a computer file. Two versions of each transcription were made. For the first of these, efforts were made to preserve the authenticity of the data as much as possible by transcribing the text exactly as it had been written, including errors and idiosyncratic use of punctuation. In the second version, spelling mistakes were corrected to allow for easier identification of words when the corpora were analysed. However, the examples of learner output presented in this thesis are all taken from the more authentic, uncorrected version of the data.

Minimal annotation was required for the output corpora. Only learners' pseudonyms and the date were added to each text. The output corpus for the control group contained 46,218 word tokens and that of the experimental group came to a total of 59,801 word tokens[26]. The number of lexical units contributed by each learner and the mean number of lexical units per learner per week is provided in Appendix I (page 455).

3.6.4. Data on sources of metaphor in output

Once learners' written homework had been collected each week, the language instructor provided responses to students' ideas and gave feedback about language errors using a simple correction code (See Appendix C, page 431). This was intended to encourage learner experimentation with language by offering feedback in a manner that was supportive and that required learners to reflect on their mistakes without feeling that they had been penalised for them.

[26] These values are token counts in the corpora. They are different from the total number of lexical units identified by MIPVU.

Also at this stage, examples of metaphorical language were noted for further investigation with the learners. The purpose of this was to enrich the output corpora data by obtaining feedback on the sources of metaphorical language that learners had chosen to use. As corpora offer evidence only of output, and not of the processes that influence lexical choice, it was felt that obtaining such feedback from learners could usefully supplement the data.

Multiple criteria were used to select words for inclusion on the feedback forms. Firstly, vocabulary from outside the first 1,000 most frequent words on the BNC-COCA frequency lists (available at http://www.lextutor.ca/vp/comp/) were considered of interest since they may have been words that learners were attempting to add to their productive vocabularies. Also, since the course itself was an obvious potential source of vocabulary for learners, metaphorical words or phrases that learners used which had previously appeared in the classroom input were selected. Learners' first language was another potential influence, so metaphorical words that were known to also function metaphorically in the L1 were chosen. L1 usage was confirmed by checking the online Japanese-English dictionary *Jishou* (http://jisho.org), which provides many translated example sentences. Examples in this category include the word *boom*, which exists as the loanword ブーム (*buumu*) in Japanese, though maintaining only its metaphorical *dramatic increase* meaning, and *born*, the translation of which (生まれる- *umareru*) is also used in Japanese to refer to the coming into existence of things other than living creatures. A final category for selection was metaphorical language that was being used unconventionally in English. Words coded as metaphorical which did not match the definitions provided by the Macmillan or Longman dictionaries were considered

unconventional, as were those which did match a definition in one of the dictionaries, but were being used in ways that created alternative interpretations. Example (49) includes the word *deported* being used in a way that does not appear in either of the above dictionaries, while in example (50), *absorb* appears to match sense number 5 in Macmillan (*if something absorbs you, it is so interesting or entertaining that it takes all your attention*), but the sentence as a whole seems to imply a more literal meaning.

(49) Asasyoryo broke these rule, and he was deported from sumo's society. [EXP3]

(50) Sumo makes old people absorb. [CON23]

In the next lesson, the notebooks were returned to learners and a short period of class time was used for students to read the instructor's comments and to edit their work according to the correction code. This also provided an opportunity for the language instructor to ask learners who had produced metaphors that had been selected for further investigation about their reasons for choosing that particular word or phrase. These learners were asked to complete a bilingual form which required them to state whether they had used an L1 or L2 dictionary entry, had translated from their L1, had used language from class, had already known that language, or had produced the selected language for some other reason. Learners were allowed to select one or more of these options on the form, a monolingual example of which is given in Figure 3.1.

One important feature of these forms was that they could be completed in a short period of time without disrupting the class itself. Fatigue effects and low response rates are a noted issue with questionnaire-based research (Dörnyei, 2003,

As part of my research, I am interested in why learners choose to use particular language in their writing. Please look at the underlined part of the sentence below and think about why you chose to use those words. Then answer the question below.
Language from homework Over many years, Sumo has faced many <u>obstacles</u> along its' history.
Why did you use the underlined word(s) in the sentence above? You can check more than one reason if you wish. I found the word(s) in a dictionary… in English □ in my own language □ I translated from my own language (without using a dictionary). □ (Please provide the expression in your native language.) …….. I used word(s) we studied in class. □ I knew this language before I began this class. □ Other □ (please explain): ……..

Figure 3.1. Source of metaphorical language feedback form

p. 14). Because of this, the form used check boxes for simplicity, while providing two open-ended options to allow for all possible responses. The response choices were written to be as specific, easy to answer, and clear as possible (Brown, 2001, p. 18 - 19). Dörnyei (2003, p. 10) warns that the user-friendly nature of survey data can lead to responses that are somewhat superficial. However, they served as a useful complement to the corpus data, in particular by allowing those learners who were drawing on their L1 to provide translations of their intended meaning.

Other issues that Dörnyei describes with questionnaires are the potential for prestige bias and self-deception (2003, p. 12 - 13). Prestige bias refers to situations when respondents select an answer choice not because it is true of them, but because it is perceived as being more desirable. Care was taken in developing the form to avoid word choices that would indicate positive or negative connotations. The most likely answer choice to suffer from this form of bias seems to be 'I used word(s) we studied in class.' It is possible that learners felt their efforts would be

appreciated if they indicated that they had tried to use words from class. In the case of self-deception, respondents unconsciously provide inaccurate information by suppressing details that may be embarrassing, uncomfortable, or otherwise negative in favour of a more desirable response. It is commonly held that our motivations can bias our beliefs (Mele, 1997, p. 93), and if learners perceive one or more of the answer choices as being less prestigious in some way, it is possible that they may avoid them. These questions will be returned to in section 7.6.1 (page 347).

Language examples were selected from as wide a range of participants as possible each week in order to obtain feedback from the entire class and to avoid overly burdening particular individuals. In the control condition, 111 metaphor forms were completed over whole semester - a mean of 7.4 forms per week. Learners received a mean of 4.83 forms each per semester (SD: 2.08, Min.: 1, Max.: 8). Experimental group learners completed 142 forms during the semester, with a mean of 9.47 per week. There was a mean of 6.17 forms completed per learner (SD: 1.67, Min.: 3, Max.: 9). In total, 157 metaphor sources were provided in the control condition, and 197 were provided by experimental group learners.

3.7. Summary

This chapter has described the context for the study and the participants involved. It has outlined the procedure for the control and experimental groups and explained the rationale for each condition, and the methods of data collection and their limitations have been described. The following four chapters will present different aspects of the data analysis. Chapter four provides a comparison of metaphor produced by the two study groups. Chapter five examines the relationships between

learner ability and metaphor production. Chapter six offers an in-depth examination of the relationship between classroom input and learner output. Chapter seven investigates factors affecting learner production of metaphor: word frequency, part of speech, phraseology and the influence of the L1. Chapter eight then draws together the conclusions of the study, considers its broader implications, and offers suggestions for future research.

CHAPTER 4. METAPHOR PRODUCTION UNDER DIFFERENT TREATMENT CONDITIONS

4.1. Introduction and rationale for the study

This chapter will examine the production of metaphorical language under both the control and experimental conditions. Metaphors were identified using a variation of the MIPVU, and the data were then filtered to allow for more fine-grained analyses of differences in production between the two groups, taking into account parts of speech, target language production, language variety, and the effect of writing topic.

Despite the growth in interest for studies on metaphor production by language learners, there are still large gaps in our understanding of how learners develop the competencies to make use of metaphor in their own output. Several studies have now applied explicit identification procedures to learners' written output (Chapetón, 2010; Hoàng, 2015; Littlemore et al, 2014; Nacey, 2013; Turner, 2014), but much remains to be investigated. Studies set in classroom contexts are rich in promise, firstly as learners spend considerable time in these environments, but also because the availability of reference materials both inside and outside of class may support learners' experimentation and acquisition of new language (Dziemianko, 2010, p. 262 - 267). While the decision to allow for the use of dictionaries, class notes or other reference materials in this study limits what can be said about learners' actual competence, this approach allows for the examination of how learners go about adding expressions to their productive repertoire.

Boers (2004, p. 215 - 217) has suggested that awareness raising should be a regular activity in the classroom in order for long-term changes in language production to manifest themselves. A "general foregrounding of metaphor," (MacArthur, 2010, p. 157) would allow for variation in learning styles and would bring metaphor to learners' attention as part of the general experience of learning a language. By taking a more naturalistic approach, it is possible that such longitudinal studies could be more revealing of the processes by which learners begin to develop metaphorical competence than tightly controlled 'one-off' studies that allow little time for changes in cognitive behaviour to occur.

Little is known about how a sustained approach to metaphor awareness raising will impact on learner production. While it might be assumed that the teaching of any language feature would lead to increased output, the possibility that learners will avoid language that they regard as difficult should not be overlooked, and an examination of the overall quantity of metaphor produced seems an appropriate place to start. However, this can only be considered a very blunt appraisal of developments in learner language. This chapter will then consider several more detailed elements of learners' output.

Greater use of metaphor from open-class parts of speech has been observed in the writing of higher proficiency learners completing more demanding tasks (Littlemore et al., 2014). Since vocabulary development is seen as an incremental process (Schmitt, 2010, p. 19 - 22), it may be the case that as learners become aware of the range of meaning encompassed by a particular lexical item, they can apply it to new contexts and awaken its metaphoric potential. The relationship between learner ability, metaphor and word class will be explored in further detail in chapter

five, while this chapter examines the word classes of metaphors produced in the control and experimental conditions as whole groups.

An important question to consider is not only whether awareness-raising activities lead to increased production of metaphor, but also whether learners are able to draw on their knowledge independently. The ability to apply topic and vehicle combinations appropriately was identified by Low (1988, p. 132 - 133) as one component of metaphoric competence, and as with vocabulary teaching in general, the sheer scale of this challenge dictates that learners must be guided towards developing a 'sense' for appropriacy and not allowed to become reliant solely on the teacher for their lexical growth. It would be of interest then, to examine firstly whether the explicit teaching of target metaphor leads to increased production of those metaphors, and secondly whether evidence can be found that learners who receive this explicit instruction are more able to extend their production of metaphor to include other untaught members of particular conceptual metaphors.

Finally, the influence of topic on metaphor production will be addressed. Applications of MIPVU to different genres of English have revealed stark differences in metaphor use (Steen et al., 2010) and the various writing tasks in examinations may also elicit markedly different amounts of metaphor (Turner, 2014, p. 117 - 127). Genre is a different notion from topic, one that pertains to the conventional features of a particular mode of discourse, including its purpose, participants, and key elements (Biber & Conrad, 2009, p. 15 - 19; Semino, Deignan & Littlemore, 2013, p. 43), but these findings suggest that the influence of topic should not be overlooked. The degree to which topic affects output is an important consideration because many studies on metaphor production are reliant on sampled data from banks of

essays or corpora. If topic matter were shown to influence metaphorical output, this would greatly impact the generalizability of findings in all but the most controlled sampling conditions. The specific research questions that will be addressed in this chapter are as follows:

a) Will there be differences in the amount of metaphor produced between the control and experimental conditions as measured by MIPVU? Will learners writing in classroom conditions produce greater levels of metaphor than learners in other studies writing under examination conditions?
b) Will the written output of the two groups show any differences in the amount of metaphors used from different parts of speech? Is there a shift towards more metaphor from open-class parts of speech?
c) Will the experimental group produce more of the targeted metaphorical expressions than the control group? Will they produce metaphorical expressions that are related to the target language? That is, will they be able to extend metaphor themes appropriately?
d) What is the effect of topic on metaphor use in written output?

4.2. Methodology

4.2.1. Data sources

The results in this chapter are based on the output corpora described in section 3.6.3 (page 118). The analyses reported here were carried out on the versions of the corpora with corrected spellings, but reported examples of student output come from the uncorrected versions.

4.2.2. Application of MIPVU

The metaphor identification carried out in this chapter followed the procedure outlined by Steen et al. (2010) with the following exceptions:

- Some target metaphors that were coded as MRWs would not have been tagged as metaphorical under MIPVU owing to...
  - their being target metaphors that were listed in an entirely separate section of the dictionary (as opposed to separate numbered senses of the same section), or...
  - their being target metaphors whose dictionary definitions conflated literal and metaphorical senses.

In total, there were four exceptions in the control group data and 14 in the experimental condition. These exceptions respectively comprise 0.01% and 0.02% of the lexical units in each MIPVU analysis. All exceptions are listed in Appendix D (page 433), and raw data from this chapter can be found in Appendix J (page 459).

4.3. Overall levels of metaphor in the control and experimental groups

The overall levels of metaphor production for both groups showed that while all three types of metaphor were more frequent in the experimental condition, there appeared to be a very similar level of metaphor production between the two groups.

Independent samples Mann-Whitney Tests[27] found no significant differences in metaphor rates between the groups (Table 4.1).

Table 4.1. Mann-Whitney tests for metaphor rates between the two conditions

Metaphor type	Mean (S.D)		Median		Mean rank		*U*	*z*	*p*	*r*
	Con.	Exp.	Con.	Exp.	Con.	Exp.				
Indirect	11.41% (1.65%)	11.47% (1.46%)	11.15%	11.20%	23.09	23.91	255.000	-.209	.835	0.03
Direct	0.02% (0.04%)	0.05% (0.09%)	0.00%	0.00%	21.26	25.74	213.000	-1.389	.165	0.20
Implicit	0.13% (0.19%)	0.20% (0.22%)	0.07%	0.11%	21.65	25.35	222.000	-.938	.348	0.14
Total	11.56% (1.72%)	11.72% (1.53%)	11.71%	11.70%	23.04	23.96	254.000	-.231	.818	0.03

Note: *n* = 46 in Mann-Whitney tests (total number of participants)

These distributions of metaphor type are very much in line with other studies of learner writing, in that indirect metaphor comprises the overwhelming majority of MRWs. Turner's (2014, p. 105, p. 107) study of Japanese and French learners' examination writing found that direct metaphor respectively made up 0.02% and 0.04% of all lexical units. For implicit metaphor, the figures were 0.24% and 0.15%.

Figure 4.1 compares this data to other studies that have applied variations of MIPVU. In the figure, white columns indicate studies on native speaker data, black columns are for studies on English language learners whose L1 is not Japanese, and grey columns are for studies on Japanese learners of English. The two groups in this study are on the far right. The values reported in Figure 4.1 are the total number of MRWs as a proportion of the total number of lexical units, hence the difference from the values reported in Table 4.1.

[27] Small *n* sizes meant that non-parametric tests were used in this study.

It should be noted that when learners who are mostly of A2 level are writing on academic topics in a CLIL environment (with the aid of lesson materials and dictionaries), their metaphorical output is comparable to learners of far higher ability writing under examination conditions (as in the Littlemore et al. (2014) and Turner (2014) studies). This serves to highlight the impact of contextual factors on metaphor production. While the studies on Greek, German, French, and Japanese A2-C2 level learners in Figure 4.1 offer insights into metaphor production in an unaided setting, the present study is able to show how learners perform when given a degree of linguistic support and a genre that invites (or possibly even requires) metaphorical output. The native speaker production in the academic, fiction and news genres investigated by Steen et al. might also be considered to have been written under the same circumstances as the learners in this study. Clearly, there is a need for further investigation into metaphor production among different groups and contexts.

It is worth considering why little difference appeared in the overall levels of metaphor production between the control and experimental groups in this study. Both groups were encouraged to experiment with new language when completing the weekly writing activity, but only the experimental group received explicit awareness-raising activities on how metaphorical language is used to describe key course concepts. Given that the writing prompts were identical, it may be that the topics themselves necessitated a particular kind of language in written responses - certain concepts simply cannot be described without metaphor. Another possibility is that learners in the experimental group avoided using metaphorical language that was perceived as difficult. Studies have shown that learners can display a tendency

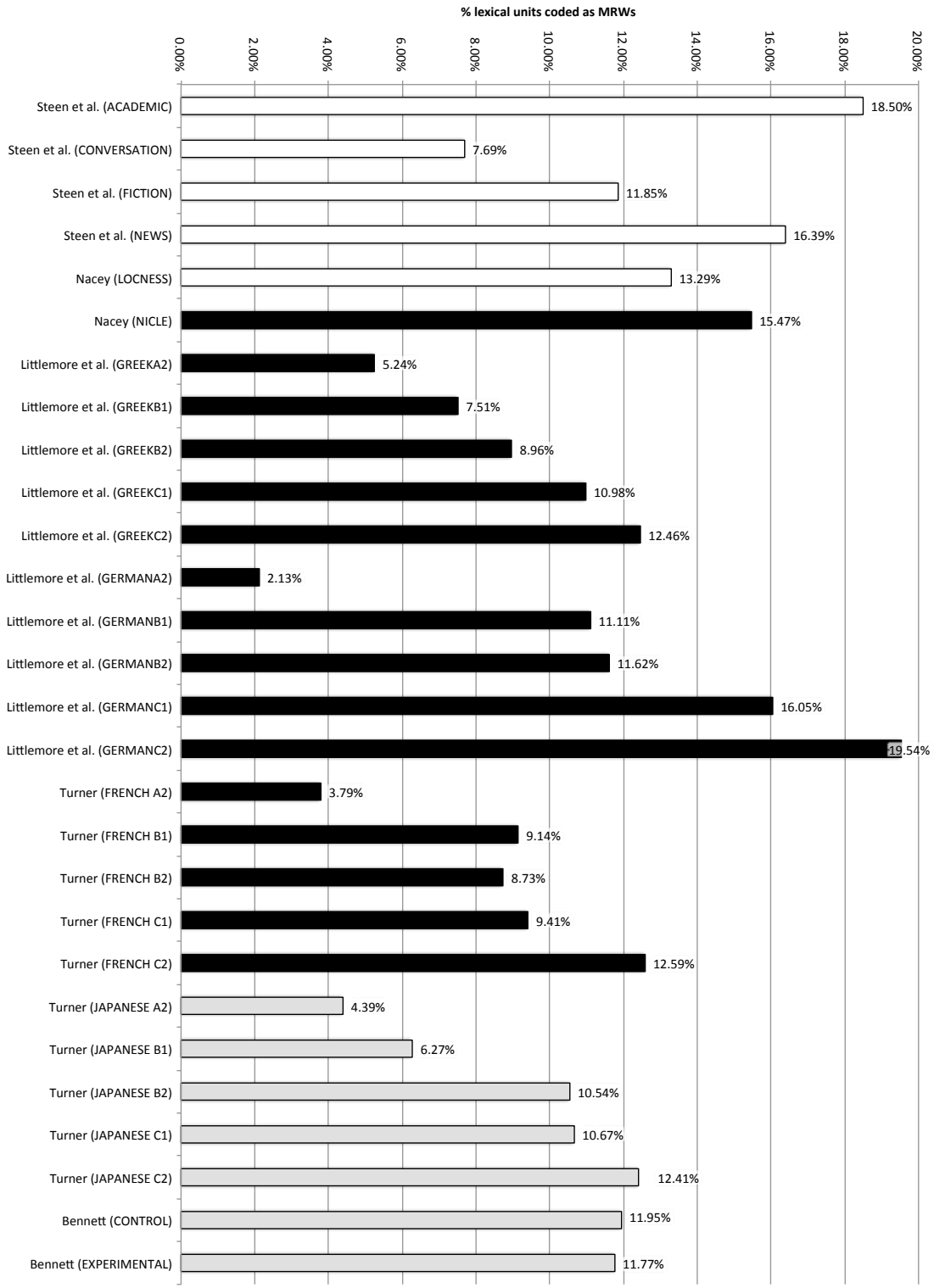


Figure 4.1. Percentage of lexical units coded as MRWs in this study and other MIPVU-related studies[28, 29]

[28] Data adapted from: Littlemore et al., 2014, p. 125 - 126; Nacey, 2013, p. 136 - 138, 149; Steen et al, 2010, p. 202 - 207; Turner, 2014, p. 90-91.

[29] The NICLE data does count uses of *of* or *for* as MRWs. This is to harmonise the coding with other studies (see Nacey, 2013, p. 137 - 139).

to stick to tried and trusted phrases as safe options in their linguistic output (Hasselgren, 1994, Groom, 2009, p. 32 - 33). In spite of the encouraging feedback they received and the instructions to experiment without fear of being penalised for inaccuracies, learners may have chosen to cling on to their 'lexical teddy bears'. A final possibility, and one that will be explored in the coming sections and chapters, is that, as a deliberately inclusive measure of metaphor, MIPVU operates on too broad a scale to detect the more nuanced changes that may be taking place in language development. In order to investigate learner production of metaphor in more detail, finer-grained analyses are required.

4.4. Metaphor production at the part-of-speech level

Rates of metaphor use across individual parts of speech varied only slightly. Table 4.2 shows the median and mean percentages of lexical units from each part of speech that were coded as MRWs across all participants. The values suggest that metaphor rates for individual parts of speech remained similar in the two conditions, and indeed none of the differences were significant according to independent samples Mann-Whitney tests (Table 4.3). Figure 4.2 displays the distribution of open- and closed-class metaphoric densities in box and whisker plots. The dots represent the mean values, while the boxes display the interquartile range, the range within which 50% of the data lie either side of the median value, which is shown by the lines through the middle of each box. The whiskers above and below each box show maximum and minimum values. The mean and median values for open-class metaphoric densities in the two conditions seem quite similar between the conditions, and indeed an independent samples Mann-Whitney found no significant

Table 4.2. Median and mean metaphor rates for parts of speech in both conditions

Part of speech	Median metaphor rate		Mean metaphor rate (SD)	
	Con.	Exp.	Con.	Exp.
Nouns	7.12%	6.48%	7.02% (2.09%)	7.32% (2.20%)
Verbs	20.19%	22.27%	20.95% (4.59%)	21.62% (4.60%)
Phrasal verbs	75.00%	55.56%	69.58% (31.57%)	59.84% (32.63%)
Adjectives	7.14%	8.11%	7.61% (4.10%)	8.54% (2.87%)
Adverbs	1.56%	2.08%	1.90% (1.77%)	2.01% (1.34%)
Conjunctions	0.00%	0.00%	0.10% (0.24%)	0.16% (0.36%)
Determiners	8.97%	9.90%	8.75% (3.56%)	9.96% (4.69%)
Prepositions	39.52%	40.18%	41.08% (4.94%)	39.97% (5.30%)
Other	3.13%	3.64%	3.24% (1.85%)	3.63% (1.93%)

Table 4.3. Mann-Whitney tests for metaphor rates across parts of speech in the two conditions

Part of speech	Mean ranks		*U*	*z*	*p*	*r*
	Con.	Exp.				
Nouns	23.22	23.78	258.000	-.143	.886	0.02
Verbs	22.70	24.30	246.000	-.406	.684	0.06
Phrasal verbs	23.93	20.33	191.500	-.958	.338	0.14
Adjectives	20.87	26.13	204.000	-1.329	.184	0.20
Adverbs	22.72	24.28	246.500	-.397	.691	0.06
Conjunctions	23.04	23.96	254.000	-.333	.739	0.05
Determiners	21.74	25.26	224.000	-.890	.374	0.13
Prepositions	24.70	22.30	237.000	-.604	.546	0.09
Other	22.13	24.87	233.000	-.692	.489	0.10

Note: *n* = 46 for all parts of speech except for phrasal verbs. Three participants in the control group did not produce any phrasal verbs (*n* = 43)

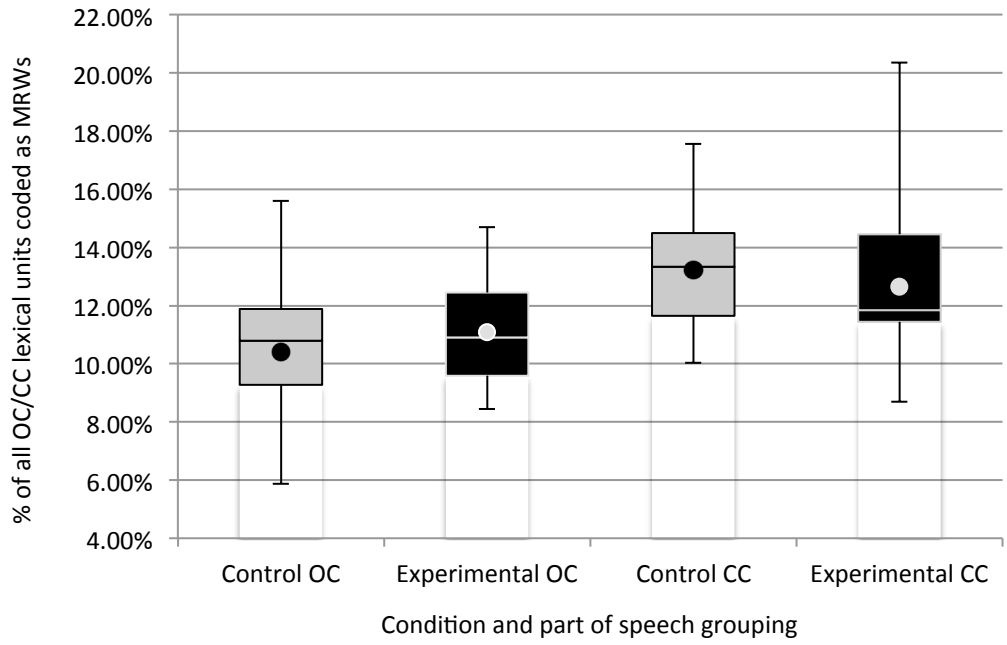


Figure 4.2. Distribution of open- and closed-class metaphor rates in the two conditions

difference between these values (Con. *Md* = 10.79%, *n* = 23, Exp. *Md* = 10.91%, *n* = 23, *U* = 213.500, *z* = -1.120, *p* = .263, *r* = 0.17). The values for closed-class metaphor produced in the experimental condition appear to be a little lower than the control, but again, Mann-Whitney tests indicated that these differences were not significant (Con. *Md* = 13.34%, *n* = 23, Exp. *Md* = 11.84%, *n* = 23, *U* = 210.000, *z* = -1.197, *p* = .231, *r* = 0.18).

Overall, in this data set, we do not see strong signs of a shift from the grammatically-required use of closed-class metaphor towards usage that is determined more by lexical choice. This may imply that the experimental treatment was ineffective or that there was some avoidance of metaphor, perhaps due to its perceived difficulty. Alternatively, the majority of learners may still have been at a level of proficiency where they were not able to demonstrate a wider range of lexical use in their written output. However, studies that have found shifts toward greater

use of open-class metaphor (e.g., Littlemore et al. 2014; Turner, 2014) have been based on larger samples of learner data from a broader range of ability than that of this study, and it seems most likely that fundamental changes in learner output can only be observed over these longer spans.

It has been noted that learners will gradually produce more open-class metaphor as they reach higher levels of proficiency and begin to engage with tasks that require discussion of more abstract topics, although separating the effects of topic and proficiency remains a challenge (Littlemore et al., 2014, p. 128). In this study, however, the writing task topics were controlled between the two conditions, and it therefore seems more likely that the increased metaphor production compared to learners of similar ability seen in Figure 4.1 was due to the academic nature of the writing topics. These would have been more similar to those of intermediate to advanced-level examinations than the more personal writing topics expected for an A2-level examination (Littlemore et al. 2014, p. 131 - 137; Turner, p. 62 - 64). The results of Figure 4.2 may then indicate that the participants' productive vocabularies were still limited in the range to which they could use open-class lexical items.

To investigate this further, the distribution of open- and closed-class metaphors in the studies from Figure 4.1 were examined (Figure 4.3). The values reported in this case are the number of open- or closed-class lexical items that were coded as MRWs as a proportion of the total number of open- or closed- class lexical units. Although topic and genre are somewhat different notions, the wide range of values found in the different genres of the Steen et al. studies suggest that the subject matter and discourse medium should not be overlooked when investigating metaphorical

language on a broad scale. The effect of writing topics in the language learner data and of editing in some of the Steen at al. genres makes it somewhat more difficult to comment on how proficiency affects metaphor use. However, the general pattern from the four groups of English learners in the table appears to be that both open- and closed-class metaphor production increases in line with the demands of higher level examinations, and that open-class metaphor use overtakes closed-class production in the B1 to B2 range or perhaps between B2 and C1 in the case of the Greek learners. However, this is not always a smooth development, as sudden increases and occasional regressions can be seen. The Greek, German, and French learners all exhibit a sudden jump in the use of open-class metaphor at the B1 level, but for Japanese learners this jump appears one band later at B2. This may imply that the results in Figure 4.3 reflect some element of linguistic competence as well as topic if it can be assumed that all B1 learners were responding to similar examination prompts. However, this is somewhat speculative and highlights the need for further studies in this area.

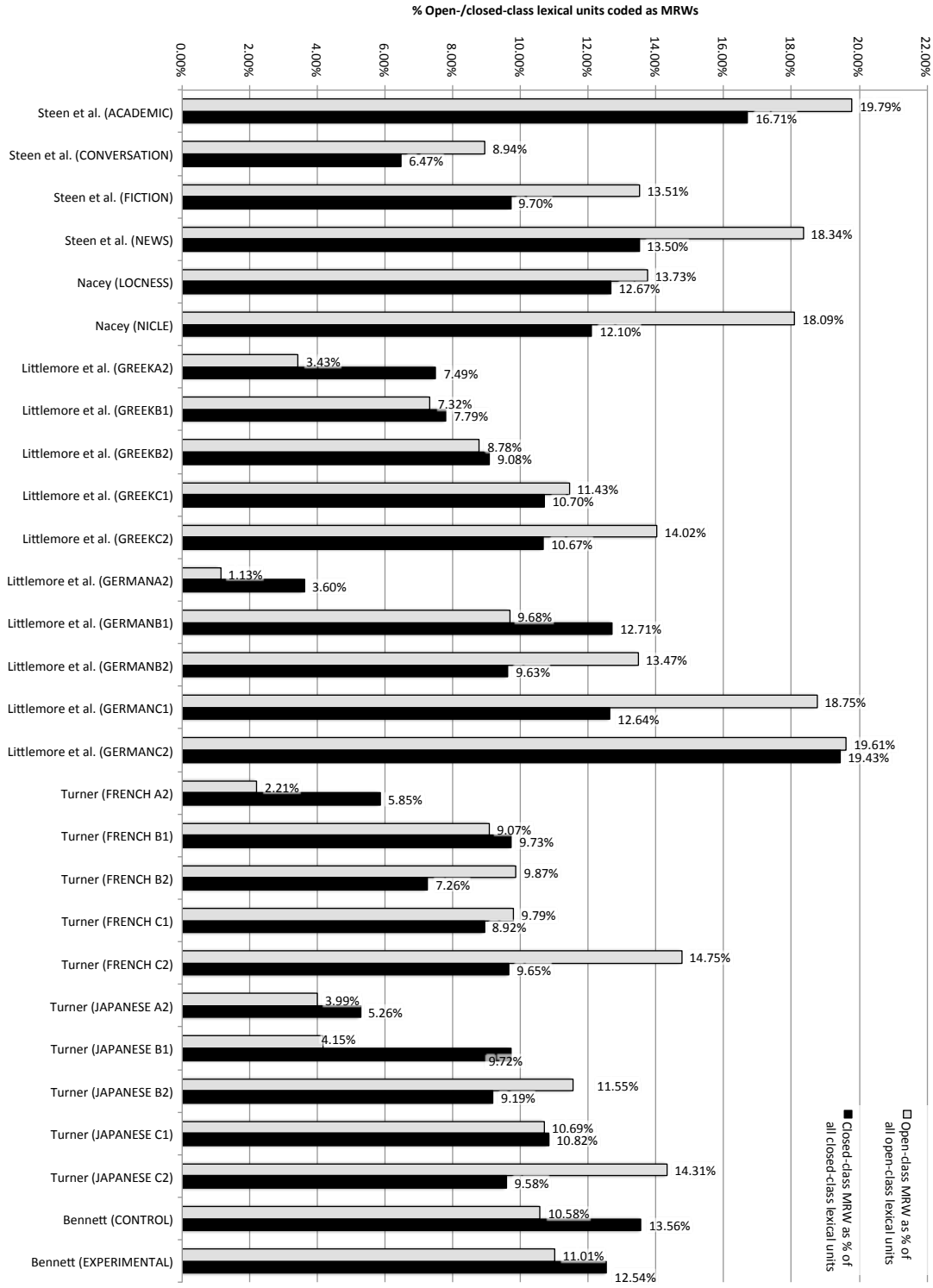


Figure 4.3. Percentage of open- and closed-class lexical units coded as MRWs in this study and other MIPVU-related studies[30,31]

[30] Data adapted from: J. Littlemore, personal communication, September, 2016; Nacey, 2013, p. 136 - 138, 149; Steen et al., 2010, p. 202 - 207; S. Turner, personal communication, March 2017.

[31] The NICLE data does not count uses of *of* or *for* as MRWs. This is to harmonise the coding with other studies as much as possible (see Nacey, 2013, p. 137 - 139).

4.5. The production of target metaphorical language

The next focus of investigation was learner production of metaphors that were targeted for explicit instruction. Both metaphors that were taught in class and those that were not taught but which could be considered part of the same metaphor theme as the taught expressions (i.e., untaught target metaphors) were analysed. Thus the aims were twofold: to examine the effect on written output of raising awareness of particular metaphorical expressions, and to assess the evidence for learners being able to apply their awareness of taught metaphor themes to extend their production to language that had not been taught in class.

The output of the control group serves as an example of learner production in an environment more similar to implicit learning. That is, the target metaphorical language was present in classroom input, but received little explicit classroom teaching (see section 6.3, page 240). For both the control and experimental groups, the amount of non-target open-class metaphorical language is also presented to act as a comparison.

As was mentioned in 3.6.2.1 (page 112), untaught target metaphors were identified with Wmatrix and linked to the conceptual metaphors that they instantiated through Steen's (1999) five-step procedure. The raw frequency counts for taught and untaught target metaphors produced by both groups are provided in Table 4.4. It can be seen that learners produced language related to all of the metaphorical themes, although there was wide variation in usage. Clearly, certain themes reflect more general concepts or topics that are given greater prominence in a course on Japanese popular culture. It is also apparent that metaphorical language

Table 4.4. Taught and untaught target metaphors produced in both conditions

Metaphor theme	Control group target metaphors		Experimental group target metaphors	
	Taught	Untaught	Taught	Untaught
PRECIOUS MATERIALS ARE VALUABLE, WONDERFUL OR BEAUTIFUL ENTITIES	1	0	16	0
Sources of metaphor: plants	5	0	8	0
MOVING FORWARD IS IMPROVEMENT	17	5	19	5
MOVEMENT IS FREEDOM	2	2	13	3
Sources of metaphor: the body	7	13	9	14
Metaphorical adjectives	70	111	107	178
CULTURES ARE LIVING THINGS	41	31	52	42
CULTURES ARE SPACES	33	12	108	7
CULTURES ARE SUBSTANCES	69	28	83	20
CULTURES ARE CONSTRUCTIONS	5	6	15	1
CULTURES HAVE MANY PARTS	55	15	33	28
CULTURES ARE POSSESSIONS	27	54	56	64
MORE OF AN ABSTRACT THING IS AN INCREASE IN SIZE OR HEIGHT	11	8	25	13
MORE IS UP AND LESS IS DOWN	13	7	10	4
STRONG EMOTIONS ARE SICKNESS/INSANITY	2	0	3	0
BEING INTERESTED IS ENTERING SOMETHING	3	1	7	0
STRONG EMOTION IS PHYSICAL FORCE	0	2	5	9
ABSTRACT ENTITIES ARE PHYSICAL STRUCTURES	24	19	30	17

reflecting these more common themes formed the bulk of target metaphor production in both groups. Since the target metaphors for the experimental group were selected based on an analysis of classroom input and learner output in the control condition, it was only to be expected that the two groups' output would shows signs of similarity.

The two output corpora differed in size by about 30%, so the raw frequencies were converted to percentages of all words in each corpus. Table 4.5 shows the median and mean metaphor rates for each category of target metaphor, while Table

4.6 shows the results of Mann-Whitney tests for taught, untaught, and combined (taught and untaught) target metaphor rates between the two conditions. The distributions of target metaphor rates are plotted in Figure 4.4. The results suggest that explicit teaching of target metaphors has had some effect on learner output; although the application of the Bonferroni adjustment meant that the difference in taught target metaphor production between the two groups did not reach significance, there was a moderate effect size. However, there was no change in untaught target metaphor production. This may imply that while explicit teaching of metaphorical language through awareness raising can affect output in the targeted forms, it does not lead to more independent use of untaught language that is related to target metaphors.

Table 4.5. Median and mean target metaphor rates in the two conditions

Target metaphor category	Median target metaphor rates		Mean target metaphor rates (SD)	
	Con.	Exp.	Con.	Exp.
Taught	0.86%	1.04%	0.86% (0.27%)	1.04% (0.27%)
Untaught	0.69%	0.70%	0.70% (0.29%)	0.70% (0.28%)
Combined	1.57%	1.75%	1.56% (0.41%)	1.74% (0.33%)
Non-target (OC)	4.57%	4.83%	4.47% (0.89%)	4.71% (0.74%)

Table 4.6. Mann-Whitney tests for target metaphor rates between the two conditions

Target metaphor category	Mean ranks		*U*	*z*	*p*	*r*
	Con.	Exp.				
Taught	19.22	27.78	166.000	-2.164	.030#	0.32
Untaught	23.00	24.00	253.000	-.253	.801	0.04
Combined	20.52	26.48	196.000	-1.505	.132	0.22
Non-target (OC)	21.17	25.83	211.000	-1.175	.240	0.17

Note: *n* = 46 for all tests
= Non-significant after Bonferroni adjustment

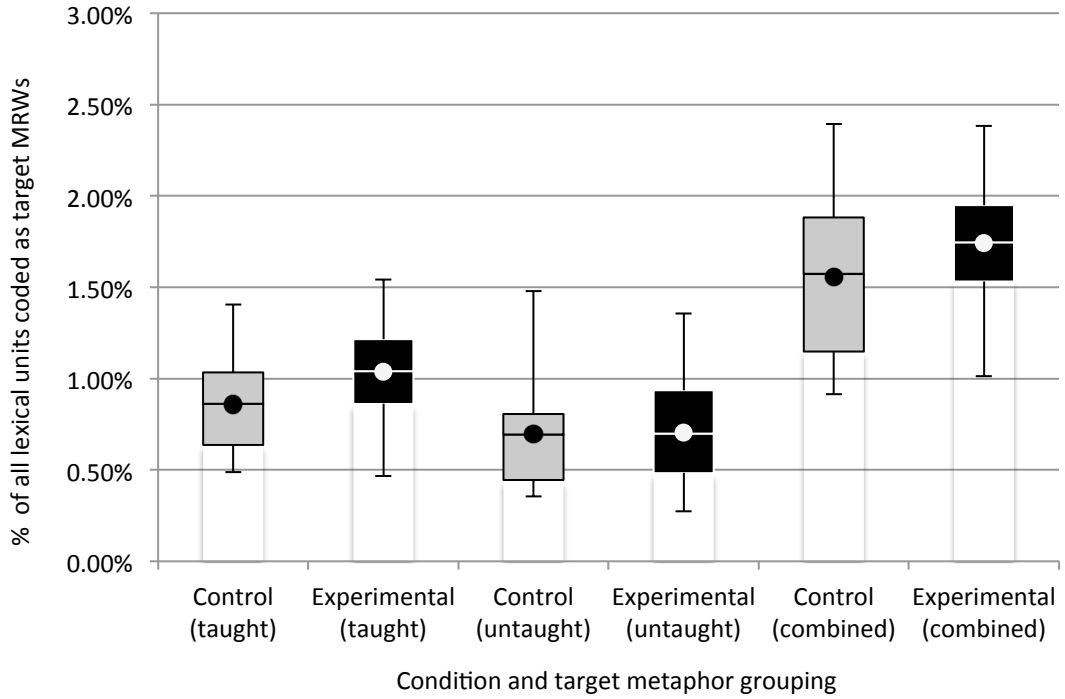


Figure 4.4. Distribution of metaphor rates for taught, untaught, and combined target metaphors in the two conditions

Although there was an increase in the amount of non-target open-class metaphor production, it was not particularly large. Nevertheless, the results imply that even in a CLIL environment, the majority of metaphor production is likely to be of language that is not directly related to course themes. Target metaphors comprised 25% and 27% of all open-class metaphor production in the control and experimental conditions respectively. Chapters five and seven will discuss the use of some non-target metaphors.

While there is some suggestion that learners respond to metaphor awareness-raising activities by producing more of the explicitly targeted expressions, there is also the question of whether their output becomes more varied. Increased diversity in production could be taken as evidence of a gain in metaphorical competence or

lexical fluency, as learners acquire a range of expressions to describe particular concepts.

However, assessing lexical variety is not a simple matter. Type/token ratios, in which the number of different word forms (types) is divided by the total number of words (tokens), have been used to assess lexical variety, but this ratio is affected by text length, with longer texts having lower values (Schmitt, 2010, p. 212 - 213). At the same time, however, not taking into account the length of a text would boost results for longer pieces of writing. In this analysis, therefore, the ratio of the total number of metaphorical target types to the total number of all lexical item types produced by each learner was calculated. Table 4.7 provides an example for participant 23 from the experimental group. This individual produced fifteen word forms related to taught target metaphors, which are listed in the second column of the table. Within the fifteen forms, there are ten unique types. Since he produced 602 unique word forms in all of his written output, the ratio for metaphor-related target types to all types is 0.0166, or 1.66%. The mean ratios for taught and untaught target metaphor types for all learners were calculated and used as a measure of metaphorical variety.

The metaphorical variety measures shown in Tables 4.8 - 4.9 again suggest that explicit instruction had the greatest effect on learner output. In addition to producing a greater number of target metaphors, experimental group learners also produced a greater variety of metaphorical word types. Metaphorical variety for untaught types was almost identical, as had been the case for overall production (Figure 4.4). Figure 4.5 displays the same data in box and whisker plots. The interquartile range for the taught metaphor types in the experimental condition is

Table 4.7. Example of calculation for target metaphor types to all types ratio

	Taught target metaphor forms	Types	Calculation
1	big	1	All types = 602
2	biggest	2	10 ÷ 602 = 1.66%
3	dropped	3	
4	explode	4	
5	free-spirited	5	
6	high	6	
7	high		
8	high		
9	hybrid	7	
10	losing	8	
11	losing		
12	losing		
13	snowballed	9	
14	snowballed		
15	steal	10	

Table 4.8. Median and mean target type ratios in the two conditions

Metaphor category	Median type ratio		Mean type ratio (SD)	
	Con.	Exp.	Con.	Exp.
Target	1.95%	2.53%	2.10% (0.73%)	2.60% (0.66%)
Untaught	1.61%	1.64%	1.61% (0.54%)	1.66% (0.58%)
Combined	3.93%	4.10%	3.71% (0.99%)	4.25% (0.86%)

Table 4.9. Mann-Whitney tests for target metaphor type ratios between the two conditions

Metaphor category	Mean ranks		*p*	*U*	*z*	*r*
	Con.	Exp.				
Ratio of taught target MRW types to all types	19.26	27.74	0.032#	167.000	-2.142	0.32
Ratio of untaught target MRW types to all types	23.13	23.87	0.852	256.000	-0.187	0.03
Ratio of combined target MRW types to all types	20.65	26.35	0.150	199.000	-1.439	0.21

Note: *n* = 46 for all tests
= Non-significant after Bonferroni adjustment

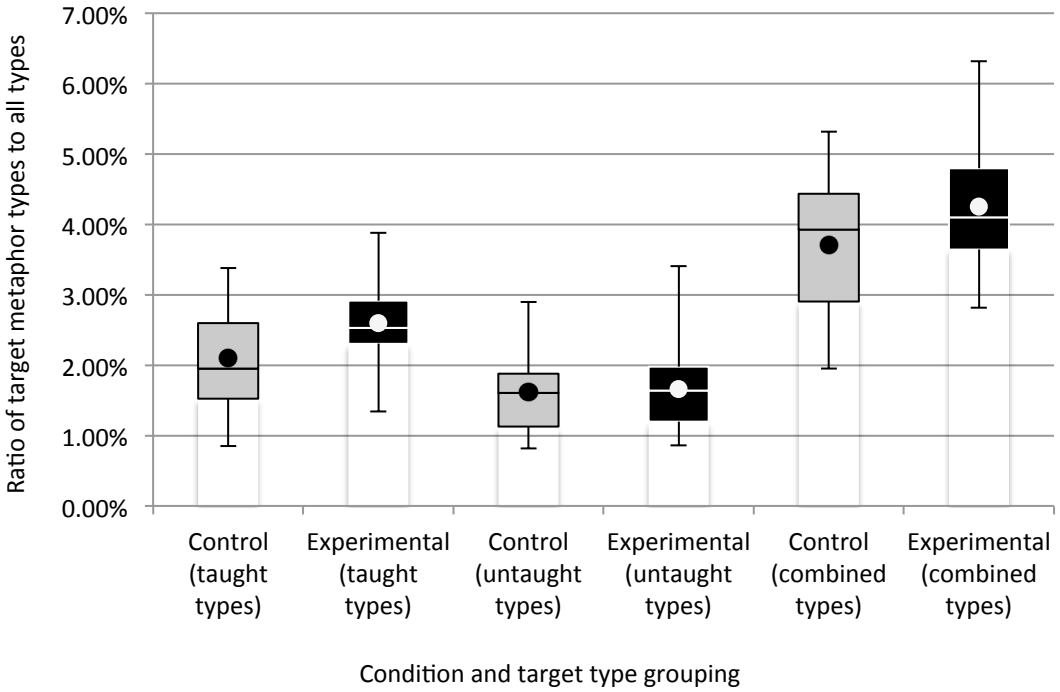


Figure 4.5. Distribution of target metaphor type ratios in the two conditions

narrower than that of the control group (Con. = 1.08%, Exp. = 0.61%[32]), suggesting that values have bunched more around the median. This point will be returned to in chapter five.

It appears then, that there is some evidence for the claim that raising learners' awareness of metaphor leads to increased production, both in terms of quantity and variety, of the particular forms taught in class. There is, however, little evidence that learners extend what they have been taught to produce other metaphorical language within each metaphor theme. Neither the quantity nor variety of untaught metaphor produced by the two groups differed appreciably. It should be noted, however, that lack of production does not mean learning has not taken place. The data in this study are only of written output, and there remains the possibility that receptive knowledge of untaught forms, may have improved. It is also possible that

[32] These values are the 25th percentile value subtracted from the 75th percentile value (i.e., from the bottom to the top of the box). See the description of box and whisker plots on page 133.

learners of higher ability than in the present study may be more able to extend metaphor knowledge appropriately, but there does appear to be merit in identifying key metaphorical patterns for explicit classroom attention. These findings are important for the growing number of programs that are adopting CLIL or English-medium approaches to language education, as learners in such programs will be confronted with genre-specific discourse norms that they will be expected to comprehend and produce.

4.6. The effect of topic on metaphor production

The degree to which the topic matter can influence metaphor production is an important point to consider in studies of language production. If it can be shown that topic does have a close relationship with metaphor use, even when learners are writing within the same broad theme, then future research must take great care to consider the prompts used to elicit responses, especially in longitudinal studies.

Mean metaphor rates were calculated for learners in each condition across the writing topics for all MRWs, open-class metaphors and both taught and untaught target metaphors. The results in Figures 4.6 - 4.9 show that metaphor rates fluctuated as the courses progressed, with similar peaks and troughs for certain topics between the two conditions. There is also some suggestion that the peaks were becoming more frequent towards the end of the course in the case of experimental condition target metaphors. Overall, the figures seem to suggest that topic does have some influence on metaphor production, and that learners may begin to use more target metaphors over time, though this will only occur after several weeks of instruction.

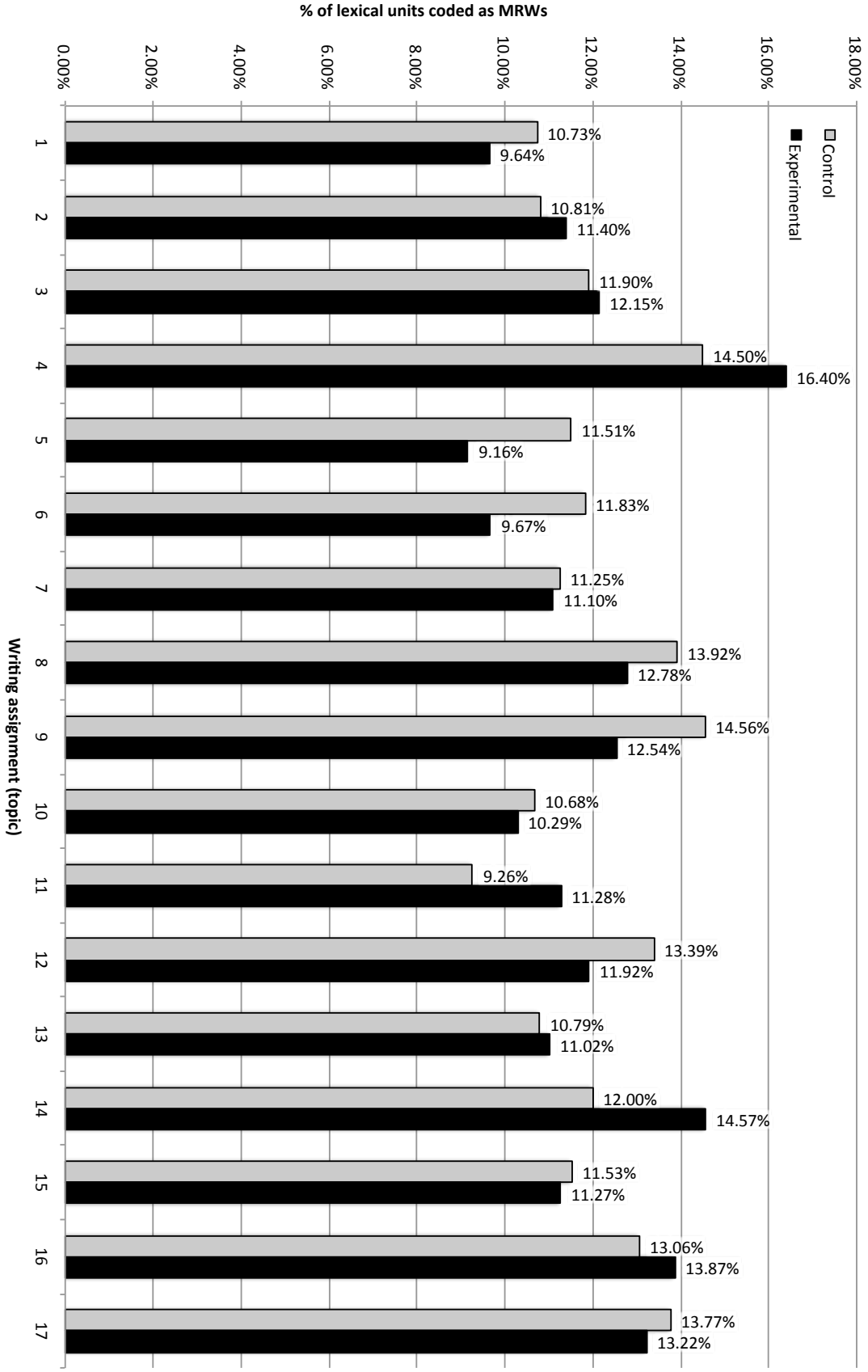


Figure 4.6. Mean metaphor rates per participant in each topic

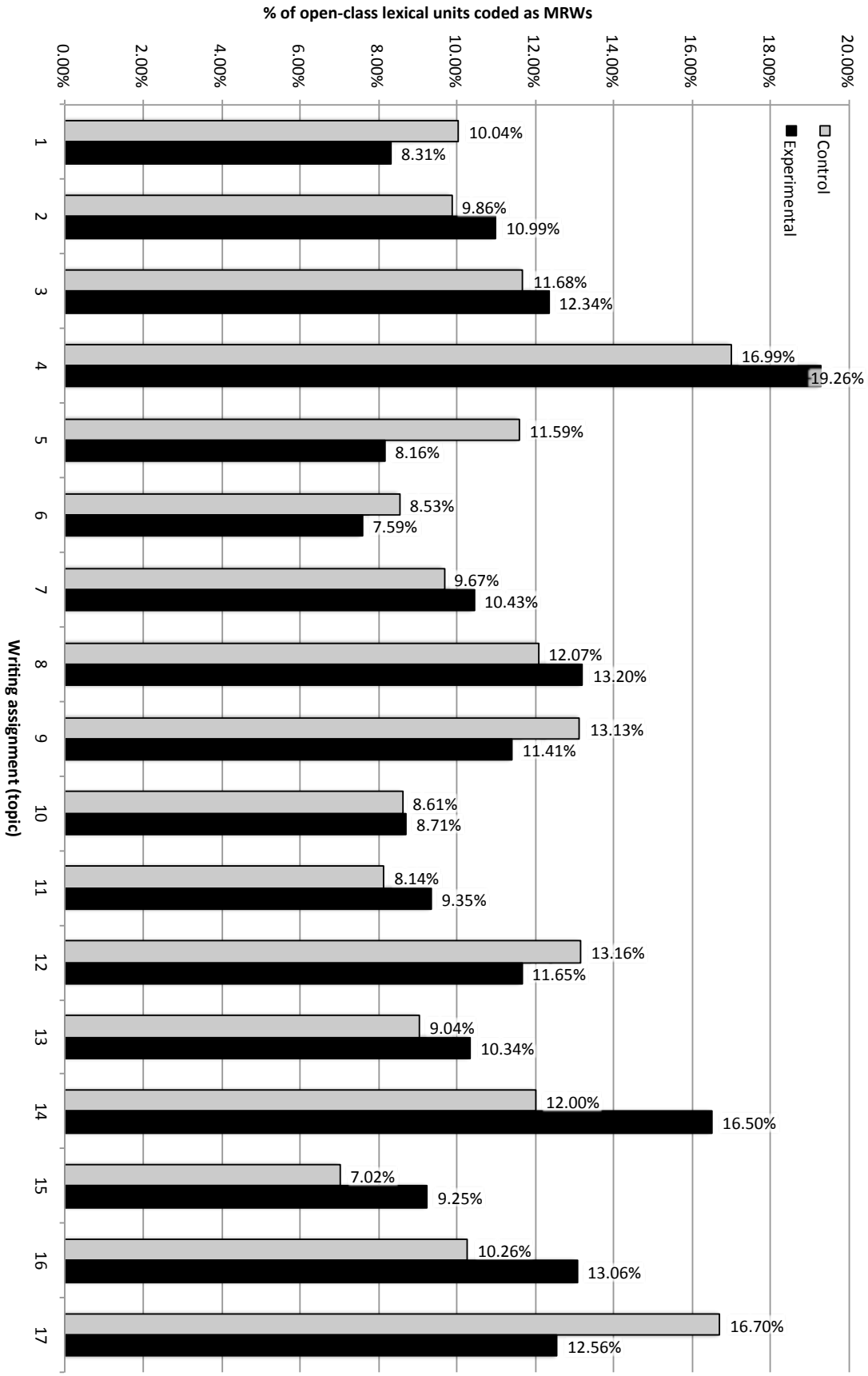


Figure 4.7. Mean open-class metaphor rates per participant in each topic

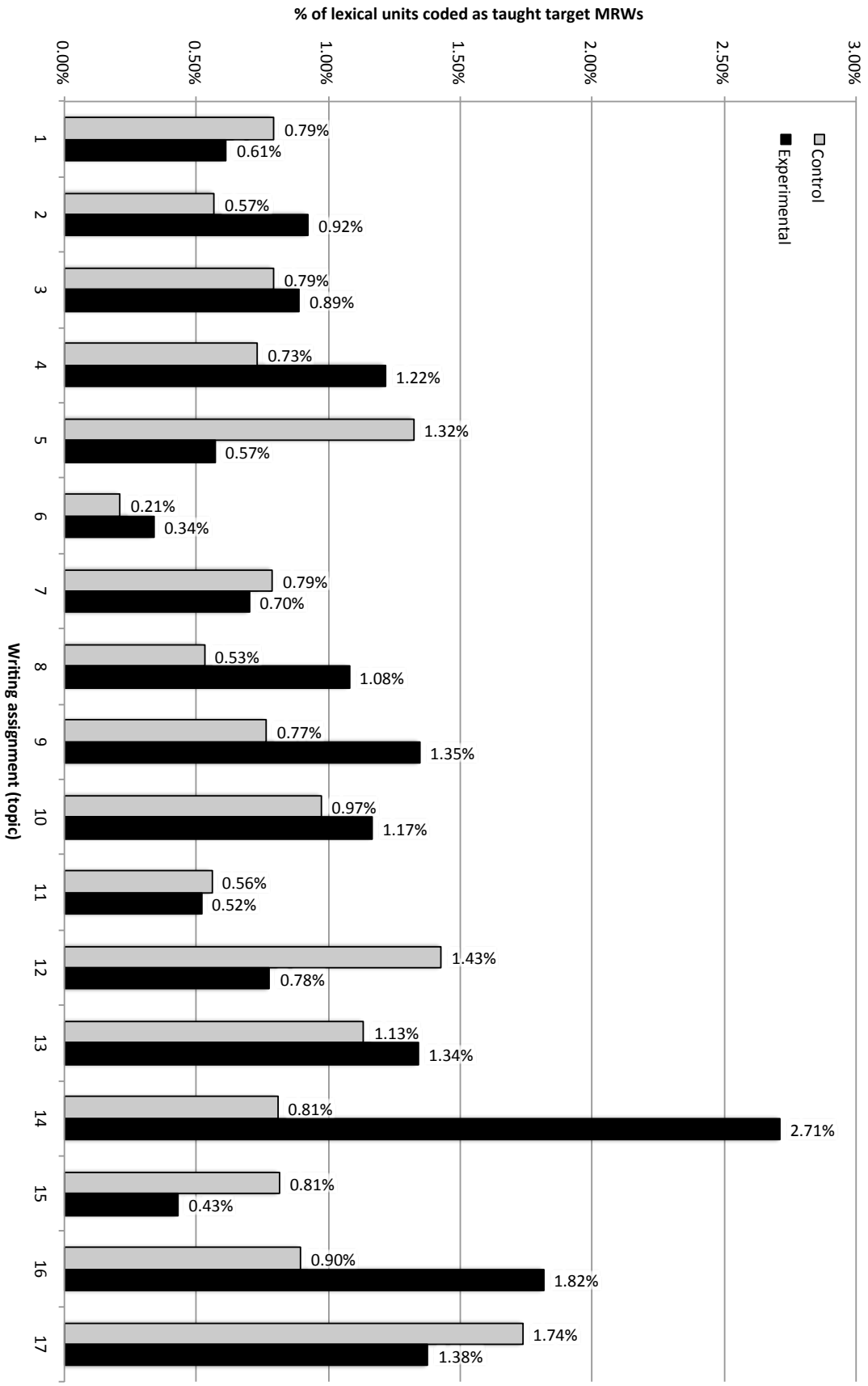


Figure 4.8. Mean taught target metaphor rates per participant in each topic

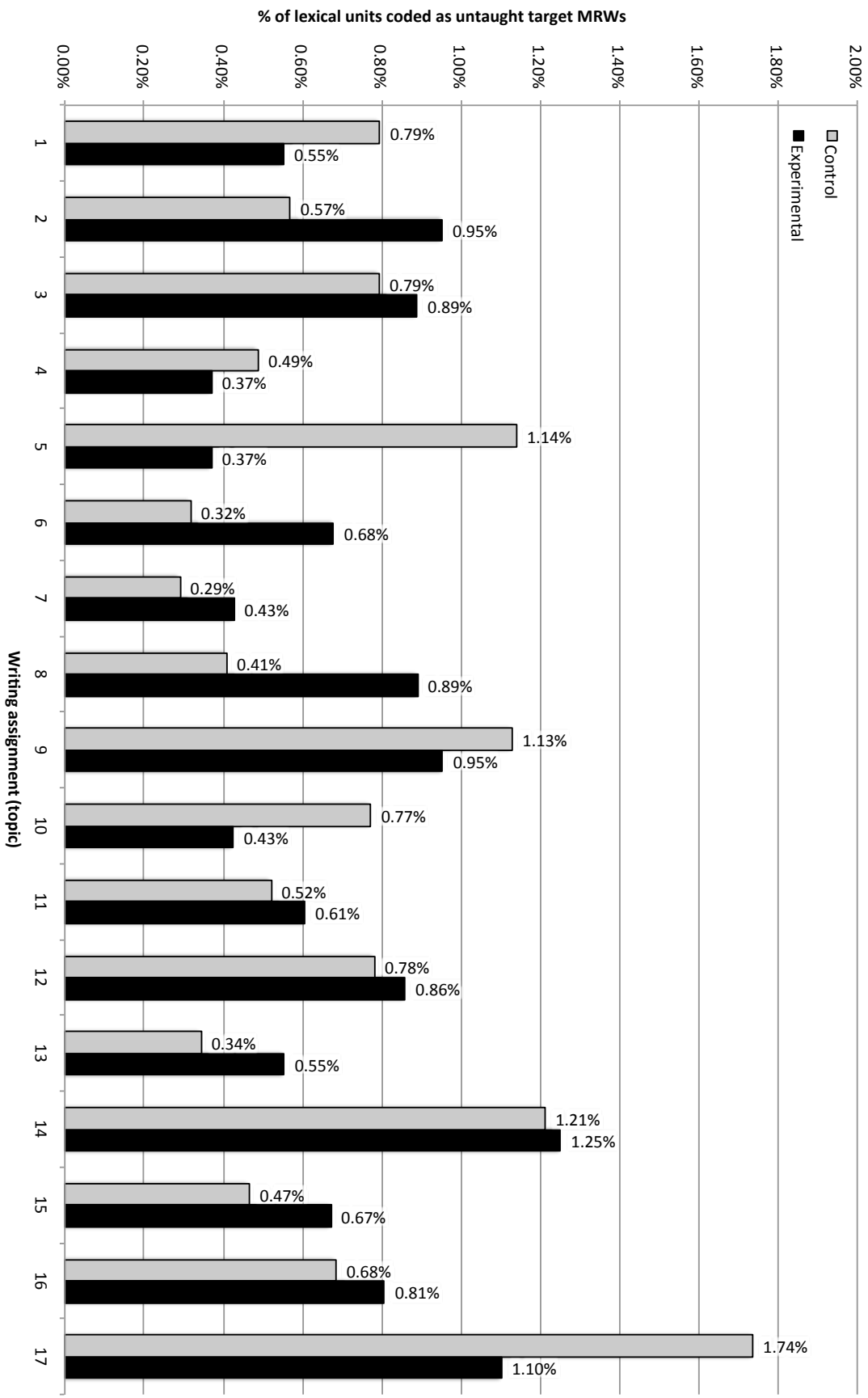


Figure 4.9. Mean untaught target metaphor rates per participant in each topic

To investigate the relationship between topic and metaphor production, Spearman's rho correlations were calculated for different categories of metaphor between the two conditions (Table 4.10). These showed a strong relationship for both the total amount of metaphor and for open-class metaphors across the topics in the two conditions. The correlations for the two target metaphor categories were more moderate and that of closed-class metaphors was weak. Even without explicit instruction in metaphorical language, the control group's total use of metaphor largely followed the same trend as the experimental group's writing, which suggests that topic is playing a part in learner production of metaphor. The strong relationship between open-class metaphor use across topics in the two conditions implies that lexical choice has a considerable influence on this.

To investigate this relationship a little more closely, the topics that were consistently high or low scoring for open-class metaphor were identified. The topics and writing prompts are shown in Table 4.11. What immediately stands out in the high metaphor group is that most of the prompts require learners to express opinions about Japanese culture, whereas in the low metaphor group the prompts are less explicit about inviting learners' views. Expressing opinions is a necessary element of argumentative essay writing, which is a writing task that has been shown

Table 4.10. Spearman correlations for learners' mean metaphor rates across topics

Metaphor type	r_s	*p*
% total MRWs	.733	.001**
% open-class MRWs	.725	.001**
% closed-class MRWs	.279	.277
% taught target MRWs	.294	.252
% untaught target MRWs	.382	.130

Note: *n* = 17 for all tests
** = Significant at 0.01 level

Table 4.11. Topics that produced high and low levels of metaphor use

Topic	Writing prompt	Details
(high metaphor rates)		
4	Give your opinions about one or more of the Japanese tribes. Can you identify with any of these groups or can you at least understand their opinions?	Expressing opinions on Japanese sub-cultures
8	What did you learn from Makiko's New World? What surprised you? What did you feel was good or positive about their lives? What did you think was not so good about their lives?	Expressing opinions on Meiji-era life
9	Discuss the popularity of Japanese versus foreign music with reference to newspaper article.	Describing popularity of music
14	Are the Japanese losing their culture? Is there too much American influence? Are the Japanese westernized or modernized?	Expressing opinions on cultural flow
17	Define 'hybridity' and 'hybridism' and give your opinion on why Japan might use hybridism to construct its identity in global society.	Expressing opinions on why Japan might try to preserve its image of being unchanged
(low metaphor rates)		
1	What is your personal image of Japanese popular culture? What does this phrase mean to you?	Providing personal impression of course theme
6	What are some similarities and differences between your life and Makiko Nakano's life?	Comparing students' own lives to Meiji era lives
10	Choose an element of Japanese culture that is a hybrid. Describe how it shows hybridity and whether it shows hybridism.	Describing a cultural entity and explaining how it has been adapted to suit Japanese culture
11	Compare sumo with other sports played in Japan. Makes comparisons in terms of traditions, history, popularity etc.	Comparing sumo with another sport
15	What element of Japanese Popular Culture would you choose to teach that was not discussed in this course?	Personal choice

to elicit large quantities of metaphor from high-proficiency learners (Littlemore et al., 2014; Turner, 2014). Such essays typically require learners to critique opposing views, use emphasis and evaluation, and to construct a coherent flow of ideas, all of which are likely to require metaphorical language.

The prompts for low metaphor topics required comparison, description and expressions of personal choice. While certainly offering scope for metaphor use, these modes may be more likely to evoke concrete than abstract referents, which would limit metaphorical output.

A further influence of topic on output is probably due to the wording of the prompts themselves. Three of the prompts that led to high metaphor production contain a metaphorical verb (*identify*, *lose*, and *construct*) which was frequently used by learners for that particular writing task. Out of the 15 total metaphorical uses of *identify* and its related forms in both output corpora, 13 appeared in topic four (87%). For *lose*, 42 out of 80 uses (53%) were in topic 14, while for *construct*, eight out of eight (100%) of uses appeared in topic 17. It is understandable that learners recycle vocabulary from writing prompts, as this is one way of showing a genuine response to the question. However, this does also highlight the influence of the writing prompt on learners' output.

4.7. Concluding comments

This chapter has compared the written output of two groups of learners: one that received explicit instruction in common metaphorical themes, and a control group whose instruction focussed on covering academic vocabulary that appeared in course input. Learners produced writing in an aided environment, with reference

and course materials both available, and the metaphors they used in doing so were then identified. MIPVU proved to be an effective method of analysing learner data, but as this is a highly inclusive identification procedure, the raw results are simply too broad to offer insights into learner development. Nevertheless, the data obtained offer a very useful starting point for more detailed further analysis. Development of metaphoric competence in classroom settings is a little-researched area of learner development, and several findings are worthy of further comment.

First, metaphor frequency rates for both groups were clearly higher than those for learners of comparable ability writing under examination conditions. It would be interesting to carry out further studies to examine the degree to which context, topic and use of reference materials contribute to this boost. In studies on learner metaphor production thus far (e.g., Littlemore et al, 2014; Turner, 2014), data obtained from examinations has been used as a proxy measure of learners' true ability, but caution should be exercised here, as such conditions are unlikely to form the majority of language learning experiences. Different task requirements may yield greater levels of metaphorical output, and obtaining measures of output when reference materials are available would refine our understanding of learner capabilities. Likewise, our understanding of the written metaphor production of native speakers is largely based on data sampled from the work of professionals with access to reference materials. Certainly, this would have been the case in most sections of the BNC which were analysed by Steen et al. (2010). We will not fully understand the range of performance that learners or native speakers are capable of until we can be sure that we are comparing like for like.

The data presented here also highlight the limitations of considering metaphor production only at the broadest level. The metaphor rates in Table 4.1 give the impression that there was little difference between the control and experimental conditions, but the subsequent analyses imply that change may have occurred at a deeper level. Language is a complex system, and evolution in one area of that system may appear alongside (or even cause), a diminishment in other areas.

Examining target metaphor production provided some evidence of a difference in the output of the two conditions. Learners in the experimental group were clearly able and willing to apply the target language that they were explicitly taught in their own reflective writing. Since the target features were selected as appropriate for the course content, this suggests that learner output in the experimental condition more closely approximated the norms of anthropological discourse on culture. However, one question worthy of further consideration is the degree to which explicit instruction limits experimentation with other language features. The finding in this case was that production of untaught target metaphors did not increase in response to the teaching of metaphoric themes. While it is a positive sign that untaught metaphor use was not adversely affected, it seems likely that learners prioritised the language that they studied in class rather than making efforts to draw on the metaphor themes in their own ways. Perhaps classroom activities that require learners to come up with their own expressions within a target metaphorical theme might act as a counterbalance to any tendency for explicit instruction to have a normalising effect on learner output.

It should also be borne in mind that in free-production activities, genre-specific metaphor is likely to comprise only a relatively small proportion of output. Even with

the rise in target metaphor production seen in this study, actual frequency of production rose only from around once every 115 words to once every 96 words on average, while non-target open-class metaphors appeared once every 21 to 22 words. Variation in these figures for other fields is possible of course, but in situations where non-advanced learners are studying in content-based programs, it is probably better that awareness raising of metaphor be seen as one element of well-balanced vocabulary instruction, rather than a substitute for vocabulary building.

Another claim related to target metaphor production is that greater lexical variety should be considered a marker of developments in interlanguage just as much as increased frequency of production. Key features of a powerful lexicon are variety and flexibility of use (Hasselgren, 1994; Nation, 2013, p. 50, 263), and learners who overcome their dependence on certain safe options are increasing their lexical, and in this case, metaphorical, fluency.

Finally, there is a suggestion that writing topic may be related to metaphorical output. Healthy correlations were found between the output of both conditions for the total amount of metaphor and open-class metaphors. Stronger claims than this, however, would require a study designed to separate out the effects of teaching and topic. Since the target metaphors for the experimental group here were based partially on the control group's output, there is a possibility that the teaching in the experimental condition exacerbated a tendency that was already present in the learners to produce certain metaphors, rather than the topic itself driving this trend. In any case, the results here imply that future studies on the development of

metaphoric competence should attempt to control for the effects of topic and teaching effects.

One of the clear limitations of the data presented in this chapter is that learners' ability has not been taken into account. While there were no significant differences between the two groups' overall English proficiency, there remains the possibility that the treatment affected learners of different abilities in varying ways. It is this question that will be addressed in the next chapter.

CHAPTER 5. METAPHOR PRODUCTION AND LEARNER ABILITY

5.1. Introduction and rationale for the study

Metaphor operates at all levels of language, from the prosaic to the inspired. As such, we would expect to find it in the output of learners of all abilities, and some studies have confirmed this (Littlemore et al., 2014; Turner, 2014). However, the question of how metaphor production develops in learners remains an under-explored area of research. This chapter will attempt to shed new light in this area.

The above studies used data obtained from examinations that are tied to CEFR descriptors of language proficiency. Although the CEFR makes very little explicit mention of metaphor itself, it does contain descriptions of the communicative functions and linguistic competencies learners at each band would be expected to display in their output. Table 5.1 lists the descriptors for learners in the A2 - B2 levels for written production. We can see that learners at the A2 level are expected only to produce texts describing their own immediate lives using limited connecting devices. At level B1, there is limited scope for engagement with the abstract world as hopes and dreams become possible subjects, and there is more expectation that cohesive text will be produced, but only when we reach the B2 level do we find topics that offer greater latitude for the production of metaphor. At this level, learners are expected to write on a wide range of topics in detail, offering evaluations and viewpoints, and either supporting or distancing themselves from arguments.

Table 5.1. CEFR descriptors for written production[33]

CEFR band	Category	Descriptor
A2	**Global scale**	Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	Overall written production	Can write a series of simple phrases and sentences linked with simple connectors like 'and', 'but' and 'because'.
	Reports and essays	No descriptor available
B1	**Global scale**	Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.
	Overall written production	Can write straightforward connected texts on a range of familiar subjects within his/her field of interest, by linking a series of shorter discrete elements into a linear sequence.
	Reports and essays	Can write short, simple essays on topics of interest. Can summarise, report and give his/her opinion about accumulated factual information on familiar routine and non-routine matters within his/her field with some confidence. Can write very brief reports to a standard conventionalised format, which pass on routine factual information and state reasons for actions.
B2	**Global scale**	Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	Overall written production	Can write clear, detailed texts on a variety of subjects related to his/her field of interest, synthesising and evaluating information and arguments from a number of sources.
	Reports and essays	Can write an essay or report which develops an argument systematically with appropriate highlighting of significant points and relevant supporting detail. Can evaluate different ideas or solutions to a problem. Can write an essay or report which develops an argument, giving reasons in support of or against a particular point of view and explaining the advantages and disadvantages of various options. Can synthesise information and arguments from a number of sources.

[33] Adapted from: Council of Europe, 2001, p. 26 - 27; p. 61-62

Such descriptions are valuable for sketching out the degree of textual and conceptual sophistication learners might be expected to demonstrate, but as a whole, the language used remains somewhat removed from the level of detail required to analyse lexical or metaphorical development within the CEFR framework. For this reason, more specific descriptors have been proposed for vocabulary size (Milton, 2010) as well as range and complexity of metaphor in language production (Littlemore et al., 2014, p. 141 - 142), as shown in Table 5.2.

Vocabulary size, or breadth, is an excellent predictor of overall proficiency, and knowledge of the most frequent 2,000 word families in particular has been shown to correlate strongly with performance on reading, listening and notably, writing examinations (Stæhr, 2008, p. 144 - 145). Metaphor itself is so pervasive that it can be found throughout the lexico-grammatical continuum, but as Table 5.2 suggests, for lower proficiency learners, it may appear primarily in grammaticalised forms such as prepositions. Only at more intermediate stages does open-class metaphor become more prevalent.

If we attempt to draw all of these findings together, we might speculate that while some learners at the A2 level possess sufficient vocabulary to perform well on written tasks, for the most part their output will be characterised by simple constructions, with metaphor mainly restricted to closed-class forms. When they reach B1, learners should all possess vocabularies that allow for self-expression on topics both within and outside their realm of interest. Metaphor will become more prominent in the output of such learners, though still chiefly in conventionalised expressions. By the B2 stage, learners' vocabularies will be rich enough for them to

be able to deploy metaphor for evaluative or text structuring purposes and in both complex and creative constructions.

Table 5.2. Proposed CEFR descriptors for metaphor and vocabulary[34]

CEFR band	Category	Proposed descriptor
A2	**Metaphor**	Learners should be able to make use of a limited range of metaphorical prepositions.
	Vocabulary size	1500 - 2500 word families
B1	**Metaphor**	Learners should be able to use a limited number of conventional metaphors, with appropriate phraseology to present their own perspective. They should also be able to make limited use of personification metaphors. They may be starting to use a small number of metaphor clusters.
	Vocabulary size	2750 - 3250 word families
B2	**Metaphor**	Learners should be able to make use of a limited number of conventional and creative open-class metaphors. They should be able to use metaphors for evaluative and discourse organising purposes. They should be starting to use personification metaphors more extensively. Metaphorical clusters are more in evidence at this level. Some are coherent, whereas others contain mixed metaphors.
	Vocabulary size	3250 - 3750 word families

One limitation to studies that have drawn data from examination papers is that it is difficult to separate evidence of increased metaphorical production from the effects of the writing task and context. Learners writing at the A2 or B1 levels are typically asked to produce shorter pieces of writing such as letters or postcards, while from the B2 level, there is a shift towards lengthier compositions that require discussion of a global issue. The latter tasks conceivably offer more scope for

[34] Adapted from: Littlemore et al., 2014, p. 141-142; Milton, 2010, p. 224

learners to engage with abstract topics and use rhetoric to justify their views, and it may not be a surprise then that they also use more metaphor to do so.

Studies on learner writing have accounted for the effect of topic by either limiting their sampling to particular topics (Nacey, 2013, p. 132 - 134; Turner, 2014, p. 60 - 66) or by identifying certain keywords that could be searched for in a larger corpus in order to extract writing on similar issues (Littlemore et al., 2014, p. 120 - 121). One intention of this study is to complement other investigations into learner development by controlling for the effect of writing topic; learners were all given the exact same prompts in both conditions. It must be conceded, however, that controlling for topic comes at the expense of a loss of control for proficiency. As will be shown in Table 5.3, learner ability was not evenly distributed within each condition.

A further difference to consider is the context in which writing takes place. Studies based on the Cambridge Learner Corpus (Littlemore et al., 2014; Turner, 2014) drew on data from examination conditions, in which dictionary use is not permitted[35]. Rather than writing under examination conditions, learners in this study were responding to course content with access to reference materials and while being encouraged to experiment with new language. For the control group, instruction centred on more traditional teaching of academic and general purpose vocabulary. As the input corpus confirmed, metaphor was certainly present in the classroom, but no attempt was made by the instructors to teach this systematically. The experimental condition switched the emphasis from regular vocabulary

[35] Based on the list of frequently asked questions for the Cambridge English examinations (http://www.cambridgeopencentre.org/asp_pages/faqs.asp)

instruction to awareness raising of metaphor. As such, the control condition to some extent reflects implicit learning of metaphor, while the experimental group was given explicit instruction. Three limitations should be noted here, however. The first is that the lack of a pre-test means we cannot be sure that learners were not already aware of expressions prior to the commencement of the course. The second is that lack of production cannot be taken as evidence of lack of learning. Clearly, it is possible that learners may have acquired a receptive understanding of particular expressions without having actually produced them in their writing. Finally, as has already been noted, the study cannot make claims about actual acquisition since reference materials were available. Instead, the focus will be on the quantity, diversity and quality of metaphor that learners produce in their writing in either condition.

A brief examination of data from chapter four will highlight the degree of individual variation in metaphor production and hints at the effect that the experimental treatment had on output. Figure 5.1 shows the percentage of open-class lexical units that were classed as metaphorical in the two conditions. This is the same data for open-class MRWs as provided in Figure 4.2 (page 135). The standard deviation measures indicate that the overall range of values has become somewhat narrower than that of the control group (Con. *SD* = 2.17%; Exp. *SD* = 1.73%), suggesting the data may have become more 'bunched'. Whether this is due to the treatment having a differential effect on learners of varying proficiencies is one question that will be considered.

This chapter expands on the findings of previous studies into learner metaphor production by considering the effects of classroom activities on the writing of

learners of different proficiencies. The use of classroom data as opposed to samples from larger corpora also allows for a more in-depth analysis of learner output, since expressions can be traced back to the classroom context, and the effects of implicit or explicit instruction considered.

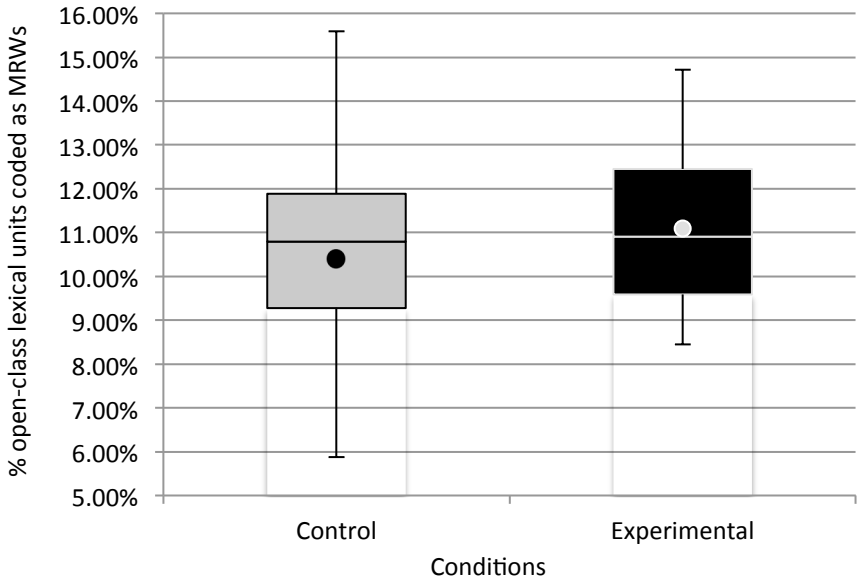


Figure 5.1. Distribution of open-class metaphor rates in the two conditions

Furthermore, it was claimed in chapter four that the raw MIPVU results might be too inclusive to properly understand developments within either condition or the effect of the experimental treatment, so this chapter will narrow the focus of the study by introducing several other variables. Firstly, by focusing on the learners themselves, the relationship between developments in metaphor production and overall proficiency can be examined. This is important to consider, as it is recognised that language development is an incremental process, and learners of different abilities may well respond differently to instruction. While retaining this attention to learner ability, the focus will then turn to aspects of language by considering open- and closed-class parts of speech separately in order to examine how the lexical and

grammatical systems respond to metaphor awareness raising. It is also important to consider the effect of explicit teaching, as the choice of target forms will have ramifications for other areas of learners' interlanguage, and learners themselves may respond differentially to particular target forms. Finally, skilful language users are able to harness metaphor to meet varying pragmatic goals, so there is a need to examine the rhetorical purposes that learners at different stages of development are able to use metaphor to fulfil. The specific research questions in this chapter are:

a) For both the control and experimental groups, will differences in overall language ability correspond to variations in metaphor use in written output and overall language ability?
b) Will learners of different levels of proficiency exhibit differences in open-class POS metaphor production in either condition?
c) Will learners of different levels of proficiency exhibit differences in target metaphor production in either condition?
d) Will learners of different levels of proficiency exhibit differences in use of metaphor for rhetorical purposes in either condition?

5.2. Methodology

5.2.1. Data sources and treatment

The learner output corpus data used in this section are the same as that for chapter four. That is, spelling mistakes were corrected in order to aid corpus analysis, but the original versions are reported here.

The main goal of this chapter is to consider how learners' written production of metaphor varies with overall language proficiency. Proficiency was operationalized as participants' maximum scores on the TOEIC test[36]. As was explained in section 3.6.1 (page 100), TOEIC is a widely-used instrument in Japan, and one that learners could therefore be expected to be familiar with. The TOEIC to CEFR conversion scale developed by Tannenbaum and Wylie (2006, 2008, p. 29) was used to obtain approximations of the CEFR bands which most likely describe participants' abilities. These CEFR estimates were used to draw comparisons with other studies that have considered learner metaphor production over a range of abilities.

One problem with the TOEIC and CEFR ability measures obtained, as Table 5.3 shows, is that while the two conditions had broadly similar distributions of learner ability, the distributions within each condition were far from equal. For this reason, participants were divided into three almost equal groups based on their TOEIC scores. Those in the higher group had scores that were on or higher than the mean of the population of Japanese TOEIC test takers in 2013 and 2014, those in the mid group had scores that were within half a standard deviation of the mean, and those in the low level group had scores more than half a standard deviation below the population mean (ETS, 2014, p. 5; 2015, p. 5).

The final section of this chapter turns to the question of how learners put metaphorical language to use. Littlemore and Low (2006b) demonstrate how metaphor can be found in all of the major competencies in Bachman's (1990, p. 84 - 109) definition of language competence: grammatical, textual, illocutionary and sociolinguistic. As such, it is no surprise that metaphor is found in learner writing

[36] Scores were obtained from the paper-based test of listening and reading skills.

from all proficiency levels. It can be seen in Table 5.4 that learners writing at higher CEFR levels incorporate a greater variety of metaphor in their output. Presumably, increased control of grammatical and lexical patterns allows learners to begin using metaphor for more complex purposes, including language to trigger vivid emotions or humorous responses and genuinely creative formations.

Table 5.3. Participants' TOEIC and CEFR groupings

	Grouping	Control	Experimental
TOEIC score	<400	6	4
	400-<500	9	9
	500-<600	4	6
	600-<700	2	2
	700+	2	2
CEFR band	A2	18	17
	B1	4	4
	B2	1	2
Ability group	TOEIC: low	7	8
	TOEIC: mid	8	7
	TOEIC: high	8	8

Table 5.4. Developments in metaphorical competence[37]

CEFR band	Function of metaphor
A2	Grammatical metaphors and fixed expressions
B1	Metaphors to express opinions; personification metaphors
B2	Persuasive and evaluative use of metaphor; some creative metaphor; open-class metaphor use overtakes closed-class metaphor use; discourse organising metaphors
C1	Phraseology becomes more appropriate; greater use of metaphors in clusters; emotive metaphor use; some direct metaphor appears
C2	Phraseologically correct usage; creative usage in both direct and indirect metaphor; metaphors to express humour and sarcasm; extended metaphors

[37] Adapted from: Littlemore et al. 2014, p. 141-142

Turner (2014, p. 285 - 291) drew on Bachman's framework for language competence to create a coding scheme for learner metaphor production that captured most of the elements of Table 5.4. This scheme contained five broad categories: interpersonal, evaluative, textual, creative/imaginative and unmarked conventional. Studies into learner use of metaphor at various proficiency levels have typically found little evidence of creative or imaginative use before the C2 level of the CEFR scale (Littlemore et al., 2014, p. 135; Turner, 2014, p. 291 - 301), so this category was not adopted. The other categories, however, were felt to be broadly in line with language produced in the learners' writing. Accordingly, an adapted version of Turner's scheme was used, with the descriptions and examples of each category as applied in this study outlined in Table 5.5. The definitions listed in the table were used to help classify metaphors that could potentially belong to more than one category, and the data sorting function of Excel allowed for previous classifications to be listed, which helped to maintain consistency in the coding.

Table 5.5. Rhetorical function coding scheme for metaphor used in this study[38]

Category	Definition	Example
Interpersonal		
a. Advising	Metaphor that recommended some course of action	*We should hand down to Japanese culture*. [CON3]
b. Manipulative	Metaphor that attempted to persuade the reader to accept an opinion	*Clearly, "Neets" are criticized by society*. [EXP3]
Evaluative		
a. Emotive	Metaphor that served to raise the emotional quality of a particular statement	*I would teach about Japanese nature because it stands on the edge of a precipice now*. [CON16]
b. Emphasis	Metaphor that emphasised a particular statement, without adding an emotional quality	*A cultural import that had a powerful effect was electric light*. [EXP13]
c. Mitigation	Metaphor that served to diminish a concept or weaken a statement	*From the past in Japan, female was lower position in the family and tended to work only inside house*. [CON14]
d. General evaluative	Metaphor that expressed an evaluation without emphasis or emotion	*I think that Makiko's diary is very valuable.* [EXP16]
Textual		
a. Structuring	Metaphor that reinforced the coherence of a text by relating ideas with discourse markers and reference	*That is to say, It is too difficult for Japanese musician to succeed in foreign country.* [EXP15]
b. Reiteration	Metaphor that added to textual coherence by restating metaphors using synonyms	*They wanted to be new and tell the other people that they are different from people before, so "gothic lolitas" started. Some radical followers hurt themselves to make sure that they are alive*. [EXP8]
Unmarked conventional	Metaphor that served an ideational function without offering evaluation, typically through conventional language patterns	*Finally, Sumo demonstrates clear links to other elements of Japanese culture*. [CON12]

5.2.2. Application of MIPVU

In sections 5.3 to 5.5 of this chapter, MIPVU coding followed the same procedure as in chapter four, detailed in section 4.2.2 (page 129).

In section 5.6 and all subsequent sections, the MIPVU system of identifying lexical units was modified to take account of the frequent use by participants of

[38] Adapted from: Turner 2014, p. 288 - 289

phrases containing metaphorical words that were not included on the BNC polyword list. Many of these phrases were used to structure text, and since section 5.6 dealt with the use of metaphor to perform rhetorical functions including textual structuring, it was decided to treat such phrases as single units. Accordingly, phrases in the output corpora data that had been given their own separate entry in either the Macmillan or Longman dictionaries were now treated as single lexical units. If such units contained a metaphorically-used word, they were tagged as MRWs. Thus, while in the original format for MIPVU coding the phrase *on the other hand* would be listed as four lexical units, two of which are metaphorical, this system would record the phrase as a single metaphorical lexical unit. The list of phrases treated as single lexical units is provided in Appendix D (page 433), and the data from this chapter can be found in Appendix J (page 459) and Appendix K (page 473).

5.3. Language ability and metaphor production

Figures 5.2 - 5.4 show the total metaphor rates for participants grouped by the different ability measures. When grouped by approximate CEFR bands, learners in the control condition displayed a similar pattern of metaphor production to other studies that have examined metaphor in learner writing at a range of proficiency levels (Littlemore et al., 2014, p. 125 - 126; Turner, 2014, p. 84). That is, a rapid increase in metaphor production at the lower levels is followed by a plateau, in this case, at the B2 level. The experimental condition, however, displayed a remarkably different pattern of scores. While A2-level learners, or those with TOEIC scores less than 500, produced more metaphor than their control group equivalents, B1-level learners' production actually declined. In the B2 range, there was very little

difference between the two groups' scores. It should be noted that less than half of the learners fell into these higher bands, but the result is surprising nonetheless.

Figure 5.4, which divides each condition into more equally-sized groups, shows that metaphor production generally increases with overall proficiency in the control condition, albeit with some variation in values at an individual level. The experimental condition plots, meanwhile, are much less consistent. The lower ability group appears to have responded positively to the metaphor awareness raising, but the position of the median line in the interquartile box for the mid group indicates that mean values are clearly skewed by a number of high scores. The high group shows a lower level of metaphor use relative to the control, but the range of scores is narrower than the experimental group mid-level, and the median value is actually higher than that group. This seems to suggest that proficiency does play a role in the amount of metaphor that learners produce, particularly under 'normal' circumstances when metaphor is not a prominent course goal. However, the experimental treatment appears to have introduced an extra form of variability into the data, perhaps because participants responded to instruction in differing ways. We should also bear in mind that the inclusive nature of the MIPVU means that the data here will include errors, compensation strategies and simple overuse, so some overestimation may be taking place. The median and mean metaphor rates for all ability groups in both conditions are given in Table 5.6.

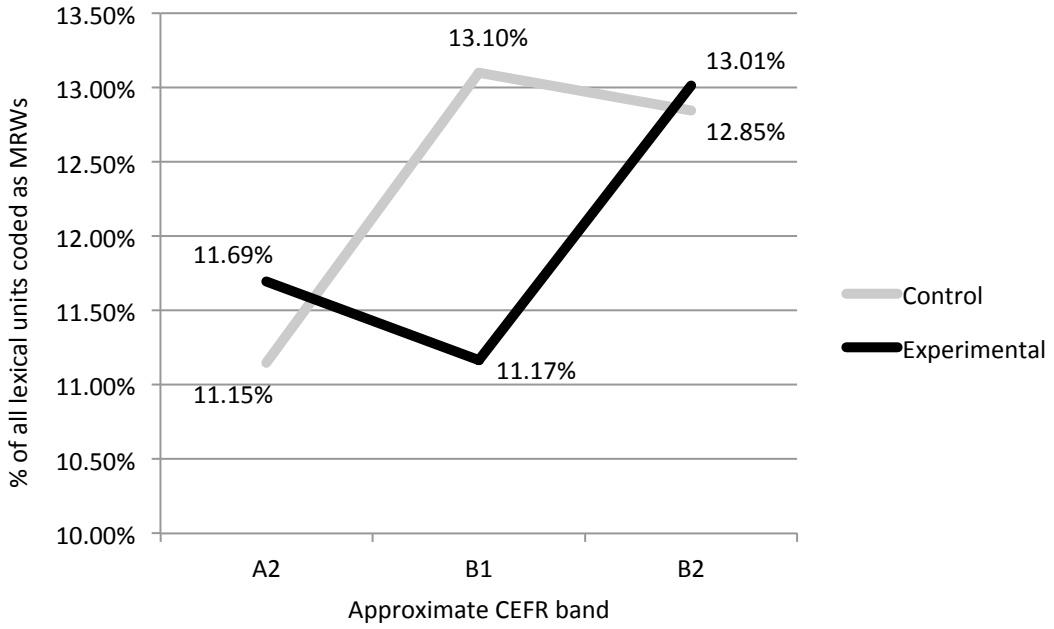


Figure 5.2. Total metaphor rates across approximated CEFR levels in the two conditions

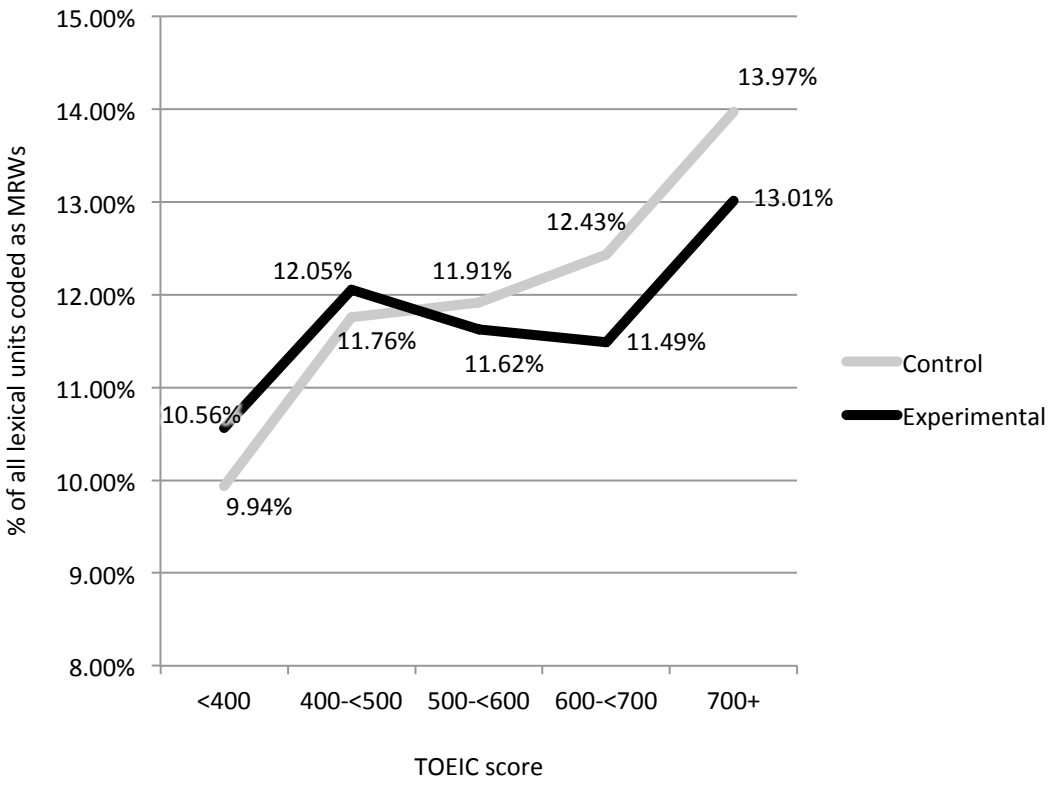


Figure 5.3. Total metaphor rates across TOEIC scores in the two conditions

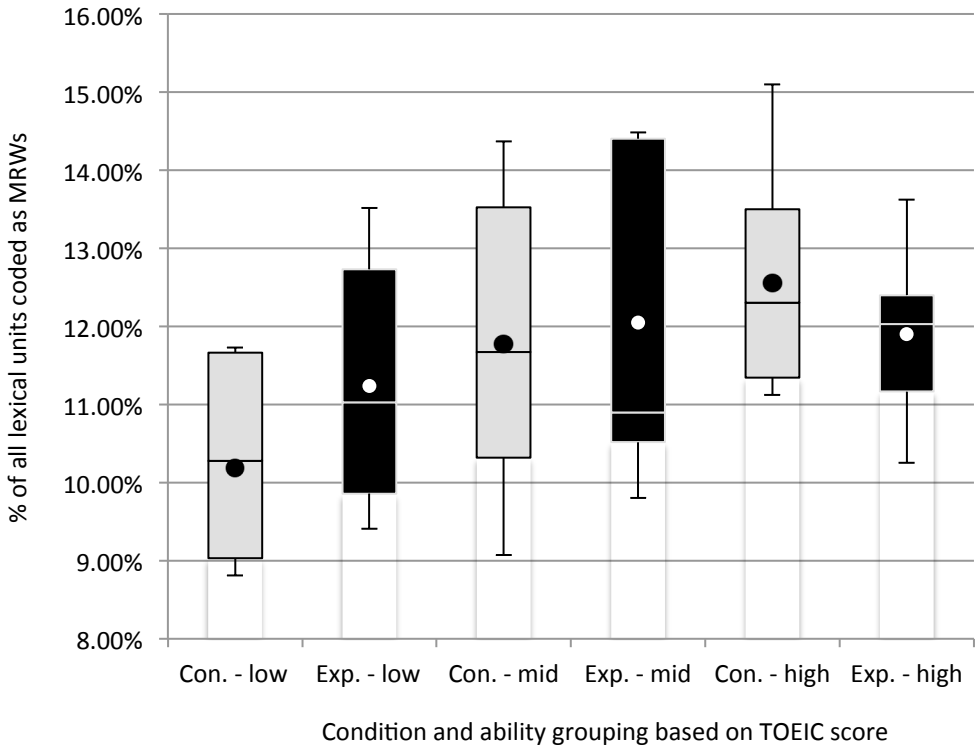


Figure 5.4. Distribution of total metaphor rates across TOEIC ability groupings in the two conditions

Table 5.6. Median and mean total metaphor rates for TOEIC ability groups in the two conditions

Ability group	Median total metaphor rates		Mean total metaphor rates (SD)	
	Con.	Exp.	Con.	Exp.
Low	10.28%	11.03%	10.19% (1.19%)	11.24% (1.53%)
Mid	11.67%	10.90%	11.77% (1.80%)	12.05% (2.05%)
High	12.30%	12.03%	12.56% (1.34%)	11.90% (1.01%)

Table 5.7. Kruskal-Wallis tests for total metaphor rates across TOEIC ability groups within conditions

Condition	Mean ranks			Chi square	*df*	*p*
	Low	Mid	High			
Control	6.43	12.88	16.00	7.639	2	.022*
Experimental	9.50	13.29	13.38	1.667	2	.434

Note: *n* = 23 for both tests
* = Significant at the 0.05 level

Table 5.8. Mann-Whitney tests for total metaphor rates across TOEIC ability groups within conditions

Group	Control condition					Experimental condition				
	n	*U*	*z*	*p*	*r*	*n*	*U*	*z*	*p*	*r*
Low-mid	15	13.000	-1.736	.083	0.45	15	18.000	-1.157	.247	0.30++
Low-high	15	4.000	-2.777	.005*	0.72	16	22.000	-1.050	.294	0.26+
Mid-high	16	24.000	-.840	.401	0.21	15	27.000	-0.116	.908	0.03

* = Significant at the 0.05 level
+ = Small effect size; ++ = Medium effect size

Table 5.9. Mann-Whitney tests for total metaphor rates across TOEIC ability groups between conditions

Groups	n	U	z	p	r
Low	15	15.000	-1.504	.132	0.39++
Mid	15	26.000	-.231	.817	0.06
High	16	21.000	-1.155	.248	0.29+

+ = Small effect size; ++ = Medium effect size

Kruskal-Wallis tests showed that there was a statistically significant difference in metaphor production within the TOEIC groupings for the control condition (Table 5.7). Post-hoc pairwise analysis with the Independent-samples Mann-Whitney test revealed a Bonferroni-adjusted significant difference between the low and high ability learners (Table 5.8). There were no significant differences between the two conditions at any TOEIC ability level grouping (Table 5.9).

With small sample sizes and no significant differences between the conditions, any claims based on these data have to be considered speculative. However, by looking at the trends presented here and considering the findings of other research into metaphor production, some tentative hypotheses might be proposed.

Firstly, the data here give the impression that the experimental treatment had a greater effect on learners of lower proficiency. This seems plausible, as those are the

learners who have most to gain in this area of competency. The moderate effect size for the low ability groups in Table 5.9 offers encouragement that a study with larger sample sizes may produce more conclusive evidence for this. However, the apparent drop in metaphor produced by higher-level learners in the experimental condition both when compared to the mid-level experimental group and the high-level learners in the control condition is curious. This may, of course, be no more than a quirk of this particular sample of learners, but another possibility is that these learners are in a different developmental stage than their classmates.

Littlemore et al. (2014, p. 132 - 133) identified B2 as the level at which learners began to diversify their range of metaphor use, rather than significantly increasing the actual quantity of metaphor produced. It was also noted that learners at this level began to experiment with metaphor much more, and consequently, made a greater amount of errors. Figures 5.2 and 5.3 actually indicate that the drop in metaphor in this study occurs a little earlier, in the 500-700 TOEIC range, or around CEFR level B1, although more proficient learners also produced less metaphor than their control group counterparts. One possibility then, is that the experimental group learners, writing in an aided environment without the pressure of high-stakes assessment, were able to demonstrate this experimental tendency one CEFR level earlier than the learners in the Littlemore et al. study, who were writing under examination conditions. Control group learners of a similar proficiency also appear to experience a slowdown in metaphor production, though not one as dramatic as in the experimental condition. Again, this has to be considered conjectural, but it offers a further point to consider in the following sections of this chapter.

5.4. Learner ability and open- or closed-class metaphor production

The findings in section 4.4 (page 133) showed a small and non-significant increase in open-class metaphor production in the experimental condition, with Figure 4.3 (page 138) seeming to indicate that greater use of open-class metaphor may be a marker of increased proficiency, with the caveat that writing topics vary in the degree to which they invite metaphorical expression. By drawing data from classroom settings, this study is able to control for the effect of topic by having learners respond to exactly the same prompts. Furthermore, the results in the previous section may indicate a differential response to the experimental treatment by learners of various proficiencies, so the rate of open- and closed-class metaphor production at different ability levels warrants further investigation.

In studies of learner metaphor production under examination conditions, it has been seen that fewer open- than closed-class metaphors appear in the work of lower proficiency learners, but relatively quickly, open-class metaphor becomes the most frequent type. This seems to occur somewhere around the B2 level for German, Greek, and Japanese learners of English, and possibly a little earlier for speakers of French (Littlemore et al., 2014, p. 127 - 128; Turner, 2014, p. 111 - 112). In order to observe developments in open- and closed-class metaphor production with proficiency, Figures 5.5 - 5.8 were plotted.

The various groupings of the four charts offer slightly different perspectives, but perhaps Figure 5.6 provides the clearest picture of development. It is noteworthy that in each condition, open- and closed-class metaphors appear to follow a similar trend. Broadly speaking, both open- and closed-class metaphor use follows an s-

shaped development curve, rising until TOEIC ≈ 500, then experiencing a dip, before beginning to climb again. There is also the possibility that this dip is more pronounced and long-lasting for closed-class metaphor production.

The studies included in Figure 4.3 suggested that at some point, open-class metaphor production overtakes that of closed-class use, with this taking place somewhere in the B1 to C1 bands. The figures below are less conclusive on this point, however. The OC and CC rates for the CEFR groupings do not cross within the A2 to B2 range (Figure 5.5). In the case of TOEIC scores, the control group values only show a crossing at the very highest level, for learners with scores above 700 points, and the experimental group values cross earlier, in the 600 to 700 range, but then re-cross (so that the CC rate exceeds OC) in the highest level (Figure 5.6).

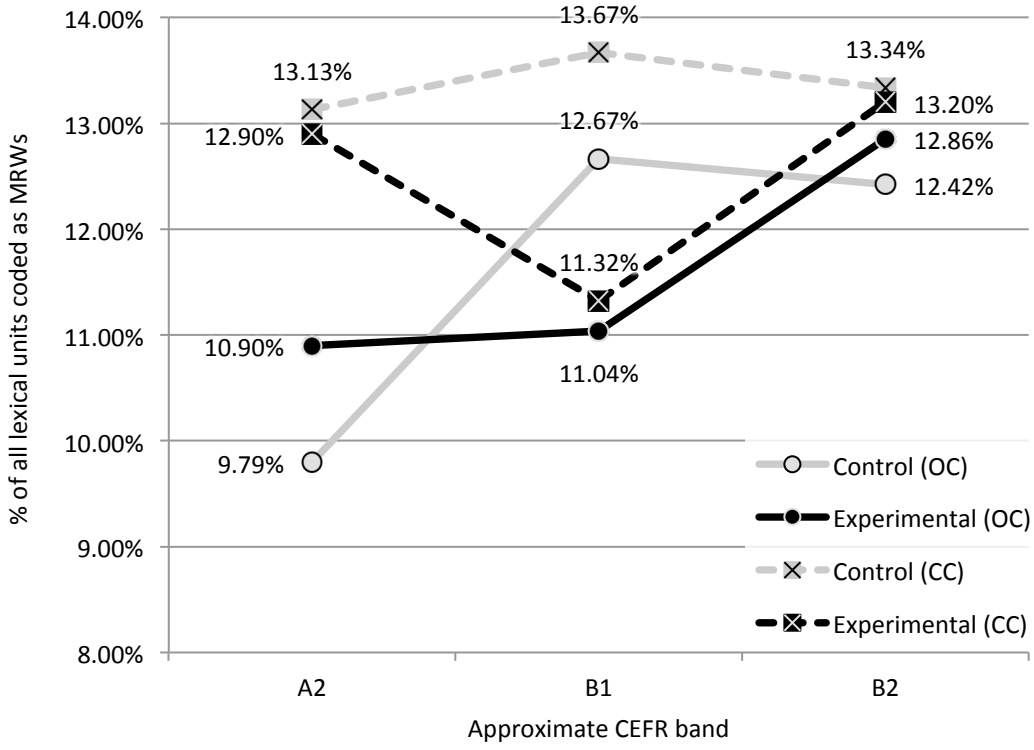


Figure 5.5. Open- and closed class metaphor rates across approximated CEFR levels in the two conditions

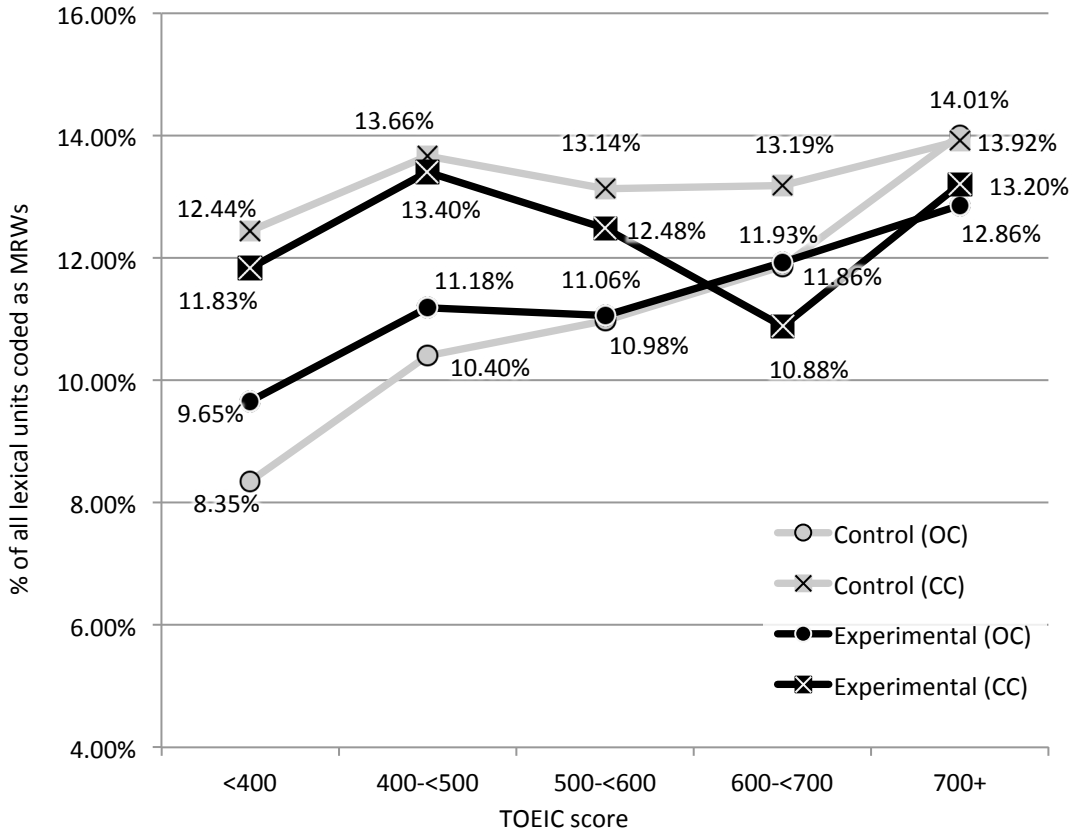


Figure 5.6. Open- and closed-class metaphor rates across TOEIC scores in the two conditions

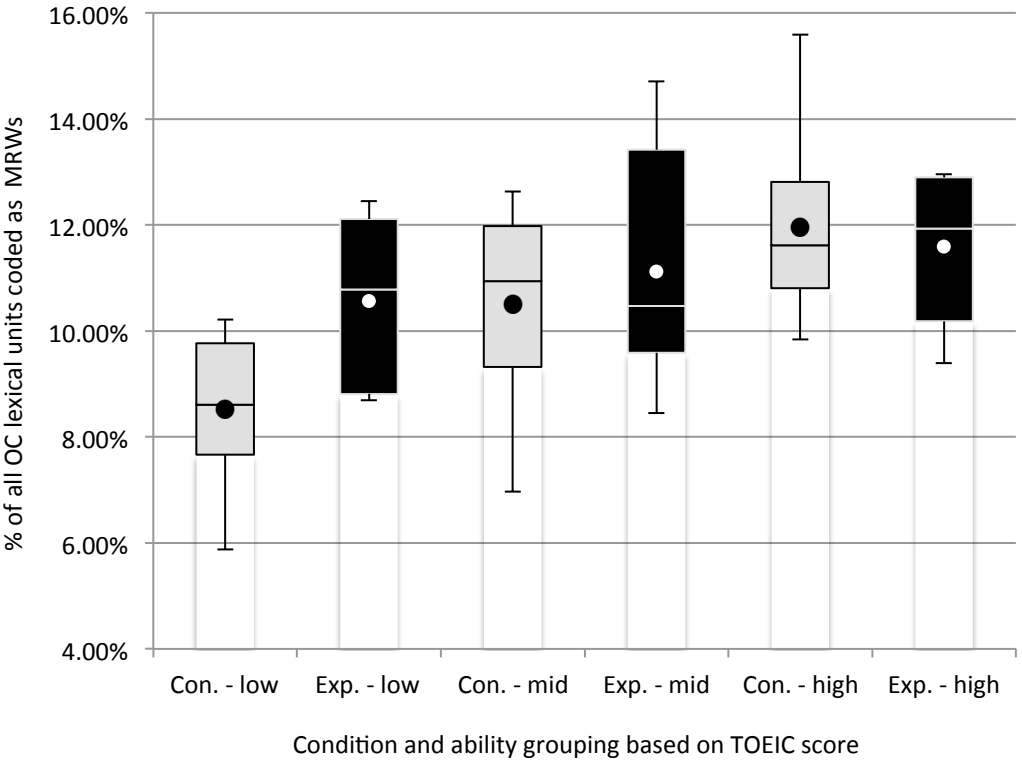


Figure 5.7. Distribution of open-class metaphor rates across TOEIC ability groups in the two conditions

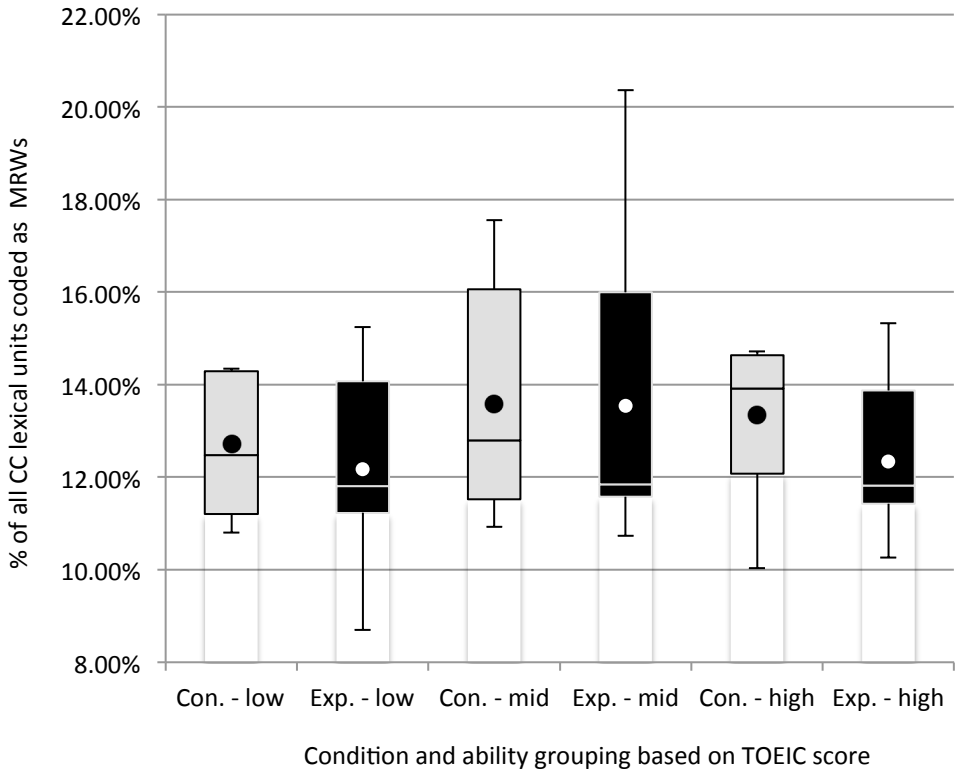


Figure 5.8. Distribution of closed-class metaphor rates across TOEIC ability groups in the two conditions

As a further step in understanding developments in metaphor production, metaphor rates for open- and closed-class lexical units produced by learners in the low, mid and high TOEIC ability groups (Table 5.10) were tested for significant differences both within and between the two conditions. For open-class metaphor, a significant difference was found in the control condition, and post-hoc analysis showed this was between the low and high groupings (Tables 5.11 - 5.12). Thus it can be claimed that there is some evidence for open-class metaphor usage increasing with proficiency under regular classroom conditions with no attempt to draw attention to metaphor. No significant differences were found for OC metaphor rates in the experimental condition or for CC metaphor rates in either condition (Tables 5.13 - 5.14). The analysis of results between equal ability TOEIC groups in each condition found a large effect size for the difference in OC metaphor

production between low-level group learners in the two conditions (Table 5.15), although with the Bonferroni adjustment applied, this was non-significant.

In summary, the trend that appears in the control group's writing is much less evident for the experimental condition. The broad conclusion from Figures 5.5 - 5.7 is that the experimental treatment had most effect on lower proficiency learners, and that this effect declined with higher ability learners. The results in Table 5.15 appear to confirm this, with the only strong effect size between the two conditions being in the low ability learners' use of open-class metaphor.

Table 5.10. Median and mean metaphor rates for open- and closed-class parts of speech across TOEIC ability groups in the two conditions

Ability group	Median metaphor rate		Mean metaphor rate (SD)	
	Con.	Exp.	Con.	Exp.
Open-class metaphors				
Low	8.61%	10.78%	8.53% (1.51%)	10.57% (1.65%)
Mid	10.94%	10.47%	10.51% (1.86%)	11.12% (2.19%)
High	11.62%	11.93%	11.96% (1.77%)	11.60% (1.39%)
Closed-class metaphors				
Low	12.47%	11.81%	12.71% (1.46%)	12.18% (2.04%)
Mid	12.79%	11.84%	13.58% (2.44%)	13.55% (3.45%)
High	13.92%	11.82%	13.34% (1.67%)	12.34% (1.69%)

Table 5.11. Kruskal-Wallis tests for open-class metaphor rates across TOEIC ability groups within conditions

Condition	Mean rank			Chi-square	*df*	*p*
	Low	Mid	High			
Control	5.86	12.50	16.88	9.919	2	.007**
Experimental	9.88	11.71	14.38	1.779	2	.411

Note: *n* = 23 in both tests
** = Significant at the 0.01 level

Table 5.12. Mann-Whitney tests for open-class metaphor rates between TOEIC ability groups within conditions

Group	Control condition					Experimental condition				
	n	*U*	*z*	*p*	*r*	*n*	*U*	*z*	*p*	*r*
Low-mid	15	12.000	-1.852	.064	0.48^{++}	15	25.000	-0.347	.728	0.09
Low-high	15	1.000	-3.125	.002**	0.81^{+++}	16	18.000	-1.470	.141	0.37^{++}
Mid-high	16	20.000	-1.260	.208	0.32^{++}	15	23.000	-0.579	.563	0.15^{+}

** = Significant at the 0.01 level
+ = Small effect size; ++ = Medium effect size; +++ = Large effect size

Table 5.13. Kruskal-Wallis tests for closed-class metaphor rates across TOEIC ability groups within conditions

Condition	Mean rank			Chi-square	*df*	*p*
	Low	Mid	High			
Control	10.00	12.50	13.25	0.924	2	.630
Experimental	11.13	13.57	11.50	0.552	2	.759

Note: *n* = 23 in both tests

Table 5.14. Mann-Whitney tests for closed-class metaphor rates between TOEIC ability groups within conditions

Group	Control condition					Experimental condition				
	n	*U*	*z*	*p*	*r*	*n*	*U*	*z*	*p*	*r*
Low-mid	15	22.000	-0.694	.487	0.18^{+}	15	23.000	-0.579	.563	0.15^{+}
Low-high	15	20.000	-0.926	.355	0.24^{+}	16	30.000	-0.210	.834	0.05
Mid-high	16	30.000	-0.210	.834	0.05	15	22.000	-0.694	.487	0.18^{+}

+ = Small effect size

Table 5.15. Mann-Whitney tests for open- and closed-class metaphor rates across equal ability TOEIC groups in each condition

Groups	Open-class MRW				Closed-class MRW			
	U	*z*	*p*	*r*	*U*	*z*	*p*	*r*
Low groups	10.000	-2.083	.037$^{\#}$	0.54^{+++}	25.000	-0.347	.728	0.09
Mid groups	27.000	-0.116	.908	0.03	24.000	-0.463	.643	0.12^{+}
High groups	31.000	-0.105	.916	0.03	22.000	-1.050	.294	0.26^{+}

Note: Low and mid groups *n* = 15, high group *n* = 16
= Non-significant after Bonferroni adjustment
+ = Small effect size; +++ = Large effect size

It is interesting that the findings of studies on German, Greek, French and Japanese learners under examination conditions, as well as Japanese learners in the classroom experiencing either implicit or explicit learning of metaphor display the same broad trends of development. On the whole, open-class metaphor gradually increases with proficiency, while closed-class metaphor either displays a rise-fall-rise s-shaped pattern or plateaus after an initial rise. The effect of explicit instruction on this potential natural tendency in learners is thus not completely clear, although it does appear more pronounced for lower proficiency learners.

5.5. Learner ability and target metaphor production

In section 4.5, we saw that learners in the experimental condition produced taught target metaphors in both greater quantity and variety. However, this chapter has found a more pronounced effect of instruction on lower ability learners than their more proficient classmates. This section will continue to explore this possibility by looking in more detail at the production of taught and untaught target metaphors by learners of different ability. It should be remembered that the control group did not receive regular explicit instruction in metaphor, and for that reason there was little difference for them between a taught and an untaught target metaphor. Nevertheless, the distinction is still useful for drawing comparisons with the output of the experimental group. Figures 5.9 - 5.12 show the development of taught, untaught and combined target metaphors across TOEIC groupings.

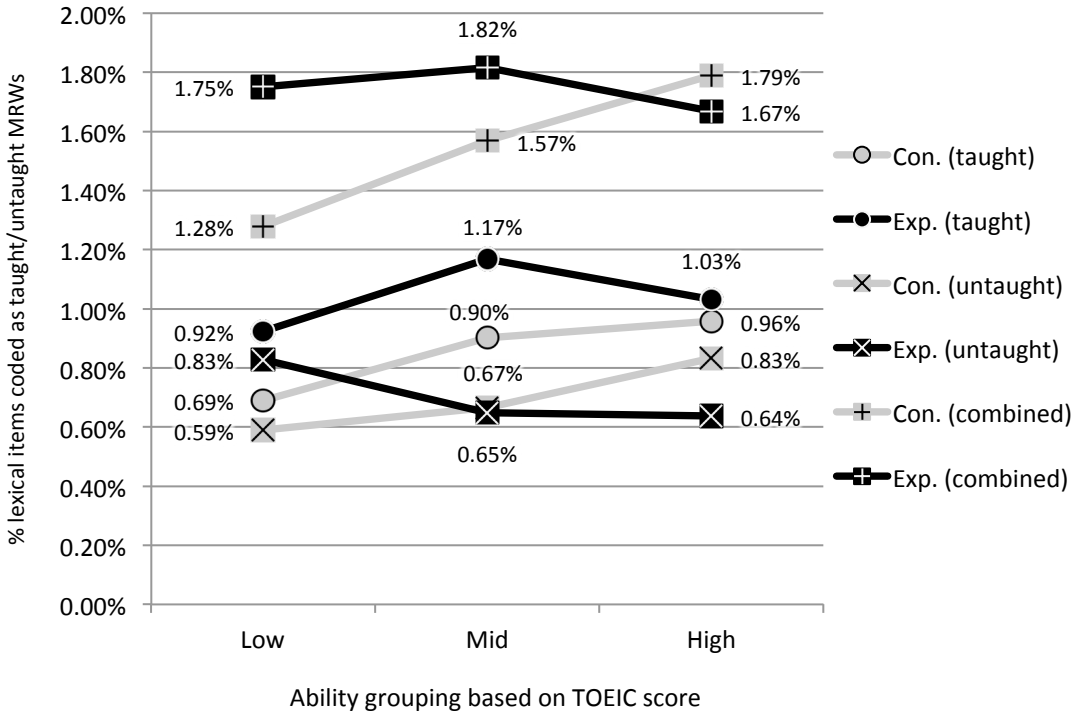


Figure 5.9. Mean target metaphor rates across TOEIC ability groups in the two conditions

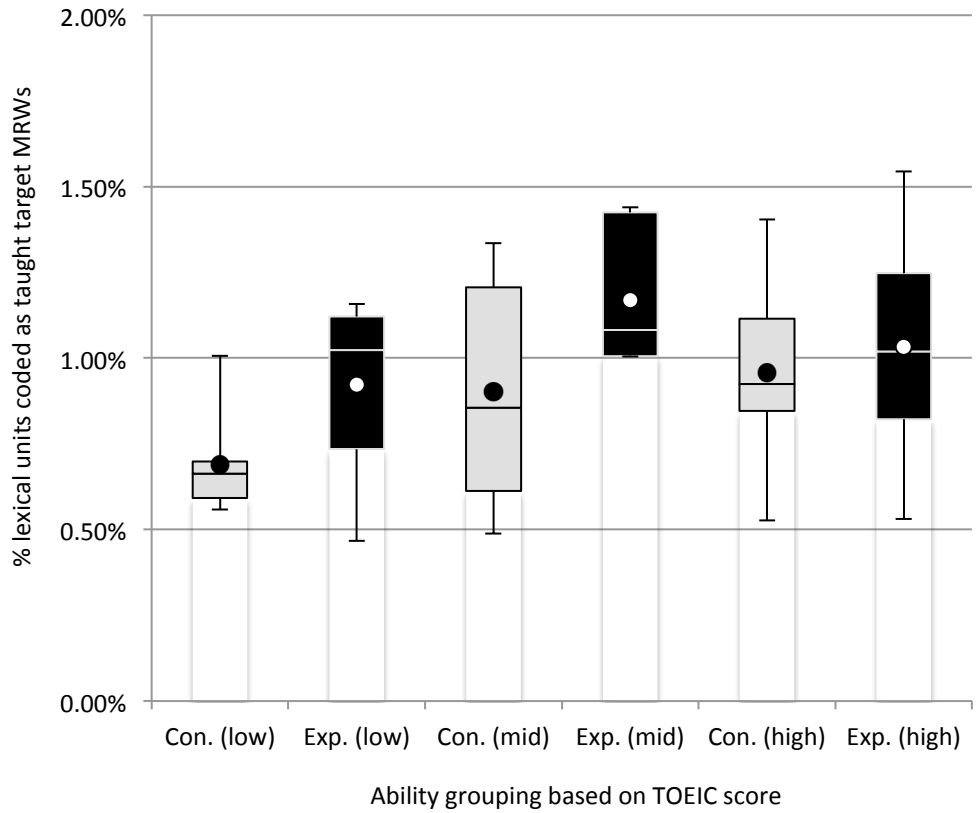


Figure 5.10. Distribution of taught target metaphor rates across TOEIC ability groups in the two conditions

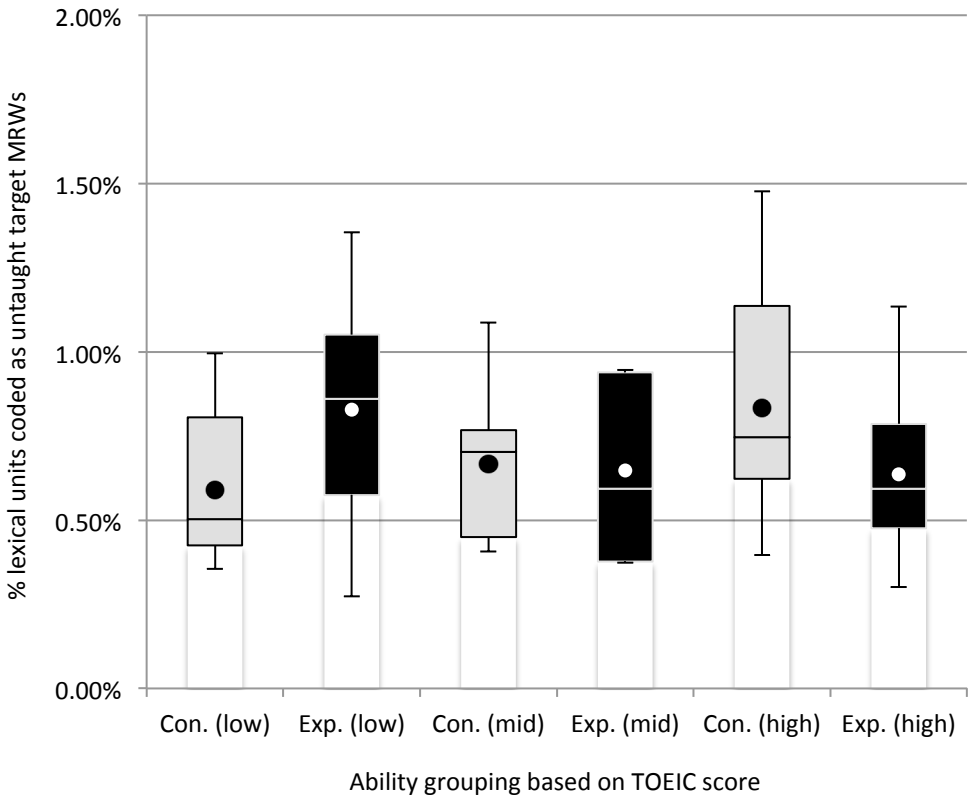


Figure 5.11. Distribution of untaught target metaphor rates across TOEIC ability groups in the two conditions

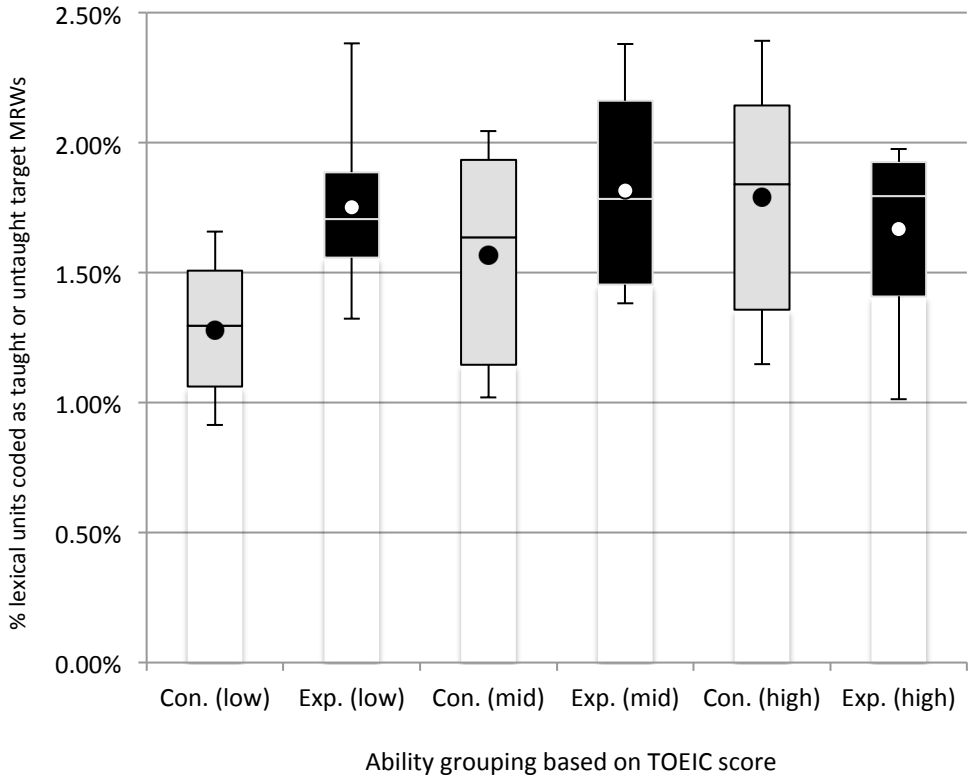


Figure 5.12. Distribution of combined taught and untaught target metaphor rates across TOEIC ability groups in the two conditions

Figure 5.9 supports the view that explicit instruction had a differential effect on taught and untaught target metaphors. For the control group, output of taught and untaught target metaphors gradually increases in line with ability. For the experimental group however, production of taught and untaught target metaphors appears to diverge. The overall quantities of taught target metaphor produced are higher in the experimental condition, which may suggest that learners did respond to the treatment, but again we see that the effect is more pronounced for lower-proficiency learners (Table 5.16).

Table 5.16. Median and mean target metaphor rates across TOEIC ability groups in the two conditions

Ability group	Median target metaphor rate		Mean target metaphor rate (SD)	
	Con.	Exp.	Con.	Exp.
Taught target metaphors				
Low	0.66%	1.02%	0.69% (0.15%)	0.92% (0.24%)
Mid	0.85%	1.08%	0.90% (0.31%)	1.17% (0.19%)
High	0.93%	1.03%	0.96% (0.25%)	1.04% (0.32%)
Untaught target metaphors				
Low	0.50%	0.86%	0.59% (0.23%)	0.83% (0.33%)
Mid	0.70%	0.59%	0.67% (0.22%)	0.64% (0.24%)
High	0.72%	0.59%	0.83% (0.36%)	0.63% (0.25%)
Combined target metaphors				
Low	1.30%	1.71%	1.28% (0.26%)	1.75% (0.32%)
Mid	1.64%	1.78%	1.57% (0.40%)	1.82% (0.38%)
High	1.84%	1.80%	1.79% (0.43%)	1.67% (0.34%)

Table 5.17. Kruskal-Wallis tests for target metaphor rates across TOEIC ability groups within conditions

Condition (category)	Mean rank			Chi square	*df*	*p*
	Low	Mid	High			
Con. (taught)	8.00	12.50	15.00	4.043	2	.132
Exp. (taught)	9.13	14.86	12.38	2.704	2	.259
Con. (untaught)	9.43	11.75	14.50	2.104	2	.349
Exp. (untaught)	14.63	11.29	10.00	1.972	2	.373
Con. (combined)	7.57	12.38	15.50	5.139	2	.077
Exp. (combined)	11.75	13.00	11.38	0.231	2	.891

Note: *n* = 23 for all tests

Table 5.18. Mann-Whitney tests for target metaphor rates across TOEIC ability groups within conditions

Group	Control condition [a]				Experimental condition [b]			
	U	*z*	*p*	*r*	*U*	*z*	*p*	*r*
Taught target metaphors								
Low-mid	17.000	-1.273	.203	0.33^{++}	17.000	-1.273	.203	0.33^{++}
Low-high	11.000	-1.967	.049$^{\#}$	0.51^{+++}	22.000	-1.050	.294	0.26^{+}
Mid-high	27.000	-0.525	.600	0.14^{+}	21.000	-0.810	.418	0.21^{+}
Untaught target metaphors								
Low-mid	23.000	-0.579	.563	0.15^{+}	21.000	-0.810	.418	0.21^{+}
Low-high	15.000	-1.504	.132	0.39^{++}	18.000	-1.470	.141	0.37^{++}
Mid-high	24.000	-0.840	.401	0.22^{+}	26.000	-0.231	.817	0.06
Combined target metaphors								
Low-mid	16.000	-1.389	.165	0.36^{++}	27.000	-0.116	.908	0.03
Low-high	9.000	-2.199	.028$^{\#}$	0.63^{+++}	31.000	0.105	.916	0.03
Mid-high	23.000	-0.945	.345	0.24^{+}	22.000	-0.694	.487	0.18^{+}

a = Control: Low-mid & low-high *n* =15, mid-high *n* = 16
b = Experimental: Low-mid & mid-high *n* = 15, low-high *n* = 16
# = Non-significant after Bonferroni adjustment
+ = Small effect size; ++ = Medium effect size; +++ = Large effect size

No significant differences were found between ability groupings in either condition, although Mann-Whitney tests showed moderate to large effect sizes between the low and high ability groups in the control condition (Tables 5.17 - 5.18).

Mann-Whitney tests with the Bonferroni adjustment applied also revealed that low proficiency learners produced significantly more combined target metaphors in the experimental condition (Table 5.19).

Table 5.19. Mann-Whitney tests for target metaphor rates across equal ability TOEIC groups in each condition

Group	Taught MRW				Untaught MRW				Combined MRW			
	U	***z***	***p***	***r***	***U***	***z***	***p***	***r***	***U***	***z***	***p***	***r***
Low	9.00	-2.199	.028[#]	0.57[+++]	15.00	-1.504	.132	0.39[++]	5.00	-2.662	.008*	0.69[+++]
Mid	13.00	-1.736	.083	0.45[++]	26.00	-0.231	.817	0.06	19.00	-1.042	.298	0.27[+]
High	29.00	-0.315	.753	0.08	21.00	-1.155	.248	0.29[+]	28.00	-0.420	.674	0.11[+]

Note: Low and mid groups *n* = 15, high group *n* = 16
= Non-significant after Bonferroni adjustment
* = Significant at the 0.05 level
+ = Small effect size; ++ = Medium effect size; +++ = Large effect size

The data presented thus far have consistently shown that lower-proficiency learners in the study, who are unlikely yet to have developed core vocabularies, show the most noticeable increase in metaphor production when provided with explicit instruction. For higher-ability learners though, it appears that metaphorical output might undergo a slight dip compared to mid-level learners. It was suggested in section 5.3 that these learners are entering a stage in which they are more inclined to add variety to their range of expression than to produce metaphor in ever-greater quantities. Certainly, the results from the control group in the figures above show that learners with more developed skills are already producing topic-appropriate metaphor without explicit teaching, so it is plausible that the next natural step for them would be to increase their range of production.

Variety of metaphor production was operationalized in the same way as in chapter four. The data presented in Table 5.20 are the median and mean ratios of all target metaphor types produced to the total number of lexical unit types produced at various ability levels. The distributions of these ratios in conditions and TOEIC ability groups are also plotted in Figures 5.13 - 5.15. Figure 5.13 suggests that a greater variety of target metaphor types are being used by high-level learners. Variety measures increase with proficiency in both the taught and combined categories, but there appears to be no clear trend in the variety of use of untaught metaphors in either condition.

Table 5.20. Mean ratios of target metaphor types to all types produced for learners across TOEIC ability groups in the two conditions

Group	Median target metaphor type ratio		Mean target metaphor type ratio (SD)	
	Con. (SD)	Exp. (SD)	Con. (SD)	Exp. (SD)
Taught target metaphor types				
Low	1.61%	2.43%	1.63% (0.48%)	2.31% (0.55%)
Mid	1.91%	2.63%	2.17% (0.71%)	2.67% (0.66%)
High	2.50%	2.69%	2.44% (0.78%)	2.82% (0.72%)
Untaught target metaphor types				
Low	1.46%	1.61%	1.45% (0.63%)	1.78% (0.78%)
Mid	1.66%	1.50%	1.65% (0.38%)	1.53% (0.35%)
High	1.70%	1.75%	1.71% (0.64%)	1.65% (0.55%)
Combined target metaphor types				
Low	2.93%	3.84%	3.08% (0.86%)	4.09% (1.04%)
Mid	3.68%	4.10%	3.82% (0.87%)	4.20% (0.70%)
High	4.38%	4.49%	4.15% (1.03%)	4.47% (0.87%)

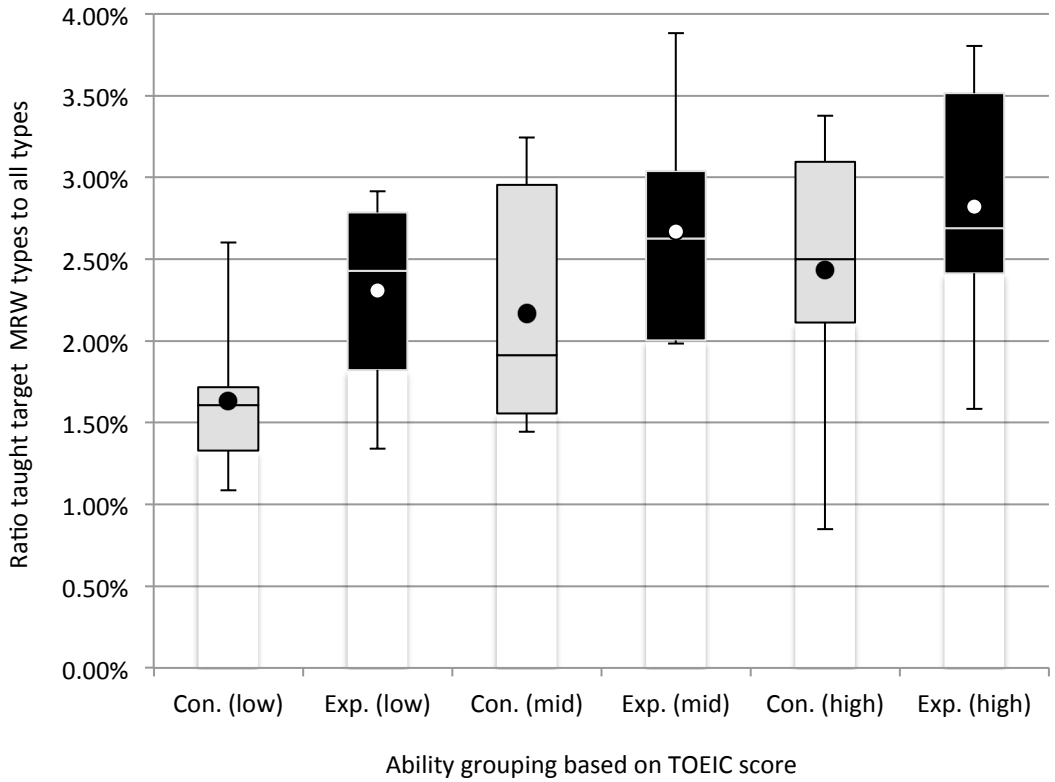


Figure 5.13. Distribution of taught target metaphor type ratios across TOEIC ability groups in the two conditions

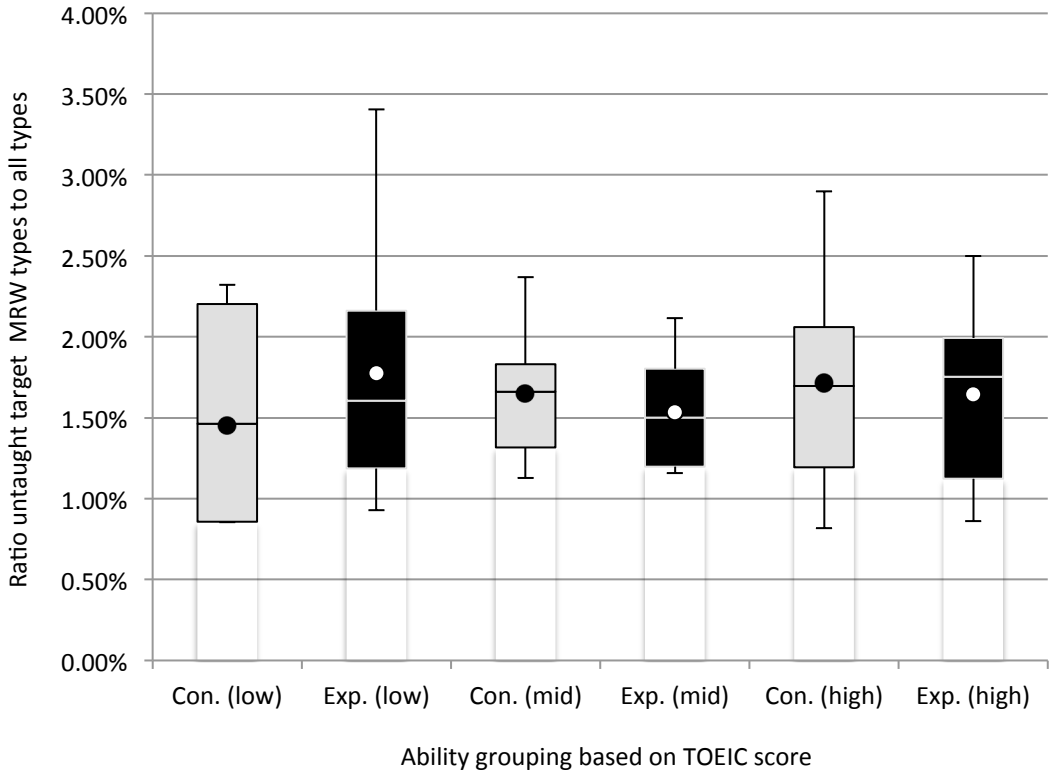


Figure 5.14. Distribution of untaught target metaphor type ratios across TOEIC ability groups in the two conditions

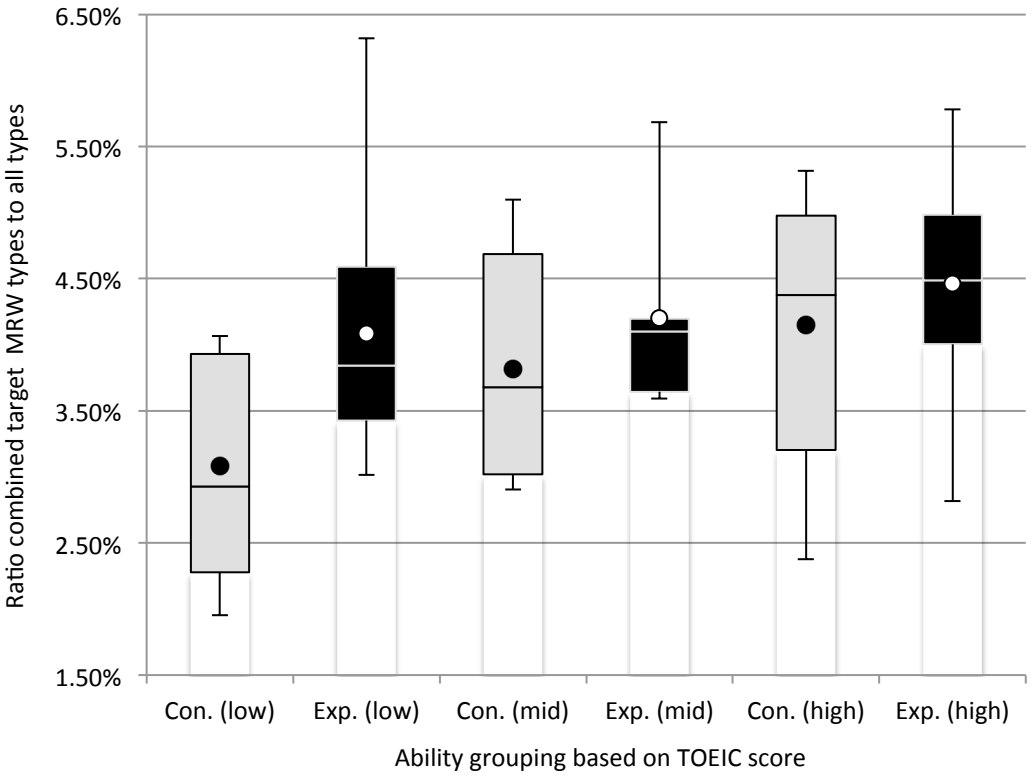


Figure 5.15. Distribution of combined target metaphor type ratios across TOEIC ability groups in the two conditions

Figures 5.13 - 5.15 broadly suggest that variety of metaphor use increased with proficiency in both conditions, and explicit instruction helped to increase variety, as well as quantity, of production. Despite these trends, none of the differences for any category of metaphor either within or between the two conditions were found to be significant (Tables 5.21 - 5.23). Nevertheless, several of the *p* values approach significance at the 95% confidence level and have moderate to strong effect sizes, indicating that it would be worth carrying out an expanded study with larger samples in order to provide more robust evidence.

The effect sizes in Table 5.23 suggest that explicit instruction in metaphor had the greatest effect on the writing of lower proficiency learners, although this does not appear to have carried over into the use of untaught target metaphors. Regarding higher ability learners, the figures above and effect sizes in the tables

Table 5.21. Kruskal-Wallis tests for target metaphor type ratios across TOEIC ability groups within conditions

Condition (type)	Mean ranks			Chi-square	*df*	*p*
	Low	Mid	High			
Control (taught)	7.43	12.63	15.38	5.229	2	.073
Experimental (taught)	9.25	12.71	14.13	2.178	2	.337
Control (untaught)	9.86	12.88	13.00	1.006	2	.605
Experimental (untaught)	12.50	11.00	12.38	0.220	2	.896
Control (combined)	7.14	13.00	15.25	5.601	2	.061
Experimental (combined)	9.38	11.71	14.88	2.648	2	.266

Note: *n* = 23 for all tests

Table 5.22. Mann-Whitney tests for target metaphor type ratios across TOEIC ability groups within conditions

Group	Control condition					Experimental condition				
	n	*U*	*z*	*p*	*r*	*n*	*U*	*z*	*p*	*r*
Taught target metaphor types										
Low-mid	15	13.000	-1.736	.083	0.45^{++}	15	20.000	-0.926	.355	0.24^{+}
Low-high	15	11.000	-1.967	.049$^{\#}$	0.51^{+++}	16	18.000	-1.470	.141	0.37^{++}
Mid-high	16	22.000	-1.050	.294	0.27^{+}	15	25.000	-0.347	.728	0.09
Untaught target metaphor types										
Low-mid	15	19.000	-1.042	.298	0.27^{+}	15	24.000	-0.463	.643	0.12^{+}
Low-high	15	22.000	-0.694	.487	0.18^{+}	16	32.000	0.000	.1.000	0.00
Mid-high	16	30.000	-0.210	.834	0.05	15	25.000	-0.347	.728	0.09
Combined target metaphor types										
Low-mid	15	13.000	-1.736	.083	0.45^{++}	15	19.000	-1.042	.298	0.27^{+}
Low-high	15	9.000	-2.199	.028$^{\#}$	0.57^{+++}	16	20.000	-0.926	.355	0.23^{+}
Mid-high	16	25.000	-0.735	.462	0.19^{+}	15	17.000	-1.273	.203	0.33^{++}

= Non-significant after Bonferroni adjustment
\+ = Small effect size; ++ = Medium effect size; +++ = Large effect size

Table 5.23. Mann-Whitney tests for target metaphor type ratios across equal ability TOEIC groups in each condition

Group	Taught types				Untaught types				Combined types			
	U	*z*	*p*	*r*	*U*	*z*	*p*	*r*	*U*	*z*	*p*	*r*
Low	11.000	-1.967	.049[#]	0.51[+++]	19.000	-1.042	.298	0.27[+]	16.000	-1.389	.165	0.36[++]
Mid	15.000	-1.504	.132	0.39[++]	23.000	-0.579	.563	0.15[+]	21.000	-0.810	.418	0.21[+]
High	23.000	-0.945	.345	0.24[+]	31.000	-0.105	.916	0.03	27.000	-0.525	.600	0.13[+]

Note: Low and mid groups *n* = 15, high group *n* = 16
= Non-significant after Bonferroni adjustment
+ = Small effect size; ++ = Medium effect size; +++ = Large effect size

imply that such learners' written output was affected least by the experimental treatment. At the same time, it should be remembered that the higher proficiency learners in the experimental group produced the greatest variety of taught and combined target metaphors despite their lower overall usage of these categories relative to experimental group mid level learners or high level learners in the control group (Figures 5.10 and 5.12). This again raises the possibility that they lie in a different developmental stage in which change is more evident in terms of variety than quantity.

5.6. Language ability and use of metaphor for rhetorical purposes

The few studies that have examined development in learners' use of metaphor for rhetorical purposes have suggested that this is a very gradual process encompassing perhaps the entire spectrum of ability (Littlemore et al., 2014, p. 130 - 137; Turner, 2014, p. 291 - 305). These studies have found that what begins as only limited, rule-bound production of metaphorical prepositions leads to conventionalised forms and text structuring devices, use of expressions to subtly shift meaning, and finally to the ability to employ metaphor in a creative or imaginative fashion. Thus far, the only

data we have on this process has come from examination papers, so it is possible that learners in lower-stakes settings may produce these forms of metaphor earlier than other studies have found.

The MIPVU data were adapted to accommodate a broader selection of polywords, as described in sections 3.4.2.2 and 5.2.2 (pages 68, 169). The MRWs were then coded with the rhetorical function list from Table 5.5. However, an initial inspection of the coding revealed that some of the categories were barely present or even non-existent in the data; none of the emotive, mitigation, advising, manipulative or reiteration categories covered more than 0.08% of the total word count in either condition. It is possible that this was to some degree due to the genre of writing. Learners were being asked to reflect on the concepts and themes of an academic course rather than interact with the readers of the texts (the course instructors), and for this reason, interpersonal metaphor may have felt inappropriate. It is worth noting that these five categories of metaphor were also only present at a trace level in the Japanese learner data analysed by Turner (2014, p. 293 - 301), so it is also possible that there is some cultural or linguistic effect at work. Because of these low frequencies, it was decided to collapse the categories into the all evaluative, all interpersonal, all textual, and open-class unmarked conventional groupings shown in Table 5.24 and Figure 5.16.

Rhetorical type	Condition						Log-likelihood [a]	Bayes factor effect size
	Control			Experimental				
	Frequency	% of MRWs	% of lexical units	Frequency	% of MRWs	% of lexical units		
Emotive	32	0.59%	0.07%	43	0.63%	0.08%	-0.01	-11.51
Emphasis	238	4.39%	0.54%	283	4.15%	0.50%	1.21	-10.30
General evaluative	75	1.38%	0.17%	134	1.97%	0.24%	-4.74*	-6.77
Mitigation	33	0.61%	0.08%	27	0.40%	0.05%	3.31	-8.20
All evaluative	**378**	**6.98%**	**0.86%**	**487**	**7.14%**	**0.85%**	**0.06**	**-11.46**
Advising	1	0.03%	0.002%	3	0.04%	0.01%	/	/
Manipulative	0	0.00%	0.00%	2	0.03%	0.00%	/	/
All interpersonal	**1**	**0.03%**	**0.002%**	**5**	**0.07%**	**0.01%**	**-1.94**	**-9.58**
Reiteration	0	0.00%	0.00%	1	0.01%	0.00%	/	/
Structuring	967	26.65%	2.21%	1009	14.80%	1.77%	25.51^{****}	14.00^{+++}
All textual	**967**	**26.65%**	**2.21%**	**1010**	**14.82%**	**1.77%**	$\mathbf{23.30^{****}}$	$\mathbf{13.79^{+++}}$
Unmarked conventional (OC)	2283	62.91%	5.21%	3161	46.37%	5.55%	-4.04*	-7.48

a = Log-likelihood values calculated as proportions of entire output corpora (Control: 43,480 words; Experimental: 56,972 words). Minimum frequency = 5 in either corpus.

* = Significant at 0.05 level; **** = Significant at 0.0001 level

+++ = Very strong evidence against null hypothesis

Table 5.24. Log-likelihood tests on metaphors grouped by rhetorical function between the two conditions

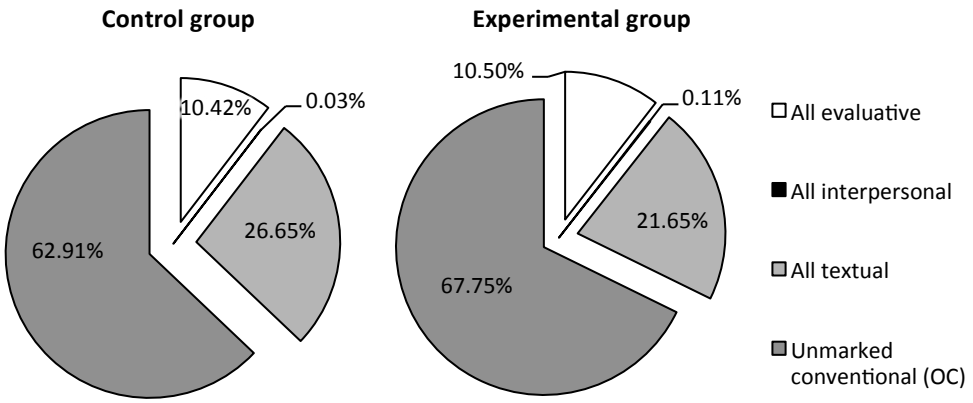


Figure 5.16. Proportions of metaphors grouped by rhetorical function in the two conditions

As Figure 5.16 shows, the distribution of rhetorical types across conditions was similar, with the only clear differences being that textual metaphors were much more frequent in the control condition, while unmarked conventional (OC) metaphors were more frequent in the experimental condition (although the effect size was weak). The change in unmarked conventional metaphor was perhaps to be expected, as the bulk of the target metaphorical expressions fell into this category. Since the focus of the study was on these unmarked usage patterns, this section offers a useful opportunity to consider how an experimental treatment might influence performance in related aspects of language use.

In order to examine whether the use of metaphor to perform rhetorical functions varied with learner ability, Figures 5.17 - 5.19 were plotted. Interpersonal metaphors were not analysed since they so infrequent, but the median and mean values of each TOEIC ability group's metaphor rates for the three other rhetorical functions are provided in Table 5.25. Evaluative metaphor use increases with proficiency in the control condition, and a significant difference was found between the low and high groups in this condition (Tables 5.26 - 5.27). For the experimental

group, a more distorted pattern appears, with the low-level group producing a similar amount of evaluative metaphors to their higher level classmates. With the Bonferroni adjustment applied, the difference in the low level learners' output between the two conditions fell just outside the significant range, but there was a large effect size (Table 5.28).

For textual metaphors, the clear drop in usage from the control to the experimental condition stands out. The high-level learners in the control condition produced significantly more of this type of metaphor than did their experimental group equals (Table 5.28). No significant differences were found within the conditions, although several moderate effect sizes were obtained (Tables 5.26 - 5.27).

For open-class unmarked conventional metaphors, the gradual increase with proficiency in the control condition was again present. Once more the *p* values lay just outside the significant range, but large effect sizes were found (Tables 5.26 - 5.27). Experimental group learners produced these metaphors at higher rates than the control, but this difference diminishes with increased proficiency, and higher level learners produced marginally fewer conventional metaphors than did the mid-level subgroup. No significant differences were found within the experimental condition. This is a very similar result to that found for taught target metaphors (Tables 5.17 - 5.18).

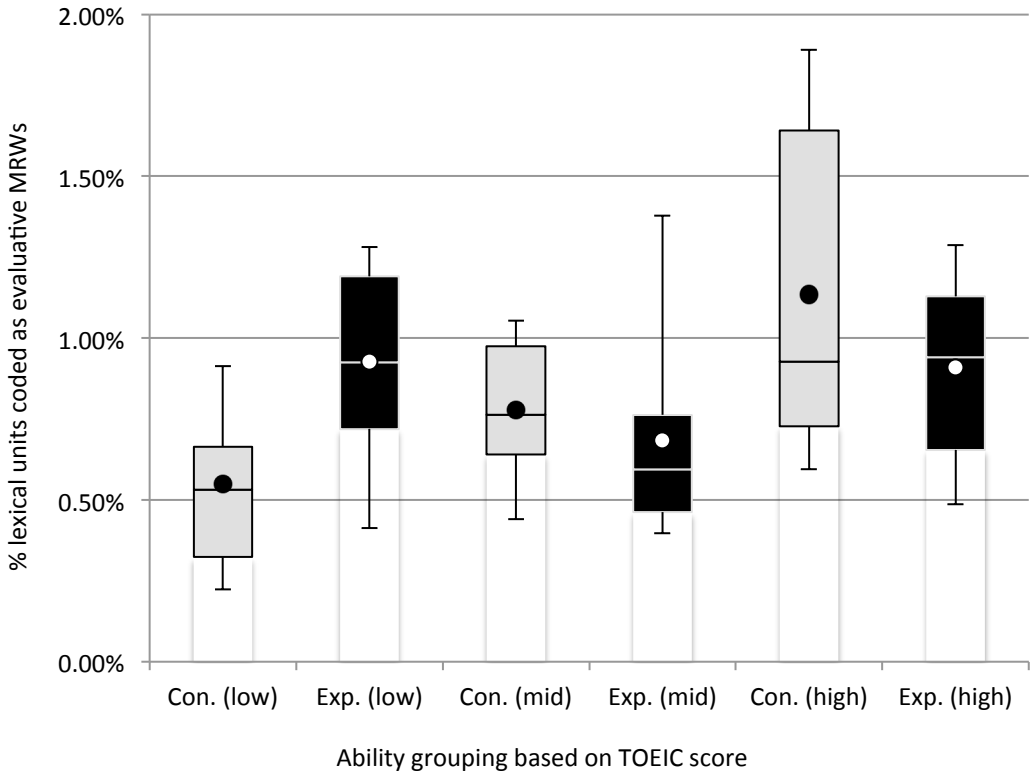


Figure 5.17. Distribution of metaphor rates for evaluative metaphors across TOEIC ability groups in the two conditions

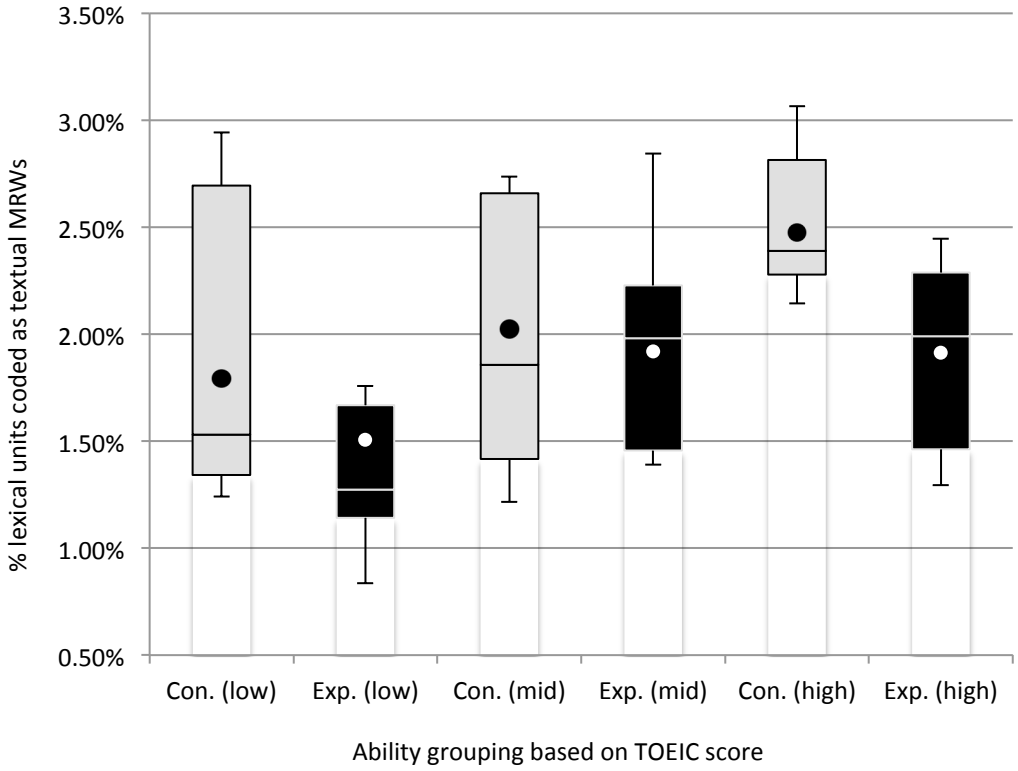


Figure 5.18. Distribution of metaphor rates for textual metaphors across TOEIC ability groups in the two conditions

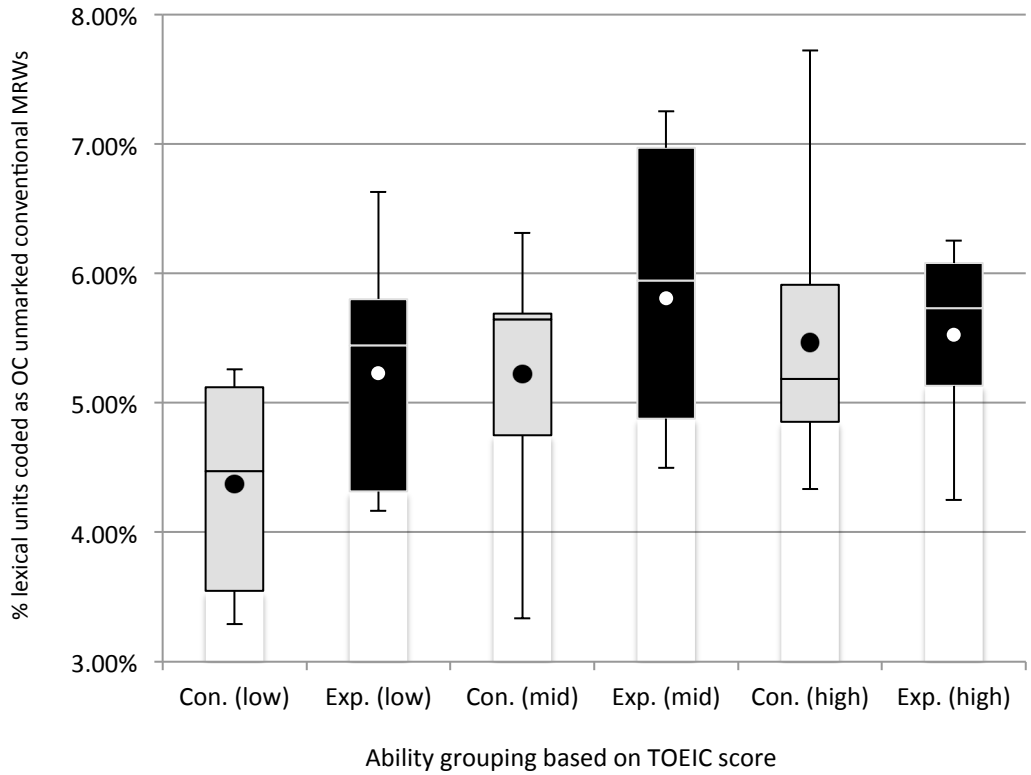


Figure 5.19. Distribution of metaphor rates for open-class unmarked conventional metaphors across TOEIC ability groups

Table 5.25. Median and mean rhetorical function metaphor rates across TOEIC ability groups in the two conditions

Ability group	Median rhetorical function metaphor rate		Mean rhetorical function metaphor rate	
	Con.	Exp.	Con.	Exp.
Evaluative metaphors				
Low	0.53%	0.93%	0.55% (0.23%)	0.92% (0.30%)
Mid	0.76%	0.59%	0.78% (0.21%)	0.68% (0.31%)
High	0.93%	0.94%	1.13% (0.50%)	0.91% (0.27%)
Textual metaphors				
Low	1.53%	1.27%	1.79% (0.70%)	1.51% (0.67%)
Mid	1.86%	1.98%	2.02% (0.63%)	1.92% (0.50%)
High	2.39%	1.99%	2.47% (0.31%)	1.91% (0.43%)
Open-class unmarked conventional metaphors				
Low	4.47%	5.44%	4.37% (0.73%)	5.23% (0.85%)
Mid	5.64%	5.94%	5.23% (0.92%)	5.81% (1.06%)
High	5.18%	5.73%	5.47% (1.04%)	5.52% (0.66%)

Table 5.26. Kruskal-Wallis tests for metaphor rates grouped by rhetorical function across TOEIC ability groups within conditions

Condition (function)	Mean ranks			Chi-square	*df*	*p*
	Low	Mid	High			
Con. (evaluative)	6.71	12.25	16.38	7.591	2	.022*
Exp. (evaluative)	14.00	8.14	13.38	3.288	2	.193
Con. (textual)	8.57	11.38	15.63	4.142	2	.126
Exp. (textual)	7.75	14.00	14.50	4.837	2	.089
Con. (unmarked conventional OC)	7.00	14.38	14.00	5.481	2	.065
Exp. (unmarked conventional OC)	9.75	14.00	12.50	1.533	2	.465

Note: *n* = 23 for all tests
* = Significant at 0.05 level

Table 5.27. Mann-Whitney tests for metaphor rates grouped by rhetorical function across TOEIC ability groups within conditions

Group	Control condition					Experimental condition				
	n	*U*	*z*	*p*	*r*	*n*	*U*	*z*	*p*	*r*
Evaluative metaphors										
Low-mid	15	13.000	-1.736	.083	0.45^{++}	15	14.000	-1.620	.105	0.42^{++}
Low-high	15	6.000	-2.546	.011*	0.66^{+++}	16	30.000	-0.210	.834	0.05
Mid-high	16	19.000	-1.365	.172	0.34^{++}	15	15.000	-1.504	.132	0.39^{++}
Textual metaphors										
Low-mid	15	20.000	-0.926	.355	0.24^{+}	15	12.000	-1.852	.064	0.48^{++}
Low-high	15	12.000	-1.852	.064	0.48^{++}	16	14.000	-1.890	.059	0.47^{++}
Mid-high	16	19.000	-1.365	.172	0.34^{++}	15	26.000	-0.231	.817	0.06
Open-class unmarked conventional metaphors										
Low-mid	15	10.000	-2.083	.037$^{\#}$	0.54^{+++}	15	17.000	-1.273	.203	0.33^{++}
Low-high	15	11.000	-1.967	.049$^{\#}$	0.51^{+++}	16	25.000	-0.735	.462	0.18^{+}
Mid-high	16	31.000	-0.105	.916	0.03	15	25.000	-0.347	.728	0.09

= Non-significant after Bonferroni adjustment
* = Significant at the 0.05 level
+ = Small effect size; ++ = Medium effect size; +++ = Large effect size

Table 5.28. Mann-Whitney tests for metaphor rates grouped by rhetorical function across equal ability TOEIC groups in each condition

Group	Evaluative metaphors				Textual metaphors				Unmarked conventional (OC) metaphors			
	U	*z*	*p*	*r*	*U*	*z*	*p*	*r*	*U*	*z*	*p*	*r*
Low	8.00	-2.315	.021[#]	0.60[+++]	19.00	-1.042	.298	0.27[+]	14.00	-1.620	.105	0.42[++]
Mid	16.00	-1.389	.165	0.36[++]	25.00	-0.347	.728	0.09	19.00	-1.042	.298	0.27[+]
High	26.00	-0.630	.529	0.16[+]	9.00	-2.415	.016*	0.60[+++]	24.00	-0.840	.401	0.21[+]

Note: Low and mid groups *n* = 15, high group *n* = 16
= Non-significant after Bonferroni adjustment
* = Significant at the 0.05 level
+ = Small effect size; ++ = Medium effect size; +++ = Large effect size

The small samples and narrow range of abilities for the participants in this study limit the claims that can be made for rhetorical use of metaphor, but the trends in the data here might permit the following cautious observations. Firstly, in the output of learners who are not being instructed in metaphor in any way, use of metaphor for the three rhetorical purposes considered here will generally increase with proficiency. Secondly, the metaphor awareness-raising activities used in this study appear to have had the most noticeable effect on lower-proficiency learners, at least in terms of increased frequency of output. These learners appear to be reaching or even surpassing their more able classmates' use of both evaluative and conventional metaphors. Finally, there is a strong suggestion that the control group produced greater quantities of textual metaphor than the learners in the experimental condition. There was a similar finding for target metaphors, with the gains in production of taught target metaphors and the lack of change in untaught metaphor usage. One possible explanation for this is that since the majority of target metaphors were in the unmarked conventional category, the increased attentional

resources learners spent on applying target forms appropriately led to a comparative inattention to other language features.

5.6.1. Lexical variety within rhetorical use of metaphor

The analysis so far has dealt with rhetorical use of metaphor at a broad-scale, considering only the quantity of production. The remainder of this chapter will consider the data at a micro-level, first by examining lexical variety within rhetorical usage, then by looking at the particular parts of speech and individual word choices made across ability levels and conditions.

The ratio of the mean number of types within each rhetorical function grouping to the total number of types produced by each learner was calculated in the same manner as in sections 4.5 and 5.5 (pages 139 & 182), and the results are presented in Figures 5.20 - 5.22 and Tables 5.29 - 5.32. The results for the control condition follow the same trends as those in the previous section. That is, the variety of metaphors used for each of the rhetorical function categories increased with proficiency. For both evaluative and unmarked conventional metaphors, there were significant differences between the variety of metaphors produced by the low and high groups (Table 5.31), and for evaluative metaphors, the mid-level group also produced a significantly greater variety of metaphors than the low-level group.

In the experimental condition, variety of metaphor also increased in line with proficiency for textual and unmarked metaphors, although none of the differences were significant. This is likely due to the elevated levels of production by lower level learners in the experimental group, as has been discussed in other sections. The low-level experimental condition participants produced a significantly greater variety of

evaluative metaphors than the low-level control group learners, while the difference in variety of conventional metaphors was just outside the level of significance once the Bonferroni adjustment had been applied. In both cases, the effect size was large (Table 5.31). As was the case with taught target metaphors, the experimental high-level group produced a greater, though not significant, variety of unmarked conventional metaphors than the mid-level group despite producing less of these forms overall. This suggests that the treatment had some success in encouraging learners to break away from tried and trusted expressions.

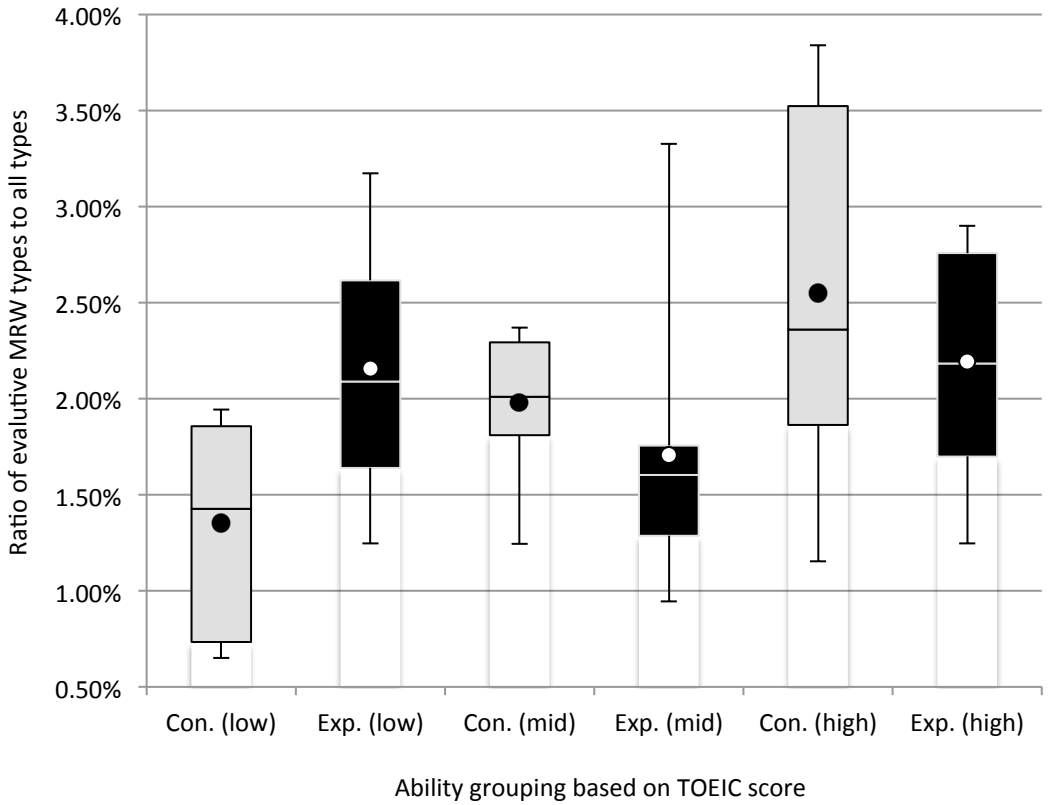


Figure 5.20. Distribution of evaluative metaphor type ratios across TOEIC ability groups in the two conditions

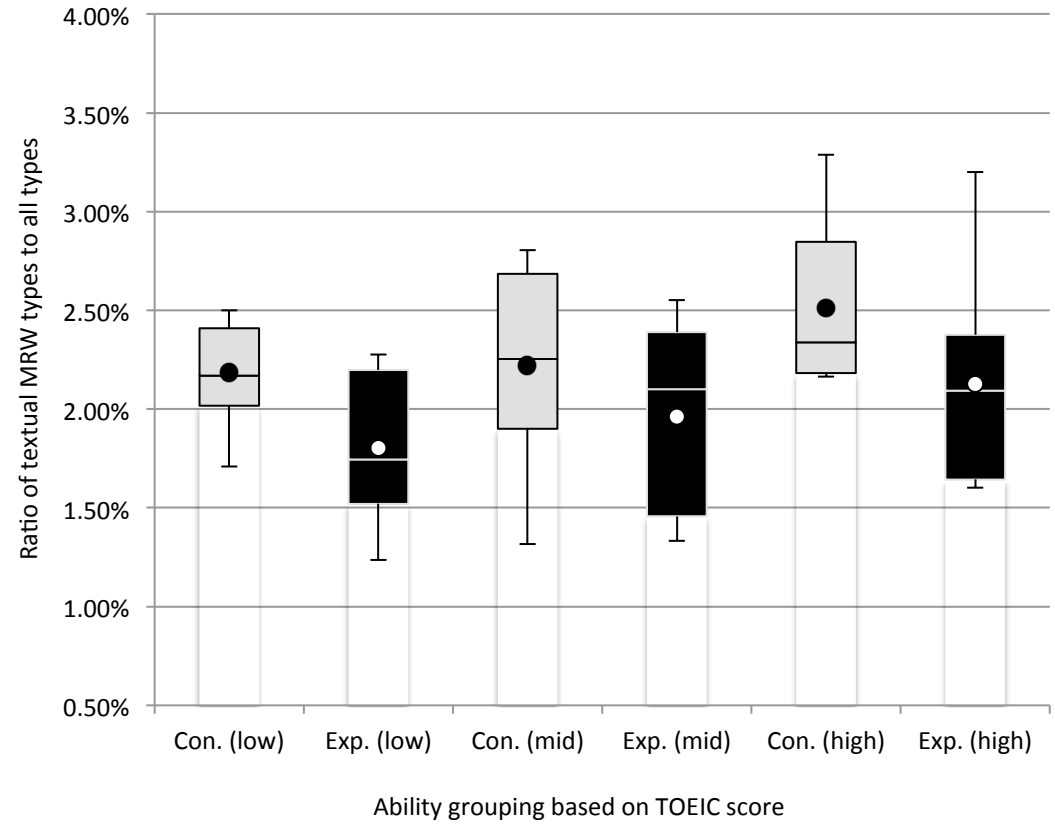


Figure 5.21. Distribution of textual metaphor type ratios across TOEIC ability groups in the two conditions

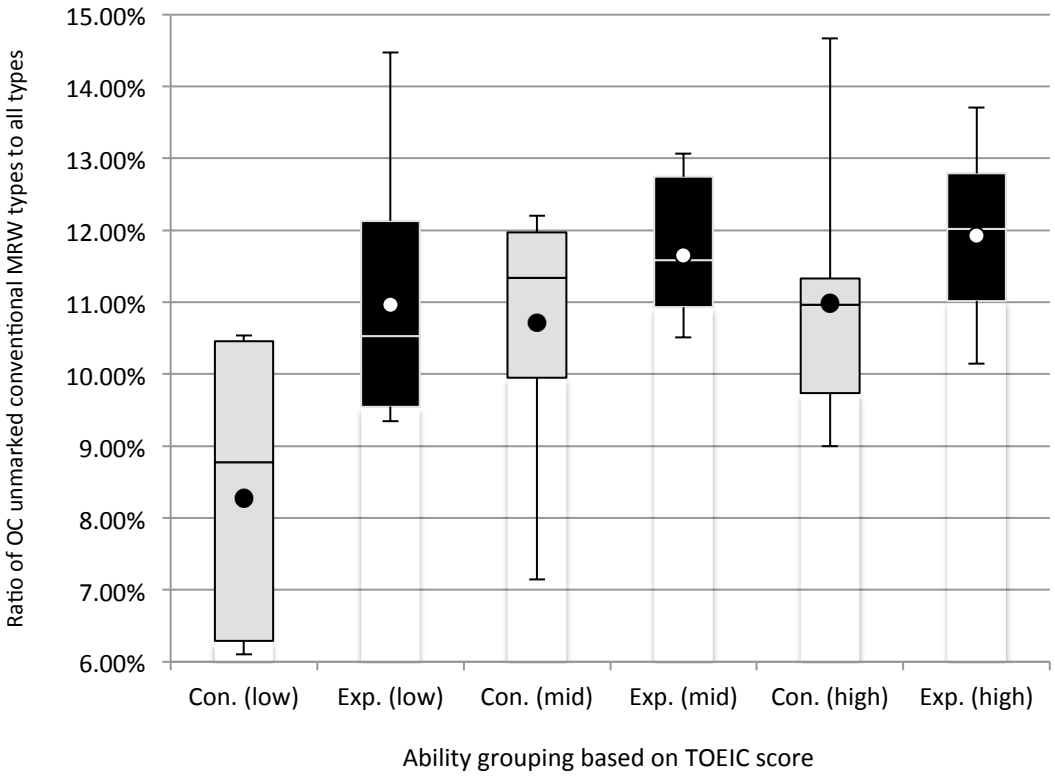


Figure 5.22. Distribution of open-class unmarked conventional metaphor type ratios across TOEIC ability groups in the two conditions

Table 5.29. Median and mean rhetorical function metaphor type ratios across TOEIC ability groups in the two conditions

Ability group	Median rhetorical function metaphor type ratio		Mean rhetorical function metaphor type ratio	
	Con.	Exp.	Con.	Exp.
Evaluative metaphor types				
Low	1.43%	2.09%	1.35% (0.53%)	2.16% (0.62%)
Mid	2.01%	1.60%	1.98% (0.36%)	1.71% (0.77%)
High	2.36%	2.18%	2.55% (0.92%)	2.19% (0.59%)
Textual metaphor types				
Low	2.17%	1.75%	2.19% (0.27%)	1.80% (0.38%)
Mid	2.25%	2.10%	2.22% (0.49%)	1.96% (0.47%)
High	2.34%	2.09%	2.51% (0.41%)	2.13% (0.54%)
Open-class unmarked conventional metaphor types				
Low	8.78%	10.53%	8.28% (1.94%)	10.97% (1.75%)
Mid	11.34%	11.58%	10.71% (1.69%)	11.65% (0.96%)
High	10.97%	12.02%	10.98% (1.72%)	11.92% (1.15%)

Table 5.30. Kruskal-Wallis tests for rhetorical metaphor type ratios across TOEIC ability groups within conditions

Condition (rhetorical type)	Mean ranks			Chi-square	*df*	*p*
	Low	Mid	High			
Control (all evaluative)	5.86	13.00	16.38	9.245	2	.010*
Experimental (all evaluative)	13.63	8.57	13.38	2.577	2	.276
Control (all textual)	9.71	10.88	15.13	2.714	2	.257
Experimental (all textual)	9.75	12.00	14.25	1.761	2	.415
Control (unmarked conventional OC)	6.14	15.13	14.00	7.615	2	.022*
Experimental (unmarked conventional OC)	8.88	13.14	14.13	2.682	2	.262

Note: *n* = 23 for all tests
* = Significant at the 0.05 level

Table 5.31. Mann-Whitney tests for rhetorical metaphor type ratios across TOEIC ability groups within conditions

Group	Control condition					Experimental condition				
	n	*U*	*z*	*p*	*r*	*n*	*U*	*z*	*p*	*r*
Evaluative metaphor types										
Low-mid	15	7.000	-2.430	.015*	0.63^{+++}	15	16.000	-1.389	.165	0.36^{++}
Low-high	15	6.000	-2.546	.011*	0.66^{+++}	16	31.000	-0.105	.916	0.03
Mid-high	16	19.000	-1.365	.172	0.34^{++}	15	16.000	-1.389	.165	0.36^{++}
Textual metaphor types										
Low-mid	15	26.000	-0.463	.643	0.12^{+}	15	24.000	-0.463	.643	0.12^{+}
Low-high	15	14.000	-1.620	.105	0.42^{++}	16	18.000	-1.470	.141	0.37^{++}
Mid-high	16	21.000	-1.155	.248	0.29^{+}	15	24.000	-0.463	.643	0.12^{+}
Open-class unmarked conventional metaphor types										
Low-mid	15	8.000	-2.315	.021$^{\#}$	0.36^{++}	15	16.000	-1.389	.165	0.36^{++}
Low-high	15	7.000	-2.430	.015*	0.63^{+++}	16	19.000	-1.365	.172	0.34^{++}
Mid-high	16	27.000	-0.525	.600	0.13^{+}	15	24.000	-0.463	.643	0.12^{+}

= Non-significant after Bonferroni adjustment
* = Significant at the 0.05 level
+ = Small effect size; ++ = Medium effect size; +++ = Large effect size

Table 5.32. Mann-Whitney tests for rhetorical metaphor type ratios across equal ability TOEIC groups in each condition

Group	Evaluative types ratio				Textual types ratio				Unmarked conv. (OC) types ratio			
	U	*z*	*p*	*r*	*U*	*z*	*p*	*r*	*U*	*z*	*p*	*r*
Low	7.00	-2.430	.015*	0.63^{+++}	14.00	-1.620	.105	0.42^{++}	8.00	-2.315	.021$^{\#}$	0.60^{+++}
Mid	13.00	-1.736	.083	0.45^{++}	21.00	-0.810	.418	0.21^{+}	21.00	-0.810	.418	0.21^{+}
High	24.50	-0.788	.431	0.20^{+}	19.00	-1.365	.172	0.34^{++}	17.00	-1.575	.115	0.39^{++}

Note: Low and mid groups *n* = 15, high group *n* = 16
= Non-significant after Bonferroni adjustment
* = Significant at the 0.05 level
+ = Small effect size; ++ = Medium effect size; +++ = Large effect size

5.6.2. Rhetorical functions and parts of speech

It seems that awareness-raising activities for content-related metaphorical themes may have as much or an even greater impact on lexical variety as they do on overall

production rates. This is particularly true for learners who are reaching intermediate ability and who already demonstrate some capability of using vocabulary in extended senses. A logical extension of this finding, then, would be to consider the particular parts of speech and word choices that are used by learners of varying proficiency across the two conditions.

5.6.2.1. Evaluative metaphors

The category of evaluative metaphors was comprised of terms that performed emotive, emphatic or mitigating functions, as well as those that provided a general value judgement. Not surprisingly, as Figure 5.23 shows, adjectives came to the fore in this regard, and there were also prominent roles for nouns, adverbs, and verbs.

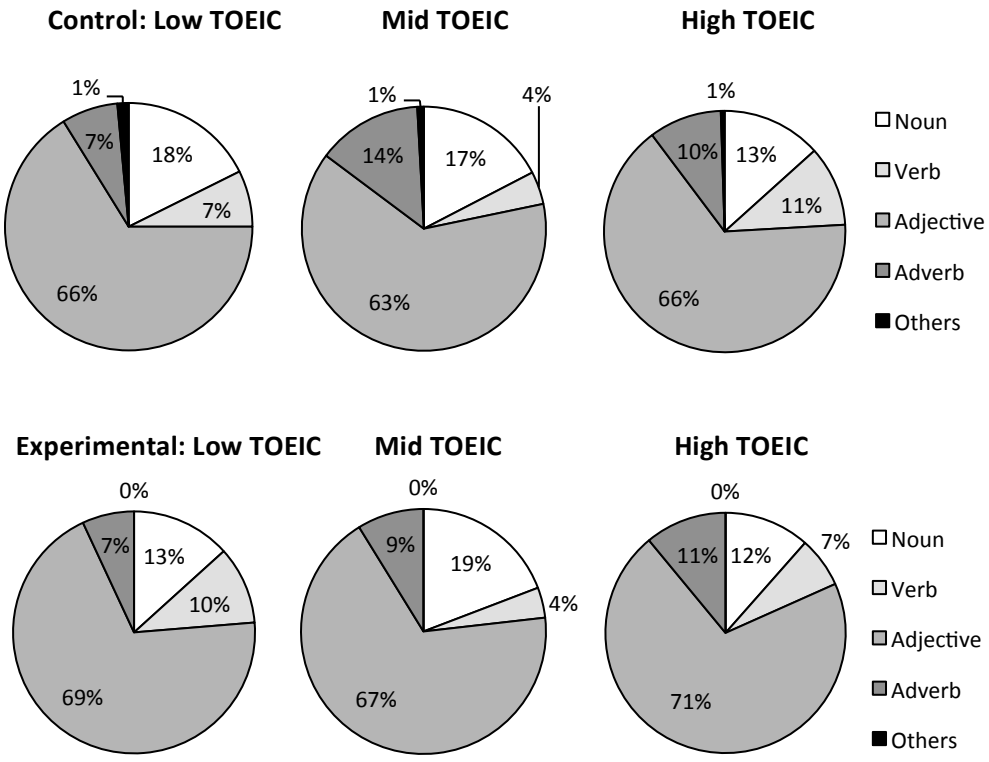
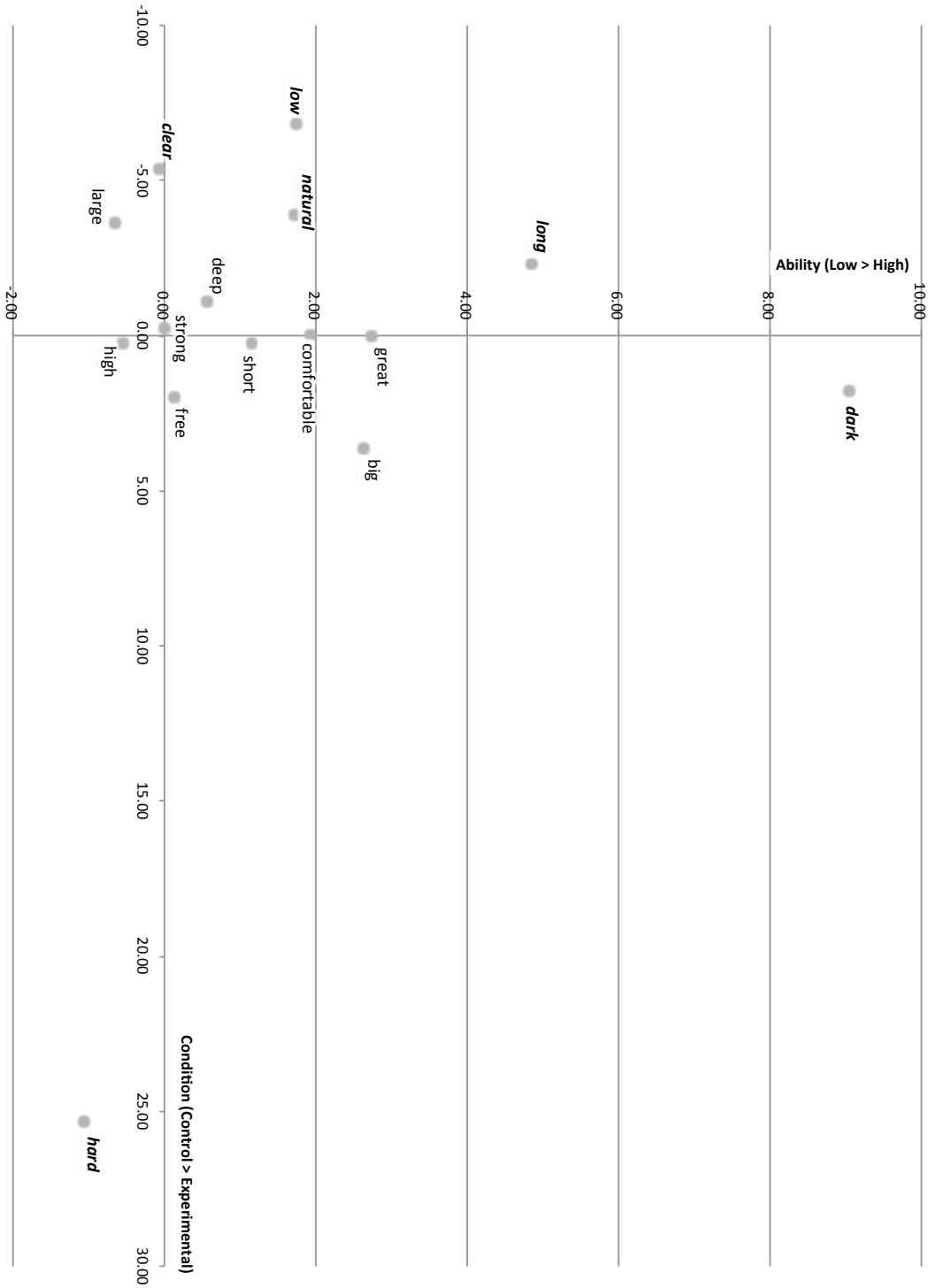


Figure 5.23. Proportions of evaluative metaphors grouped by part of speech and TOEIC ability group in the two conditions

Evaluative adjectives

A variety of metaphorical evaluative adjectives were used with moderate to high frequencies in each condition. In order to show the distribution of adjectives across conditions and ability levels, two sets of log-likelihood values were calculated for each adjective. The first value compared the relative frequency with which each word appeared in either the control or experimental output corpus, and the second reflected the distribution of each word within higher or lower ability learners' output. The results are shown in Figure 5.24, with words that significantly favoured one condition or ability range appearing in bold italic font. Full data is provided in Appendix K (page 473).

Here we can see that the majority of adjectives were quite evenly distributed between the conditions. The one adjective that strongly favoured the experimental condition, *hard*, was not a target form. In both conditions, it was used only in the sense of 'difficult' or 'full of problems' (51), so no obvious reason can be found for this change of behaviour. Two adjectives, dark and *long*, did not favour either condition, but were significantly skewed in their use towards high ability learners. *Dark* was a target expression that appeared in the course materials in the phrase *she's escaping to the dark side*, used to describe a woman's taking refuge from her daily problems in a macabre sub-culture, and later in a description of Meiji-era life as *grey and dark*. Two higher ability control group learners attempted to produce this word in a new fashion (52 - 53), and another was able to use a phrase two months after he had seen it in class (54). However, it is interesting to note that while *dark* was explicitly taught in the experimental condition, learners appeared to limit their usage to the patterns in which the forms appeared. 89% of uses of this word were in



Note: Significant values (log-likelihood ≥ |3.84|) appear in bold italic font.
Note: Minimum frequency: 5 occurrences in either corpus

Figure 5.24. Scatter plot of evaluative metaphorical adjective use across ability range and conditions

the phrase *dark side*, and the remainder used *grey and dark*. While *long* was frequently used in both conditions, it had a very restricted range of use, with 48% of its uses immediately preceding history, and another 30% of occurrences before *time*. Neither *clear* nor *natural*, both of which were favoured by control group learners, were taught explicitly in class, but *low* was a target form that was rarely used by experimental group learners.

(51) It must be Japanese tradition because elder Japanese people like making us do a really hard training. [EXP19] [High TOEIC]

(52) ...those who are Gothic Lolita are apt to have some dark parts in their mind such as rebellious mind. [CON1] [High TOEIC]

(53) According to the article which I read last time, people who love Gothic-Lolita fashion have something dark in their mind, but I don't think so. [CON17] [High TOEIC]

(54) These dark side of sumo should be changed better. [CON14] [High TOEIC]

Evaluative nouns

A wide variety of evaluative noun forms were produced, mostly with only a small number of uses each. However, a small group of nouns did appear relatively often. The noun form of *boom* was commonly used in both conditions by mid and higher proficiency learners. Since it is a loanword in Japanese, it was no surprise to find it used to describe cultural trends frequently in the data (55), although it was noted that it only appeared as a noun in the control group's output, while the experimental group produced the verb form frequently as well. *Power* was used by a variety of learners in several metaphorical senses: as an influential force (56), an ability (57), or

a country with economic influence (58). Finally, the target form *treasure* was used by learners of all abilities in the experimental condition, often in the pattern *a treasure of [country]*, but also in certain fixed phrases, as in (59). Although not a loanword, this phrase also exists as a direct translation in Japanese.

(55) *...vocaloid's boom will not continue in the future because vocaloid's music is sung by machine voice which is not human voice.* [CON20] [Mid TOEIC]

(56) *Vocaloid music is the power which spreads all over the world.* [CON9] [Mid TOEIC]

(57) *I think that power of observation is needed for writing a diary.* [EXP4] [Mid TOEIC]

(58) *I think that Korea is music big power, so CDs, goods and live tickets born many effects,...* [EXP6] [Low TOEIC]

(59) *...a Kabuki actor "Ebizu" is a living national treasure.* [EXP16] [Mid TOEIC]

Evaluative adverbs

Adverbs that took on an evaluative metaphorical function were examined, and patterns were identified. *Deeply* was frequently used by learners of a range of abilities in both conditions, as in examples (60) and (61). *Dramatically* (62) was also present in both conditions, with a tendency towards being used by higher proficiency learners. Certain other adverbs, however, displayed differing behaviours in the two conditions. *Strongly* appeared only three times in the control condition, both times by mid-TOEIC level learners. In the experimental group, it was more frequent, but was used exclusively by high-ability learners (63). Similarly, the adverbial form of *hard* was used only twice in the control condition by mid- and

high-level learners to eight times by experimental group members, mostly by low- or mid-level participants. The vast majority of adverbs in both conditions were used for emphasis; smaller groups had a general evaluative or emotive function, and few adverbs were used for mitigation.

(60) In summary, Sumo has essential component, and it is related to Japanese culture deeply. [CON11] [Low TOEIC]

(61) I think that Lolita fashion and Otaku are related to Japanese anime culture deeply and strongly. [EXP7] [High TOEIC]

(62) I think Japanese society has changed dramatically since Makiko's era. [CON22] [Mid TOEIC]

(63) I think Japanese things reminiscent Japanese culture strongly like kimono are jewels of Japanese culture,… [EXP22] [High TOEIC]

Evaluative verbs

Evaluative verb production was characterised by a wide variety of verbs being used at relatively low frequencies. Although Figure 5.23 suggests little change in their overall rate of use, experimental group learners may have been influenced by the target metaphor themes. 22% of evaluative verbs produced by the control group were taught target expressions, and 13% were untaught, whereas for the experimental group, the corresponding figures were 47% and 25% (64 - 66). Only one of these target metaphors, the verb form *boom*, showed significantly higher use in the experimental group, while another target form, *surround*, was used at a significantly higher rate in the control (67). The verb *dominate* was favoured in the control condition, and the concordance lines revealed that these learners may have

been responding to lexical items in the course materials they received. *Dominate* was used in the title of an article discussing trends in the popularity of western or local music, and all instances of this form appeared in the writing on that topic (68).

(64) In the future, I think [sumo's] popularity won't change and continue to fluctuate, at least, it won't skyrocket. [EXP11] [High TOEIC]

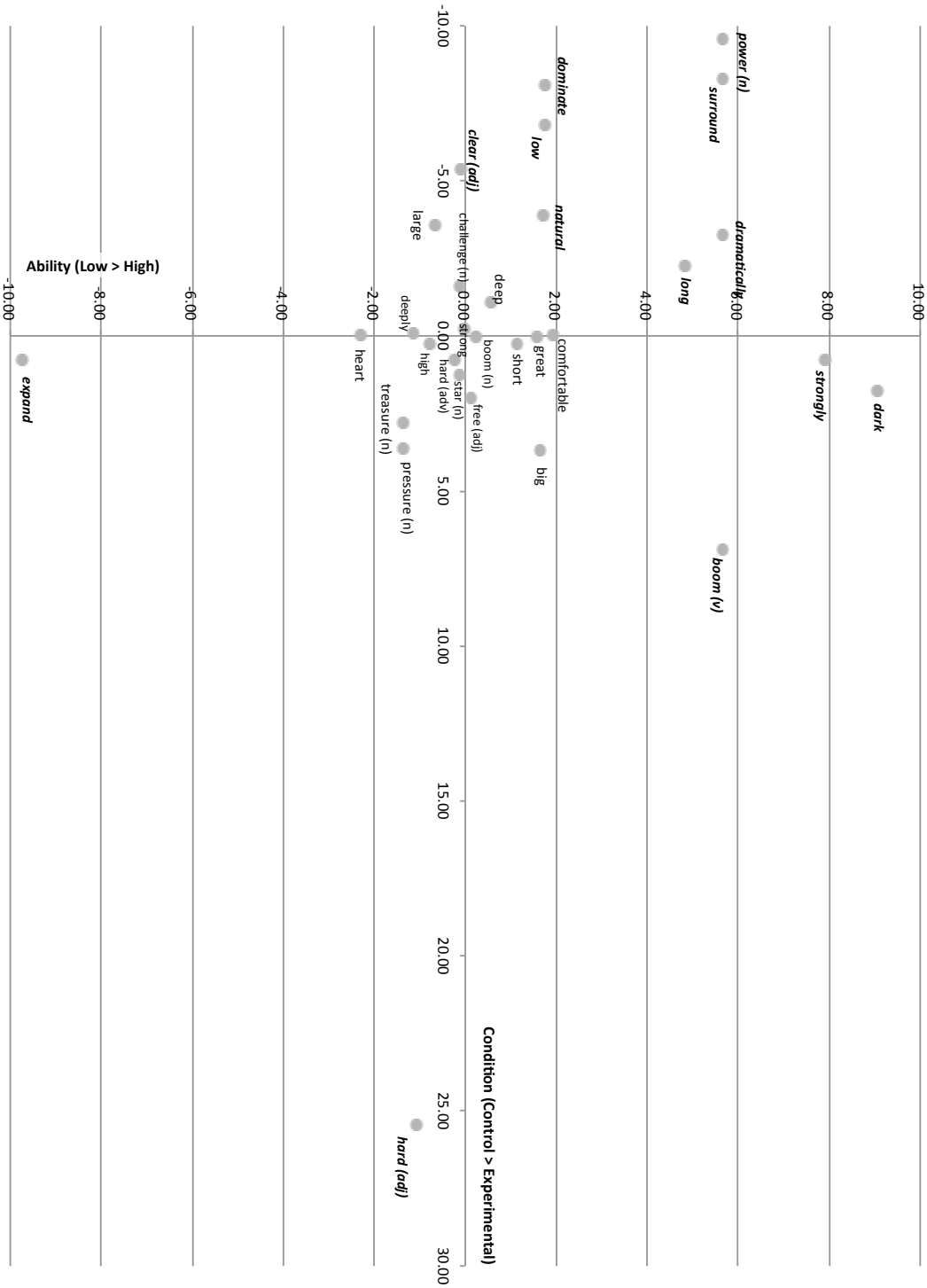
(65) In addition, there are many kinds of cuisine using kimchi in Japan, so the demand for Kimchi has mushroomed throughout Japan. [EXP7] [High TOEIC]

(66) Mass media advertise the wedding ceremonies and the popularity of the wedding ceremonies has boomed. [EXP11] [High TOEIC]

(67) Since I grew up with Japanese culture surrounding me, Japanese popular culture is the only culture I know pretty much. [CON8] [High TOEIC]

(68) British and American music used to shape the society dominate other countries. [CON21] [Mid TOEIC]

Figure 5.25 provides a summary of the distribution of evaluative metaphors between different ability groups and conditions. It seems that evaluative metaphors were generally more frequent in the output of higher ability learners, although learners across the full range of ability in the study were able to use metaphors for evaluative purposes. Several of the taught target words also showed some degree of favouring higher ability learners. *Boom* (*v*), *dark*, *long*, and *surround* were target forms that were more common in high-level learner output, while no taught target words were more frequent for lower level learners. These findings support the claim that evaluative metaphor becomes more apparent in learner writing as proficiency improves (Littlemore et al., 2014, p. 132 - 137).



Note: Minimum frequency: 5 occurrences in either corpus
Note: Significant values (log-likelihood ≥ |3.84|) appear in bold italic font.

Figure 5.25. Scatter plot of evaluative metaphor use across ability range and conditions

5.6.2.2. Textual metaphors

An examination of the parts of speech used to produce textual metaphors showed trends across both conditions, as shown in Figure 5.26. In each case, higher proficiency was associated with greater use of metaphorical pronouns at the expense of adverbs. Determiners also comprised a large proportion of textual metaphors across all ability levels, while prepositions and conjunctions were used at lower rates.

The charts give the impression that the restricted categories of pronouns and determiners were more frequent in the output of higher ability learners than the category of adverbs, which offer more potential variety in the form of phrases (see Appendix K, page 473). However, a brief comparison with COCA data suggests another explanation. Table 5.33 shows the combined frequencies of all parts of speech used for structuring metaphors by participants in this study compared with the equivalent words in the academic section of COCA. It can be seen that learner writing makes relatively greater use of adverbs, although less so in the experimental group than the control. For pronouns and determiners, learners are underusing these parts of speech in comparison with American academic writing, although it should be noted that the COCA data here includes all uses of these words, metaphorical or otherwise. This limitation notwithstanding, these values suggest that what we actually see in Figure 5.26 is learner writing going through the typical developmental stages of over- and underuse, and perhaps some suggestion that they gradually approach native-speaker norms.

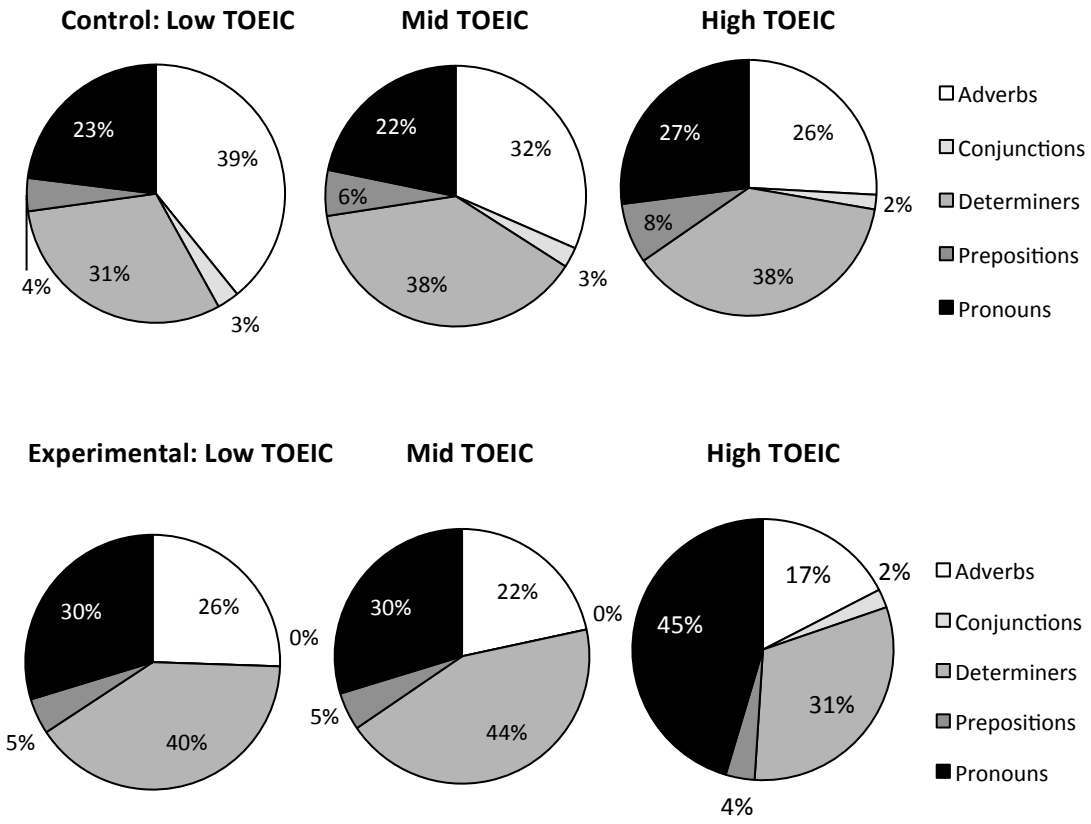


Figure 5.26. Proportions of textual metaphors grouped by part of speech and TOEIC ability group in the two conditions

Table 5.33. Frequencies of parts of speech used for textual metaphors in the two learner output corpora and COCA

Data source	Combined frequencies per 10,000 words				
	Adverbs	Conjunctions	Determiners	Prepositions	Pronouns
Control group	67.76	5.02	80.31	13.69	53.84
Experimental group	37.39	1.58	67.23	7.55	63.36
COCA	29.88	6.22	108.87	7.82	181.97

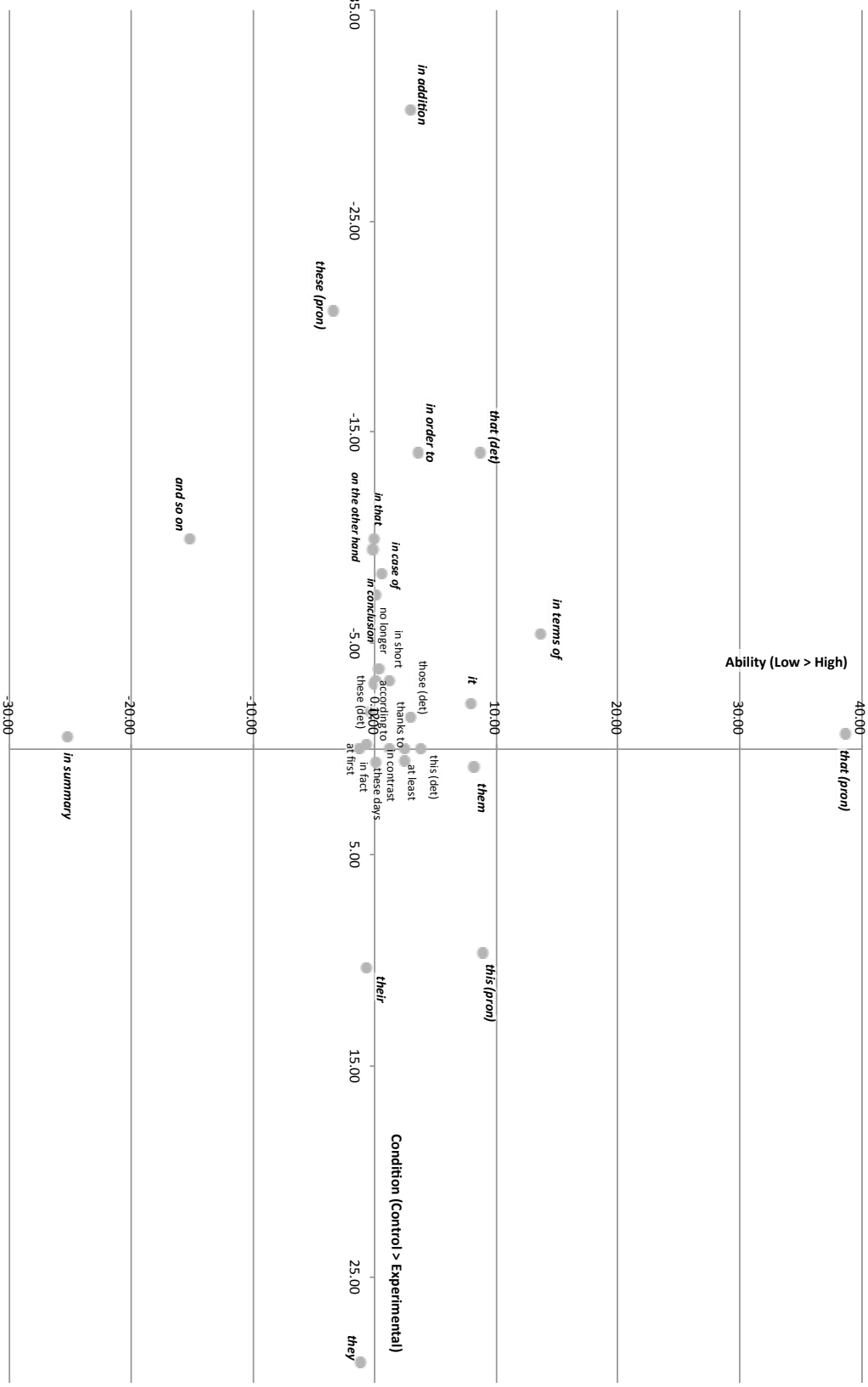
Note: COCA data refers to all uses, not just metaphorical uses.

Looking at the particular language used in writing, it is apparent that learners in both conditions were still dependent on standard expressions to structure text. Only one reiterating metaphor was found in the output from either condition, and this was derived from the course materials. In example (69), the phrase 'radical

followers' was used in the same way to refer to Gothic Lolitas as it had been in the source text.

(69) *They wanted to be new and tell the other people that they are different from people before, so "gothic lolitas" started. Some radical followers hurt themselves to make sure that they are alive.* [EXP8] [High]

Figure 5.27 plots the distributions of structuring metaphors in the two conditions. We can see that few forms favoured the experimental condition, reflecting the overall drop in production of this variety of metaphor compared with the control group. A grammatical split also appears in the data, with most pronouns and determiners being skewed toward higher ability learners and conjunctive adverbs being preferred by lower ability learners, reflecting the pattern in Figure 5.26.



Note: Minimum frequency: 5 occurrences in either corpus
Note: Significant values (log-likelihood ≥ |3.84|) appear in bold italic font.

Figure 5.27. Scatter plot of structuring metaphors across ability range and conditions

5.6.2.3. Unmarked conventional metaphors

The distribution of open-class unmarked conventional metaphors by part of speech was remarkably similar across both conditions and ability (Figure 5.28). Verbs were the dominant form in all cases, which may reflect learners' focus on describing the processes of cultural flow and interaction in their writing. Nouns were also evident in this usage, while adjectives and adverbs, having played a more prominent role in evaluative and/or structuring metaphors, were only present at lower frequencies.

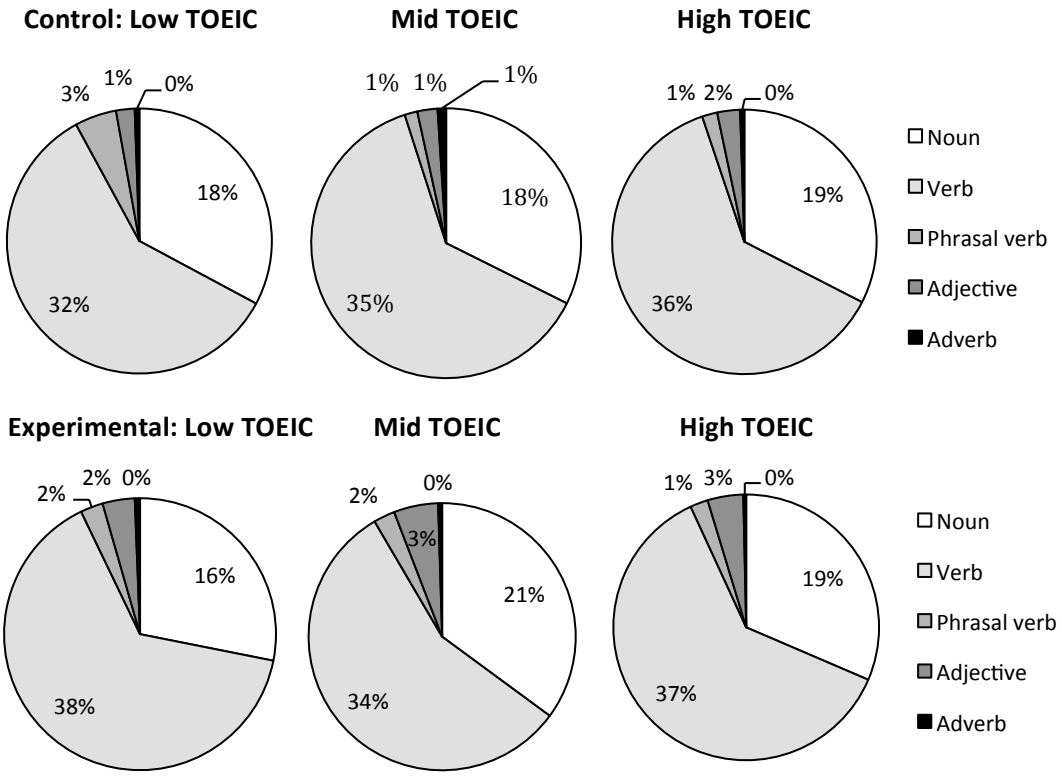


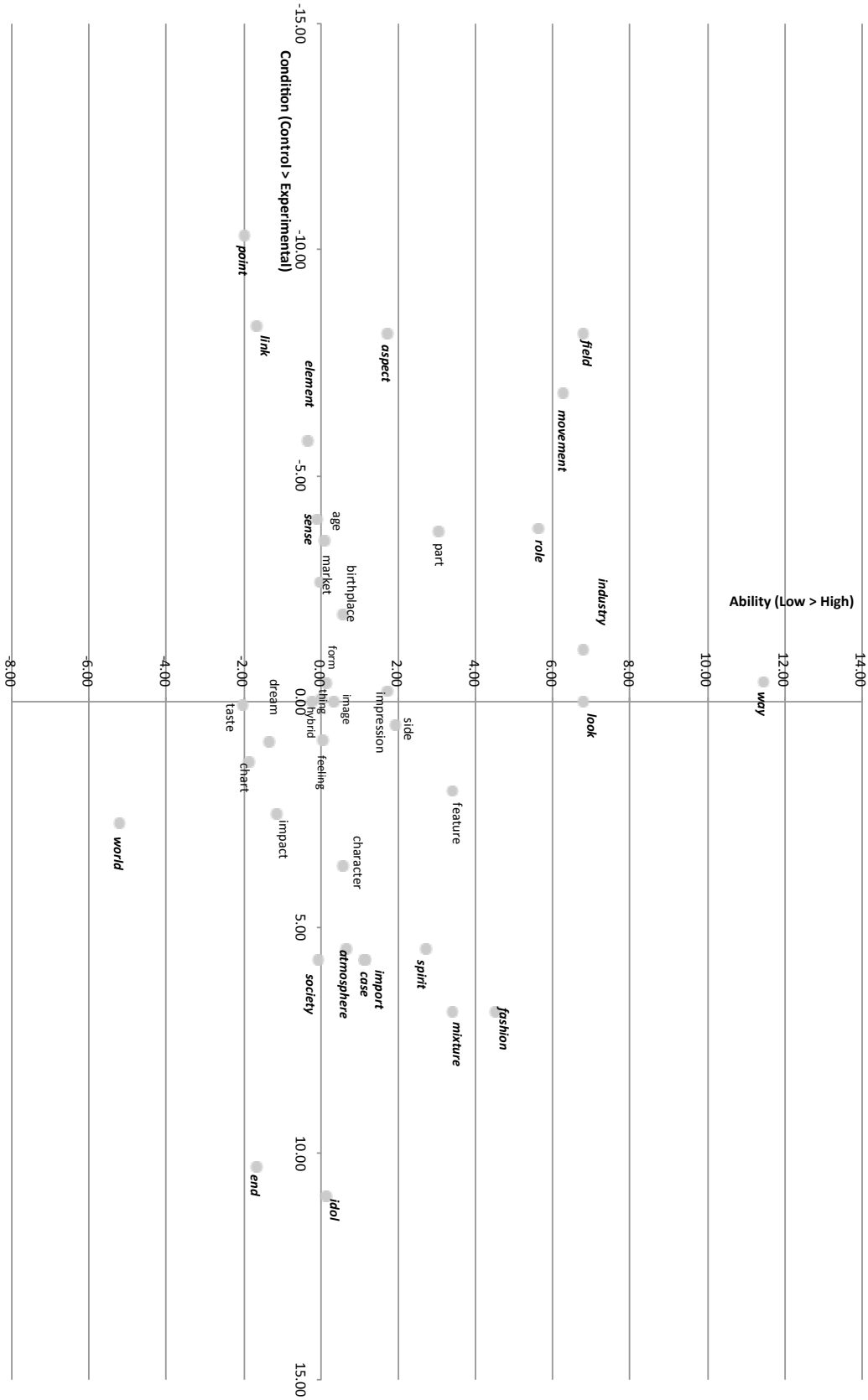
Figure 5.28. Proportions of open-class unmarked conventional metaphors grouped by part of speech and TOEIC ability group in the two conditions

Unmarked conventional nouns

The nouns that performed an unmarked conventional function in text were a combination of target forms (e.g., *movement*, *mixture*, *element*) and high frequency

words (e.g., *way*, *end*, *point*). Examining Figure 5.29, it is noticeable that target nouns were not especially favoured in the experimental group. The only target nouns to be favoured by the experimental group were *import* (70) and *mixture*, which was often used in the *mixture of A and B* pattern to describe cultural entities with a hybrid makeup (71). The other taught target nouns were *aspect*, *element*, *link*, *movement*, and *part*, all of which favoured the control group, and *hybrid*, which did not favour either condition. *Movement* was strongly biased toward high proficiency learners, who made use of its 'group of people with the same aim' sense, as in (72). *Part*, *aspect*, and *elements* were all used to express the complex nature of abstract notions (73 - 75), but it is interesting that *elements*, the least frequent of the three according to COCA, showed more inclination to be used by lower ability learners.

The data also demonstrate that learners in both conditions were able to employ a number of nouns metaphorically without explicit teaching. In some cases, such as *impression* and *idol*, it is possible that they did so without being aware of the literal senses of these words, but with terms like *birthplace* (76) it seems plausible that learners were aware of both literal and metaphorical senses. What is also clear from the data is that generalisations about metaphor production based on learner ability and word frequency cannot easily be made. While Figure 5.29 does suggest that higher ability learners will make more use of metaphorical nouns on average, some of the words that most favoured high proficiency learners were high frequency terms such as *way*, *field*, *movement*, and *look*. It is extremely unlikely that lower proficiency learners were unaware of the literal senses of these words, yet they did not draw on them in metaphorical senses as much as high level learners in their written work.



Note: Significant values (log-likelihood ≥ |3.84|) appear in bold italic font.
Note: Minimum frequency: 5 occurrences in either corpus

Figure 5.29. Scatter plot of open-class unmarked conventional metaphorical noun use across ability range and conditions

(70) *Japanese people believe that Valentine's Day is a cultural import.* [EXP15] [High TOEIC]

(71) *Harajuku fashion is so individualistic fashion and the mixture of many kinds of fashions and contrasting things,...* [EXP10] [Mid TOEIC]

(72) *If the new fashion movements occur in the U.S. many people will follow them.* [CON17] [High TOEIC]

(73) *...people think that rice is a part of Japanese culture.* [CON4] [High TOEIC]

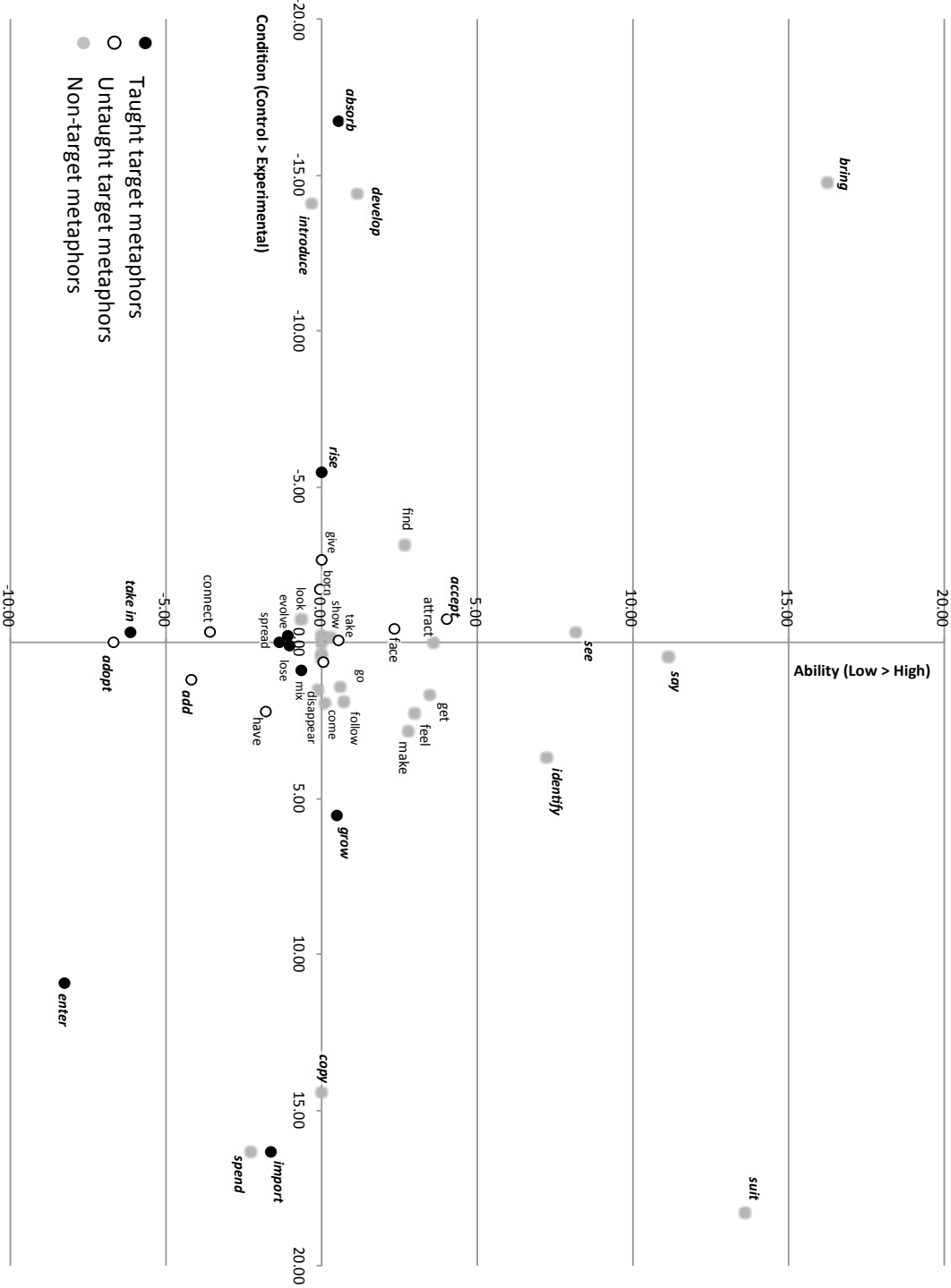
(74) *Secondly, Sumo has changed in several aspects.* [CON23] [Mid TOEIC]

(75) *Sumo has historical elements, and it symbolizes Japan.* [CON11] [Low TOEIC]

(76) *Karuta is a familiar game for Japanese, but actually the birthplace of Karuta is Portugal.* [EXP1] [High TOEIC]

Unmarked conventional verbs

A great number of verb forms were used in unmarked conventional senses. Many appeared at low frequencies, but a select number of the more common uses are plotted in Figure 5.30. In the figure, verbs have been marked according to whether they were taught or untaught target metaphors or non-target metaphors. Like targeted noun forms, some target verbs in the experimental group appear to have higher rates of usage than the control, but generally the taught and untaught target metaphors and non-target forms are distributed reasonably equally across the conditions. This suggests that experimental group learners were both picking up and putting aside particular taught target expressions, and that they were apparently not extending the knowledge of metaphor themes to produce their own untaught target metaphors more than learners in the control group.



Note: Minimum frequency: 10 occurrences in either corpus
Note: Significant values (log-likelihood ≥ |3.84|) appear in bold italic font.

Figure 5.30. Scatter plot of open-class unmarked conventional metaphorical verb use across ability range and conditions

The words that were favoured in the experimental condition tended to be simple, quite high frequency words that all learners would probably already have

had some knowledge of, such as *import*, *spend*, *copy*, and *enter*. Thus, it is quite easy to imagine low or mid-proficiency learners having little difficulty in putting these words to use in their new senses. Examples (77 - 80) show learners employing these terms to describe cultural flow. While it is clear that these learners have not yet mastered the appropriate phraseological patterning, these examples may be evidence of the first steps that learners take in broadening their lexical repertoire.

(77) Foreign cultural entities are importing. Many people have various opinions about this. [EXP13] [Mid TOEIC]

(78) I think maybe it is difficult for me to spend life vigorously in hard environment like air raid. [EXP23] [Low TOEIC]

(79) Japan have copied other country's culture, but they didn't just copy and they adapted it to suit their own cultural needs. [EXP18] [Mid TOEIC]

(80) I think that foreign cultures should enter in Japanese culture, otherwise meal for use a knife and fork, or foreign music, clothes, and so on were not entered in Japan. [EXP20] [Low TOEIC]

Figure 5.30 shows that learners from across the range of ability covered by the study make use of unmarked metaphorical verbs in their writing. There is also a split in the types of metaphor along the ability axis. All taught target verbs were either favoured by lower ability learners or were clustered around the zero-point of the vertical axis, while the verbs that significantly favoured high ability learners were non-taught target forms and the untaught target metaphor *accept*. The most likely explanation for this is that explicit instruction in this study made target forms more

accessible for learners of all proficiencies, which effectively allowed lower ability learners to catch up with higher level classmates.

As with nouns, we can see in Figure 5.30 that the forms which high ability learners made most use of are often surprisingly simple, relatively frequent words. Verbs such as *bring*, *say*, and *see* were overwhelmingly favoured by high-level learners (81 - 83). It is interesting to find that when looking at productive language, one of the markers of more advanced proficiency is learners making greater use of the extended meanings of certain extremely frequent words, rather than just using a wider variety of low frequency words.

(81) Overall, the effects of cultural hybridity and hybridism bring evolutions into Japan. [CON14] [High TOEIC]

(82) The Japan Times News in June 7, 2013 says that British and U.S. music no longer dominates the global music market... [CON1] [High TOEIC]

(83) We can see hybridism in terms of Japanese belief. [EXP17] [Mid TOEIC]

Examples (84 - 87) are presented to show language that was not explicitly taught in class. The use of *introduced* to describe cultural borrowing appears to be something that many learners were familiar with. Despite not being taught, it appeared in the output of learners of all abilities in both conditions. A possible reason for its greater use in the control condition is that experimental group learners were provided with alternative expressions which could fulfil the same function, such as *enter* or *import*. In fact, in both conditions, 13 of the 23 participants (57%) produced a form of *introduce*, but control group learners did so at an average of 3.31 uses per person, while experimental group members produced *introduce* 1.77 times

per person. That the same proportion of experimental group learners demonstrated an ability to use this word form but did so at just over half the rate of the control group lends weight to this hypothesis.

The high rate of use for forms of *adopt* by lower level learners relative to those of higher proficiency is something of a mystery. It does not exist as a loanword in Japanese, nor is it particularly high frequency. Neither *adopt* nor *suit* (85 - 87) were taught in class, so their being significantly favoured by particular groups serves as a reminder that there will always be a certain randomness in any data set. Learners develop knowledge from all of their interaction with a language, and it is possible that the use of these words was driven by experiences outside of this classroom.

(84) During Meiji era, many foreign cultural entities were introduced into Japan. [CON21] [Mid]

(85) We're adding things to our culture but is no change, because Japan adopt a foreign style without change a Japanese style. [EXP12] [Low]

(86) Although Japanese people adopted the Western elements in Japanese weddings, they sometimes misunderstand Western style. [CON23] [Mid]

(87) Sumo has changed a lot to suit needs of young generation. [EXP7] [High]

This section has examined the production of parts of speech and rhetorical functions across the two conditions and the range of learner abilities. One finding was that when viewed at this fine level of detail, the data reveal a mixed effect of teaching on learner output. In some cases, experimental group learners made more extensive use of target forms, as might be expected, and this allowed lower level learners in particular to produce a wider range of metaphor than they likely would

have without the treatment. Other outcomes suggest that experimental group learners may have actually reduced their use of particular target words relative to the control group. The teaching effect does not only boost output - there is also the possibility that learners' overuse of structuring metaphors was 'corrected' somewhat in the experimental condition, although the reasons for this are not clear. Another finding was that while learner ability has an effect on output, this does not always occur in predictable ways. The fact that higher level learners generally made more use of evaluative metaphors has been suggested in other studies, but this section has drawn attention to the increased use of relatively basic words in their metaphorical senses as a marker of high ability learners.

5.7. Concluding comments

This chapter has expanded on the findings of chapter four by considering some of the variables that may influence metaphorical output. Firstly, learner ability measures were used to filter the data and reflect on how metaphorical production develops as learners gain in proficiency. The second variable to be considered was the method of instruction, with the output from a more typical content-based course compared with that of a course with a strong focus on metaphor awareness raising. Various elements of the written output were then analysed with these two variables in mind. Production of metaphor was examined by looking at parts of speech, target language, and rhetorical function. Finally, the analysis shifted focus down to the level of individual word forms and their relationship with learner ability and classroom instruction.

The first main finding of this section was that under typical classroom conditions, metaphor production generally increases with learner ability. By and large, this holds true for all aspects of output that were considered, from all MRWs to open-class metaphors, target language and the use of metaphor for various rhetorical functions. However, this increase is not perfectly linear; once learners reach a level of around B1 on the CEFR or somewhere in the 500 - 700 range on the TOEIC test, there appears to be a levelling-off in the rate at which metaphor appears in output. It has been suggested that this is a period in which learners, having already acquired the ability to produce some conventional metaphorical language, then begin to diversify their range of use and to experiment with other rhetorical functions rather than to keep producing more metaphor (Littlemore et al., 2014, p. 133). The findings of this chapter support this hypothesis, and may in fact strengthen it since the data in this study allowed for the writing topic to be more tightly controlled than was possible in previous studies based on examination writings.

It appears, however, that the experimental awareness-raising condition disrupts this trend somewhat. Lower-level learners (those in the A2 CEFR band or with TOEIC scores less than 500), who are least likely to produce metaphor in regular settings, appear to gain the most 'benefit' in terms of increased quantity of production. This leads to the narrowing of the spread of values seen in Figure 5.1 for the experimental condition. On average, mid- and high-level learners increased their production of taught target metaphors over that of the control group, but the actual amount of overall metaphor they produced was usually close to, or even less than, that of control group learners of equivalent ability.

This lack of a clear gain for mid- and high-level learners in the experimental group over the control, in conjunction with the levelling-off in metaphor production for intermediate-level learners that was seen in both conditions, led to a hypothesis that the experimental treatment had a differential effect on learners. It was proposed that more proficient learners, being able to produce metaphor to some degree even without the treatment, might have devoted their attention to adding variety to their output instead.

This was to some degree confirmed by the analyses. Although the differences were non-significant, higher-level students produced a greater variety of metaphorical word forms than low- or mid-level learners. The difference between the control and experimental groups again showed a stronger effect for low-level learners; however, all groups were affected. Thus, one tentative conclusion of this chapter is that the rate of increase in metaphor production will gradually slow down as learners develop, being replaced by a growth in variety of metaphor use. This appears to hold true for learners in this particular study. However, one limitation is that very few participants were in the high-intermediate or advanced range; there were only three B2-level participants, and none in the C1 or C2 levels. For this reason, it would be very useful to carry out a similar study with more proficient learners.

When the analysis moved on to more narrowly-defined elements of language, the complexity of the situation became more apparent. Clear differences could be seen in the usage of target forms in the two conditions. Control group learners, for whom there was essentially no difference between taught and untaught target metaphors, produced these types in roughly equal quantities. In the experimental condition, we see gains in production of taught target language; however, this

comes at a cost, as untaught forms decrease in production. Whether this inhibits language development or not is debatable, though. Certainly, it could be argued that there is convincing evidence for the benefits of explicit instruction in target language forms, whether they are grammatical, lexical or metaphorical (Ellis, 2002; Laufer, 2005), and that emphasising any language feature by necessity entails de-emphasising others. With regard to metaphor instruction, however, there remains the question of whether learners were producing language as a result of their greater awareness of metaphor themes or simply because they were reproducing the vocabulary items that had appeared in class. The data here serve as a reminder that instruction should allow a place for learners to push at the boundaries of their knowledge and to strive to make use of language in new ways.

The findings of the analysis of rhetorical functions showed that in this particular genre and context, interpersonal metaphors play only a very limited role. Evaluative and structuring metaphors were found in the output of all learners, but the overwhelming majority of uses were unmarked conventional metaphors. This finding was expected, as learners were mostly writing in response to the themes of the course content, and the target language was selected in order to help learners express abstract concepts in a more natural fashion rather than to fulfil evaluative or text structuring functions.

Again, the gains that the experimental group made in one area, unmarked conventional metaphors, have to be balanced with a drop in another area, in this case, structuring metaphors. However, instead of considering this a loss, it might be seen as more of an adjustment to more closely reflect the norms of academic discourse in English. The analysis showed that one part of speech tended to

dominate each function; adjectives performed most evaluative functions, pronouns or determiners did most textual structuring, and verbs were the primary drivers of unmarked conventional language. This may be a natural feature of language - these parts of speech are carrying out their most typical role. There is also a suggestion, however, that this degree of dominance shifts with ability. In the case of structuring metaphors, pronouns and determiners actually became even more prevalent, with lower level learners' dependence on fixed adverbial phrases such as *and so on*, *on the other hand* and *in summary* gradually diminishing. Therefore, rather than signifying a loss in textual coherence, this drop in structuring metaphors may be evidence of learners gradually letting go of their 'lexical teddy bears' (Hasselgren, 1995).

This focus on the distribution patterns of individual words highlighted language systems in considerable flux and warned of the limitations of simplistic models of learner development (de Bot, Lowie, & Verspoor, 2007; Larsen-Freeman, 2006). For each of the rhetorical functions, it could be seen that among the forms that were most favoured by high ability learners were items such as *way*, *power*, *this*, *that* and *bring*. While it is true that advanced learners will typically show greater knowledge of infrequent words, it should be noted that they also appear to make more varied use of some extremely high frequency words and are less dependent on recycling the same safe options in their output. Lexical knowledge is frequently described using the breadth and depth distinction, which corresponds to a larger number of words known and greater knowledge about those words. But what this distinction hides is that when we examine learner production, less can often be more, as

learners are weaned off their favourite expressions in favour of a broader range of use.

One caveat to this argument is that from an empirical point of view, it is only to be expected that a distribution of data will contain a cluster of scores around the mean and a smaller number at the extremes. Words that strongly favour one ability group or condition could be no more than statistical outliers. However, other studies have shown that development in learner vocabularies (Bell, 2009, p. 116 - 127; Caspi & Lowie, 2010, p. 56 - 57; Groom, 2009, p. 23 - 24) is characterised by periods of fossilisation, fluctuation, and overuse, and the findings here are entirely consistent with this.

Clearly, developments in metaphorical output are a complex matter. This issue will continue to be explored in the next chapter by bringing two other variables into the analysis: classroom input and time.

CHAPTER 6. THE INTERACTION BETWEEN CLASSROOM INPUT AND LEARNER OUTPUT FOR METAPHORS OF CULTURE

6.1. Introduction and rationale for the study

The previous chapters have considered learners' written output with progressively finer-grained levels of analysis. At each step, the focus on a different aspect or sub-set of the data has brought new observations to light, and it is clear that a broad-scale investigation of learner development risks masking the more subtle changes that do not apply to all individuals or language forms in the data set.

Despite the attempts made to control for the effects of writing topic in this study, there are clearly a great many other variables at work here. Regarding the learners, it has been suggested that there is an interaction between proficiency and developments in metaphor production, with learners apparently shifting from increasingly greater raw usage of metaphor towards a widening of their lexical and functional range of use. This broad developmental pattern has been suggested by other studies on metaphor production in writing (Littlemore et al., 2014, p. 133) as well as investigations into spoken production of polysemous word forms (Crossley, Salsbury, & McNamara, 2010). At the same time, individual language forms themselves may be more or less amenable for learners of different abilities to apply in their metaphorical senses. It is quite possible that the experimental treatment had

the effect of increasing use of a particular word for one learner while at the same time reducing it for another.

One important variable that has not yet been considered is the input learners were exposed to from instructors' language and the course materials. It has been noted that in the control group, instruction focused on teaching academic and high-frequency vocabulary, while teaching in the experimental condition brought in activities to raise awareness of prominent metaphorical themes in the course content. But without considering the degree of emphasis given to metaphor in general or individual metaphors, it is difficult to make claims about the relative success or otherwise of the treatment in encouraging learners to diversify their lexical output.

This leads to a consideration of the fundamental relationship between pedagogy and language acquisition. In strictly objective terms, one can never be certain that a given classroom activity is the primary agent responsible for actual learning. Lacking direct access to the mental processes that constantly restructure learners' linguistic knowledge, we are reliant for the most part on instances of language performance for data, from which we extrapolate that changes in individual behaviour or differences relative to some other group are evidence of learning. Prabhu (1995, p. 59 - 60) presents two differing conceptions of how learning might relate to instruction. The first assumes that linguistic rules have some psychological validity - that the essence of learning is to capture those rules and apply them as the situation warrants. The second treats linguistic rules as approximations of unknowable mental processes. In this view, learning takes place through abstraction of language input as learners generalize and refine their understanding through experience. The two

views are not mutually exclusive, however; it is quite possible to regard explicit instruction as helping learners towards more effective abstraction while acknowledging that it is not so much that rules are being acquired as supportive evidence provided. This seems the most appropriate view with which to approach the teaching of metaphorical vocabulary. Such language lends itself not to hard and fast rules, but to tendencies and patterns. What is necessary then, is to consider the elements of teaching that might best facilitate the development of metaphoric competence.

Ellis (2006a, 2006b) identified frequency and recency of occurrence as two of the most important factors that influence human learning. Learners are more likely to remember language they have recently encountered, and that which is not met again stands a greater chance of being dismissed as less useful. Therefore, it is also necessary to consider the classroom data longitudinally, taking note of the amount and timing of exposure to metaphorical forms and how they relate to learners' output of such forms. One issue here is that while frequency has been claimed many times to have a powerful influence on learning, and on lexical development in particular (e.g., Conklin & Schmitt, 2012, p. 50 - 56; Meara, 2010, p. 5 - 6; Milton, 2009, p. 241 - 242; Perfetti & Hart, 2002, p. 190 - 192, Rott, 1999), it is important not to oversimplify what this means. The concept of frequency can be operationalized in different ways, and studies have found differential outcomes depending on how it is defined. For example, it has been found that receptive learning gains are much greater for high frequency words than for less-frequent vocabulary. In other words, the frequency effect on vocabulary learning follows a power law, rather than displaying a linear trend (Dekeyser, 1997; Dóczi & Kormos, 2016, p. 30 - 34). Further

to this, Ozturk (2015, p. 106) found that while receptive vocabulary growth could be modelled by frequency data based on native-speaker corpora, developments in productive vocabulary were influenced more by the demands of the context in which learners found themselves. This implies a need for studies to consider frequency in more localised terms, one of the goals of this chapter.

Additionally, other factors beyond frequency and recency should be considered because, as Ellis (2006a, p. 15 - 17, 2006b, p. 170 - 173) notes, many highly frequent, regularly occurring language features are among the most difficult for learners to acquire. The reason for this, he argues, is that such features lack salience - aspects of their meaning or use are easily overlooked when learners are focussed on the broader meaning. It follows then, that increasing the salience of language features is an important step in encouraging their acquisition. This is the position taken in several papers on vocabulary learning, where the benefits of explicit instruction in lexical items over incidental learning have been well documented (Hulstijn & Laufer, 2001; Laufer, 2005; Laufer & Rozovski-Roitblat, 2015). Vidal's (2011) study into learning vocabulary in an academic environment also emphasised the degree of semantic elaboration as an important factor to consider, as well as finding that higher levels of learning were achieved from studying written material over listening to lectures, especially for lower proficiency learners (p. 242 - 246).

As well as features of the input, several other factors related to words themselves have been identified as contributing to the ease or difficulty of word learning. Laufer (1997) lists several of these, including those related to orthographic, phonological, or morphological form, similarity to other words, part of speech, and semantic features such as abstractness, specificity, idiomaticity, and polysemy. Other

studies have examined word concreteness, the degree to which words can be experienced with the senses, imageability, how easily words can be called to mind, and similarity to L1 terms (de Groot & Keijzer, 2000; Paivio, 1969). These factors have all been shown to positively influence word learning, and regularity of form also appears to have a facilitative effect (Laufer, 1997, p. 144). Regarding part of speech, it is claimed that nouns, owing to their being more concrete, are easier to learn than verbs or adjectives (Crossley, Subtirelu & Salsbury, 2013, p. 744; Rodgers, 1969, p. 331). Crossley, Subtirelu, and Salsbury go on to note that while acquisition of nouns to a productive level is predicted through frequency of exposure, the learning of verbs, with their greater degree of abstractness and multiple meanings, is more strongly influenced by exposure in different contexts, rather than raw frequency of exposure (p. 744 - 745). This finding is an important reminder of the need to consider learning in a broad sense. While this study attempts to provide a rich level of detail for a single-class context, it should not be forgotten that experiences outside of this environment might have influenced learner output. Chapter seven will address external effects, as well as the influence of part of speech, learners' L1, and word frequency as estimated by native-speaker corpora. Chapter six will pick up on the finding of Ozturk (2015, p. 106) discussed earlier that localised contextual input is an important, but perhaps overlooked, consideration when examining developments in productive vocabulary.

However, one issue that was identified earlier was the effect of topic on output. The production of individual metaphors varied greatly, as was shown in Table 4.3 (page 140), and several metaphors only appeared at low frequencies. Metaphoric themes refer to broader concepts, and clearly not all concepts appear with the same

frequency; some will only be brought up by particular writing topics, while others are much more general. By way of an example, language representing the metaphor BEING INTERESTED IS ENTERING SOMETHING appeared 11 times in output across the study, and 82% of those uses were in a single writing topic on Japanese sub-cultures. The strong relationship between writing topic and concepts is a challenge to investigations of longitudinal development since learners cannot be expected to make use of metaphors that are not related to the topic at hand. Because of this, it was decided to narrow the range of focus to metaphors that were used to describe culture itself. Since this was the main theme of the entire course, these metaphors were both relatively common and reasonably dispersed across the data set.

This chapter will investigate the language used to describe culture and cultural change in both conditions of the study. In order to model the linguistic environment of the two courses as accurately as possible, the language from instructors speaking in class and from both written and audio-visual course materials was transcribed to create input corpora. These were combined with the output corpora to create diachronic records of both input and output that could be used to consider the complex interactions between input frequency, dispersion and metaphor production. The specific research questions for this chapter are:

a) Would the sustained approach to metaphor awareness raising taken in the experimental condition lead to measurable differences in input in terms of the quality, frequency, and regularity of occurrence of target metaphors of culture?

b) Did the experimental treatment lead to measurable differences in learners' output of metaphors of culture in terms of quality, frequency or regularity of production?

c) Did learner output of metaphors of culture in either condition increase over time in terms of frequency or variety? Did classroom input have a cueing effect on learner output in either condition?

6.2. Methodology

6.2.1. Classroom input

The recording, transcription and annotation of the input corpora were carried out as described in section 3.6.2 (page 106). As these corpora taken together comprised over half a million words, analysis with MIPVU was not realistic. Instead, the corpora were stored as text files and semantic tags were added using Wmatrix. The tagged corpora could then be used to identify target metaphors as was explained in section 3.6.2.1 (page 112).

Once metaphors for culture had been identified in the input, each metaphor and the immediate context in which it appeared was listed in an Excel file along with the date on which it was produced, and whether it was a taught or untaught target metaphor. In order to examine the quality of the metaphorical input learners were exposed to, the original course transcripts were checked, and each instance was assigned to one of the following five categories:

- Aural (video) – language that appeared in the course video transcripts.
- Aural (teacher talk) – language that teachers used in discussing course content, without emphasis on metaphor instruction.
- Aural (explicit) – language that teachers used when there was an explicit focus on metaphor.
- Written (text) – Language that appeared in written text with a focus on course content, rather than metaphor.
- Written (explicit) – Language that appeared in written texts with an explicit focus on metaphor.

These five categories covered the various ways in which learners were exposed to metaphors in classroom input. The first, second, and fourth categories covered different forms of implicit input, while the third and fifth categories represented measures of explicit input.

6.2.2. Classroom output

The classroom output data used in this chapter are those based on the expanded list of polywords that were identified in dictionaries. That is, the same data set as was used in the rhetorical function analysis of the previous chapter. Metaphors for culture that had been identified with Steen's (1999) five-step process were extracted from each output corpus along with sufficient context to allow for comprehension. These were then copied into an Excel file for analysis, with tags added for participant, language ability, writing topic and the date of the writing sample, as well as whether each metaphor was a taught or untaught target metaphor.

6.3. Metaphors used to describe culture in the course input

The overall distributions of metaphor themes to describe culture were somewhat similar across the two conditions (Figure 6.1). The most noticeable difference was the increase in use of the CULTURES ARE SPACES and CULTURES ARE SUBSTANCES metaphor themes, which seems to have come at the expense of language exemplifying the CULTURES ARE LIVING THINGS and CULTURES ARE CONSTRUCTIONS themes. These differences were more apparent when the quantity of metaphorical input was examined. Figure 6.2, in which the frequency values have been normalised to a per-10,000-word rate, shows that input for most of the metaphor themes related to culture was greater in the experimental condition. The data in Table 6.1 suggest that quantity of input was significantly higher for four of these themes, while there was a non-significant drop in input of the CULTURES ARE CONSTRUCTIONS theme. In combination, the effect of the different style of instruction was that learners went from encountering these metaphors once every 433 words in the control group input to once every 285 words in the experimental condition. Full data is presented in Appendix L (page 478).

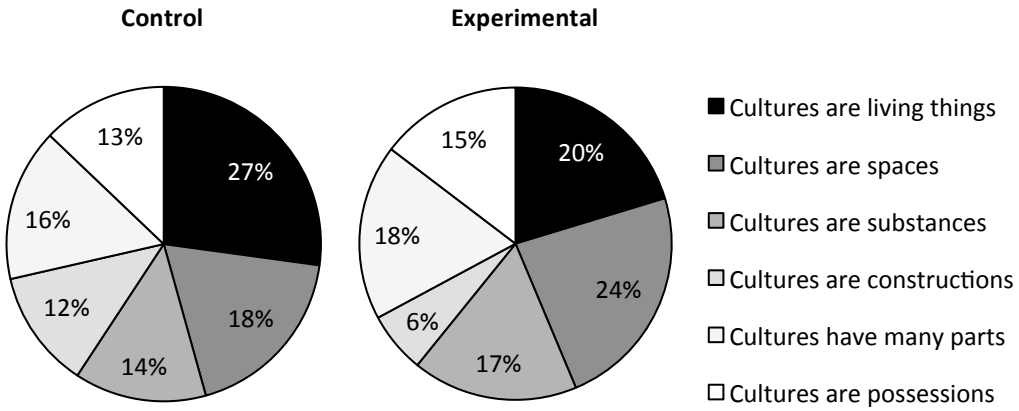


Figure 6.1. Proportions of culture metaphors grouped by theme in course input

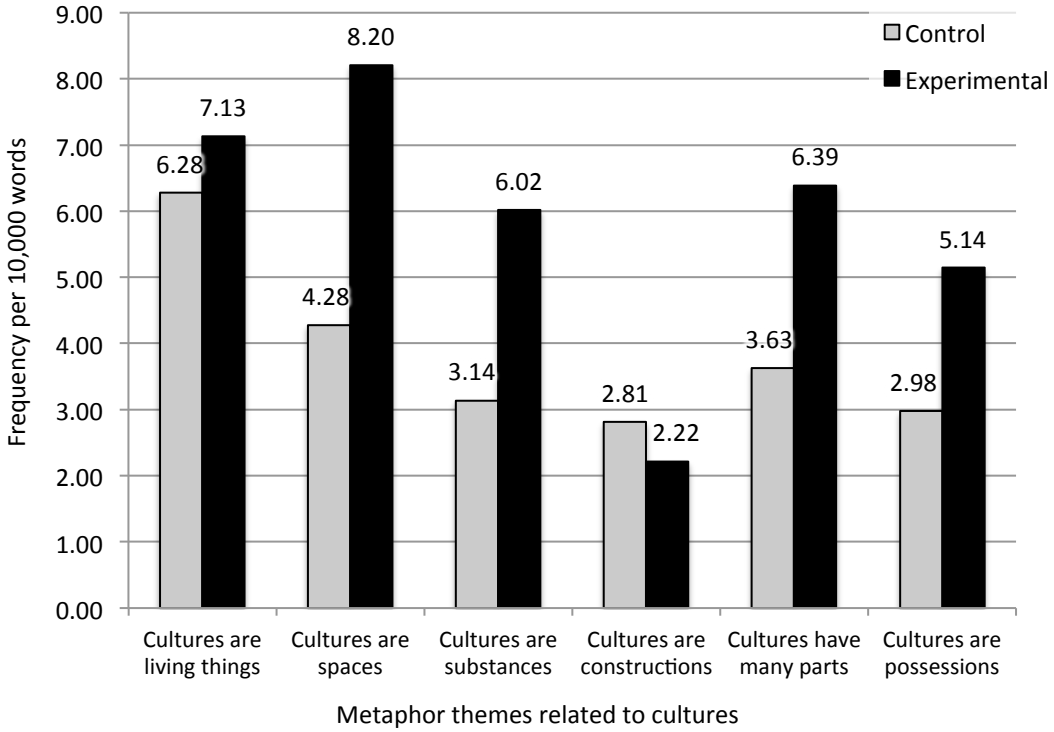


Figure 6.2. Normalised frequencies of culture metaphors grouped by theme in both input conditions

Table 6.1. Log-likelihood tests for culture metaphor themes appearing in input between the two conditions

Metaphor theme	Log-likelihood [a]	Bayes factor effect size
CULTURES ARE LIVING THINGS	-1.45	-11.75
CULTURES ARE SPACES	-33.40****	20.20^{+++}
CULTURES ARE SUBSTANCES	-24.53****	11.32^{+++}
CULTURES ARE CONSTRUCTIONS	1.89	-11.31
CULTURES HAVE MANY PARTS	-20.55****	7.34^{++}
CULTURES ARE POSSESSIONS	-15.63****	2.43^{+}
Total	-66.46****	53.26^{+++}

Note: Input corpora sizes are provided in Appendix F (page 440). Raw frequencies are provided in Appendix L (page 478).
a = Negative values indicate underuse in control condition relative to experimental condition.
**** = Significant at the .0001 level
+ = Positive evidence against the null hypothesis; ++ = Strong evidence; +++ = Very strong evidence

As was stated in section 3.5 (page 94), the selection for metaphors that would be explicitly taught in the experimental condition was partially based on an

examination of the materials and classroom instructions in the control condition. The question of whether learners would be able to extend target metaphors by producing words that they had not been taught was also of interest, and this lead to the distinction between taught and untaught target metaphors. As can be seen in Figure 6.3, words selected for explicit instruction formed the bulk of input in both conditions, and the effect of the experimental treatment appears to have been a large increase in production of these words (log-likelihood: -65.77, effect size: 52.57 (very strong)). The frequency of untaught metaphors also rose slightly, but not significantly in the experimental condition (LL: -2.20, effect size: -11.01 (very strong evidence *in favour of* the null hypothesis)).

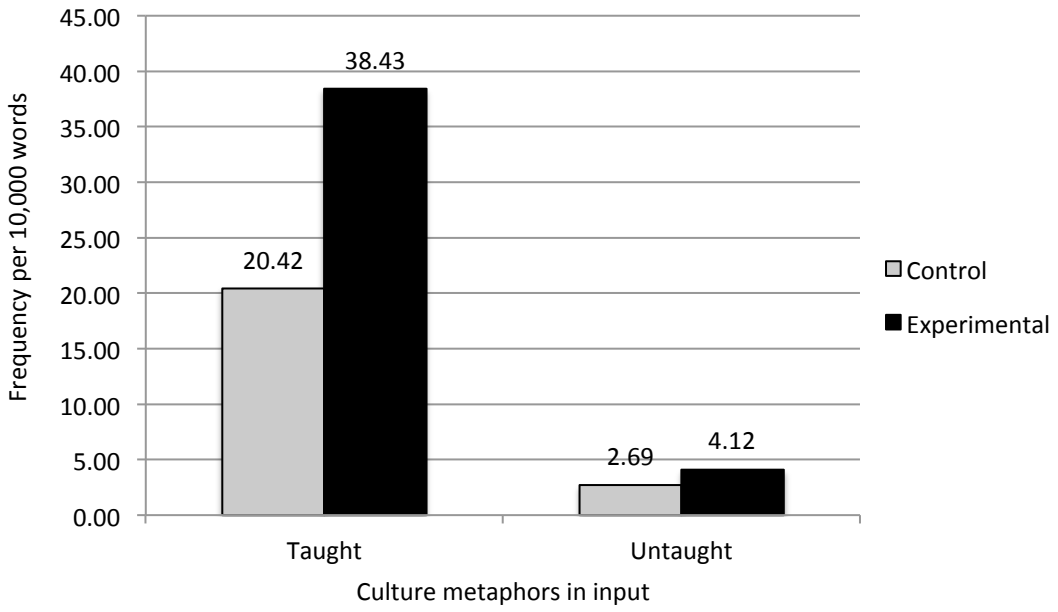


Figure 6.3. Normalised frequencies of taught and untaught target culture metaphors appearing in input in the two conditions

The treatment clearly prioritised the taught target forms, which consequently appeared far more regularly in input, but factors other than frequency are likely to have a powerful effect on vocabulary learning. Ellis (2002, p. 161 - 163; 2006a, p. 5 -

6) discussed the recency of an event as having a powerful effect on its likelihood of being recalled, and it is well noted in the literature that repetition and review are crucial elements of lexical development (Waring & Takaki, 2003; Nation, 2013, p. 107 - 110). Estimates vary for the amount and the timing of encounters necessary to develop word knowledge, but it is widely recognised that meeting the same word in various contexts is essential to build up awareness of different aspects of knowledge (Cobb, 1997, p. 313 - 314; Pellicer-Sánchez, 2016, p. 122 - 125; Webb, 2007). A word that appears 20 times in a single lesson but never again that semester might have a greater chance of being forgotten than a word appearing 20 times across several lessons. Accordingly, the input data were examined to determine how many lessons each target word appeared in. This provided a measure of the dispersion of the word forms within each input corpus, which reflected the degree to which the target words in each metaphor theme were re-appearing and consequently promoting recall.

Tables 6.2 - 6.7 provide the frequency and dispersion of the 49 taught target words which were part of the metaphor themes related to cultures and 29 untaught target words which appeared in input. In all but four cases, the frequency of exposure for taught target metaphors was higher in the experimental condition, and log-likelihood values indicated that 18 of the taught forms appeared significantly more often. Since the corpora are relatively small, however, the only strong effect size found was for the increase in input of the word *spread*. For untaught target metaphors, rates of input were much more similar between the conditions. 35% of the taught target words appeared at least 10 times in the control condition, and this rose to 67% for the experimental group, so it does appear that in many cases

learners were exposed to greater amounts of incidental input (although it should be noted that some occurrences were simple repetitions, as noted in section 3.6.2.1, page 112). Mann-Whitney tests for the normalised frequencies of target words in input between the conditions showed a significantly higher rate of input for taught metaphors for cultures, although the effect size was small (Table 6.8).

Table 6.2. Log-likelihood tests for input frequency of CULTURES ARE LIVING THINGS metaphors and dispersion over lessons in the two conditions

Target word	Frequency per 10,000 words		Log-likelihood [a]	Effect size (Bayes factor)	Dispersion (% of lessons)	
	Con.	Exp.			Con.	Exp.
Taught target forms						
alive (adj)	0.08	0.20	-1.40	-11.81	3.45%	9.68%
evolution (n)	0.00	0.20	-7.22**	-5.99	0.00%	9.68%
evolve (v)	0.53	0.81	-1.55	-11.66	20.69%	25.81%
grow (v)	0.33	1.28	-15.91****	2.71+	6.90%	32.26%
hybrid (adj)	2.12	2.52	-0.94	-15.95	31.03%	29.03%
hybrid (n)	2.49	1.58	5.51*	-7.69	20.69%	19.35%
Untaught target forms						
birth (n)	0.00	0.03	/	/	0.00%	3.23%
birthplace (n)	0.08	0.13	/	/	3.45%	3.23%
growth (n)	0.45	0.17	3.61	/	17.24%	9.68%
survival (n)	0.04	0.07	/	/	3.45%	3.23%
survive (v)	0.16	0.13	/	/	3.45%	3.23%

a = Negative values indicate underuse in control input corpus relative to experimental input corpus. Minimum frequency 5 occurrences in either corpus.
* = Significant at the 0.05 level; ** = 0.01 level; **** = 0.0001 level
+ = Positive evidence against the null hypothesis

Table 6.3. Log-likelihood tests for input frequency of CULTURES ARE SPACES metaphors and dispersion over lessons in the two conditions

Target word	Frequency per 10,000 words		Log-likelihood [a]	Effect size (Bayes factor)	Dispersion (% of lessons)	
	Con.	Exp.			Con.	Exp.
Taught target forms						
come into (ph v)	0.98	2.19	-12.55***	-0.65	34.48%	51.61%
enter (v)	0.04	0.34	-6.92**	-6.29	3.45%	12.90%
export (n)	0.29	0.50	-1.64	-11.56	6.90%	6.45%
export (v)	0.00	0.10	-3.61	-9.60	0.00%	9.68%
import (n)	0.20	0.24	-0.06	-13.41	10.34%	9.68%
import (v)	0.73	0.94	-0.69	-12.51	20.69%	25.81%
inside (prep)	0.08	0.10	/	/	6.90%	3.23%
outside (adv)	0.12	0.30	-2.10	-11.11	3.45%	9.68%
outside (n)	0.45	0.40	0.06	-13.14	13.79%	12.90%
outside (prep)	0.53	0.77	-1.22	-11.98	27.59%	25.81%
outsider (n)	0.08	0.34	-4.39*	-8.81	3.45%	16.13%
surround (v)	0.08	0.40	-6.13*	-7.07	3.45%	12.90%
take in (ph v)	0.53	0.54	0.00	-13.20	17.24%	12.90%
transcend (v)	0.00	0.34	-12.03***	-1.17	0.00%	3.23%
Untaught target forms						
bring into (ph v)	0.00	0.10	/	/	0.00%	9.68%
contain (v)	0.04	0.20	-3.06	-10.14	3.45%	9.68%
exporter (v)	0.00	0.03	/	/	0.00%	3.23%
go into (ph v)	0.00	0.10	/	/	0.00%	3.23%
leave (v)	0.00	0.03	/	/	0.00%	3.23%
outside (adj)	0.00	0.17	-6.02*	-7.19	0.00%	6.45%
penetrate (v)	0.04	0.03	/	/	3.45%	3.23%
take into (ph v)	0.04	0.03	/	/	3.45%	3.23%
within (prep)	0.04	0.00	/	/	3.45%	0.00%

a = Negative values indicate underuse in control input corpus relative to experimental input corpus. Minimum frequency 5 occurrences in either corpus.

* = Significant at the 0.05 level; ** = 0.01 level; *** = 0.001 level

Table 6.4. Log-likelihood tests for input frequency of CULTURES ARE SUBSTANCES metaphors and dispersion over lessons in the two conditions

Target word	Frequency per 10,000 words		Log-likelihood [a]	Effect size (Bayes factor)	Dispersion (% of lessons)	
	Con.	Exp.			Con.	Exp.
Taught target forms						
absorb (v)	0.61	0.77	-0.51	-12.70	13.79%	12.90%
flow (n)	0.12	0.54	-7.44**	-5.77	6.90%	19.35%
flow (v)	0.00	0.03	/	/	0.00%	3.23%
fusion (n)	0.12	0.47	-5.76*	-7.44	3.45%	16.13%
mix (n)	0.00	0.17	-6.02*	-7.19	0.00%	9.68%
mix (v)	0.20	0.71	-7.75**	-5.46	6.90%	16.13%
mixture (n)	0.86	0.50	2.49	-10.71	13.79%	16.13%
shape (n)	0.12	0.27	-1.50	-11.71	3.45%	6.45%
shape (v)	0.04	0.57	-14.32***	1.11	3.45%	6.45%
spread (v)	0.20	1.14	-18.97****	5.77^{+}	13.79%	38.71%
Untaught target forms						
combination (n)	0.08	0.07	/	-13.17	6.90%	6.45%
combine (v)	0.20	0.10	0.96	-12.24	13.79%	6.45%
form (n)	0.57	0.64	-0.10	-13.10	3.45%	3.23%
refined (adj)	0.00	0.03	/	/	0.00%	3.23%

a = Negative values indicate underuse in control input corpus relative to experimental input corpus. Minimum frequency 5 occurrences in either corpus.

* = Significant at the 0.05 level; ** = 0.01 level; *** = 0.001 level; **** = 0.0001 level

+ = Positive evidence against the null hypothesis

Table 6.5. Log-likelihood tests for input frequency of CULTURES ARE CONSTRUCTIONS metaphors and dispersion over lessons in the two conditions

Target word	Frequency per 10,000 words		Log-likelihood [a]	Effect size (Bayes factor)	Dispersion (% of lessons)	
	Con.	Exp.			Con.	Exp.
Taught target forms						
construct (v)	0.24	0.34	-0.39	-12.82	13.79%	19.35%
fit (v)	0.04	0.07	/	/	3.45%	6.45%
fit into (ph v)	0.16	0.24	-0.35	-12.85	10.34%	9.68%
make up of (ph v)	0.00	0.07	/	/	0.00%	6.45%
manufacture (v)	2.28	1.45	5.13*	-8.07	10.34%	3.23%
undamaged (adj)	0.08	0.07	/	/	6.90%	6.45%

a = Negative values indicate underuse in control input corpus relative to experimental input corpus. Minimum frequency 5 occurrences in either corpus.
* = Significant at the 0.05 level

Table 6.6. Log-likelihood tests for input frequency of CULTURES HAVE MANY PARTS metaphors and dispersion over lessons in the two conditions

Target word	Frequency per 10,000 words		Log-likelihood [a]	Effect size (Bayes factor)	Dispersion (% of lessons)	
	Con.	Exp.			Con.	Exp.
Taught target forms						
aspect (n)	0.73	1.01	-1.17	-12.04	17.24%	25.81%
component (n)	0.04	0.34	-6.92**	-6.29	3.45%	12.90%
divide (v)	0.08	0.13	/	/	3.45%	9.68%
element (n)	1.14	1.75	-3.43	-9.77	34.48%	41.94%
ingredient (n)	0.16	0.44	-3.44	-9.76	3.45%	9.68%
part (n)	1.18	2.39	-11.04***	-2.16	44.83%	61.29%
Untaught target forms						
add (v)	0.20	0.24	-0.06	-13.14	3.45%	6.45%
piece (n)	0.08	0.10	/	/	6.90%	6.45%

a = Negative values indicate underuse in control input corpus relative to experimental input corpus. Minimum frequency 5 occurrences in either corpus.
** = Significant at the 0.01 level; *** = 0.001 level

Table 6.7. Log-likelihood tests for input frequency of CULTURES ARE POSSESSIONS metaphors and dispersion over lessons in the two conditions

Target word	Frequency per 10,000 words		Log-likelihood [a]	Effect size (Bayes factor)	Dispersion (% of lessons)	
	Con.	Exp.			Con.	Exp.
Taught target forms						
borrow (v)	0.57	1.24	-6.80**	-6.40	20.69%	12.90%
claim (v)	0.04	0.44	-10.02**	-3.18	3.45%	6.45%
lose (v)	0.90	1.04	-0.29	-12.91	17.24%	12.90%
share (v)	0.37	0.74	-3.41	-9.80	10.34%	9.68%
shared (adj)	0.00	0.03	/	/	0.00%	3.23%
steal (v)	0.00	0.17	-6.02*	-7.19	0.00%	3.23%
take (v)	0.45	0.50	-0.09	-13.12	17.24%	19.35%
Untaught target forms						
accept (v)	0.16	0.27	-0.70	-12.50	10.34%	12.90%
adopt (v)	0.04	0.03	/	/	3.45%	3.23%
adoption (n)	0.04	0.03	/	/	3.45%	3.23%
exchange (n)	0.00	0.10	/	/	0.00%	6.45%
give (v)	0.00	0.03	/	/	0.00%	3.23%
have (v)	0.12	0.13	/	/	6.90%	12.90%
keep (v)	0.04	0.13	/	/	3.45%	3.23%
possess (v)	0.12	0.13	/	/	3.45%	6.45%
retain (v)	0.12	0.10	/	/	3.45%	3.23%

a = Negative values indicate underuse in control input corpus relative to experimental input corpus. Minimum frequency 5 occurrences in either corpus.
* = Significant at the 0.05 level; ** = 0.01 level

Table 6.8. Mann-Whitney tests for normalised frequency of culture metaphors appearing in classroom input between conditions

Metaphor category	*n* [a]	Median input freq.		Mean rank		*U*	*Z*	*p*	*r*
		Con.	Exp.	Con.	Exp.				
Taught	90	0.24	0.53	37.89	51.87	692.500	-2.531	.011*	0.27[+]
Untaught	48	0.08	0.12	22.85	25.68	247.000	-.710	.478	0.10[+]
Combined	138	0.16	0.29	61.70	75.68	247.000	-2.049	.040[#]	0.17[+]

a = Culture MRWs with input freq. > 0 per 10,000 words - Control: Taught = 41, Untaught = 20; Experimental: Taught = 49, Untaught = 28
# = Non-significant after Bonferroni adjustment
* = Significant at the 0.05 level
+ = Small effect size

While there does appear to be some evidence for the claim that learners in the experimental group were exposed to a greater quantity of target words, the data presented in Tables 6.2 - 6.7 are less clear on the question of regularity of occurrence. Regularity here was defined as the percentage of lessons in which the target form appeared, giving a measure of the dispersion of the word within each input corpus. While 69% of the taught target words did appear in a greater proportion of lessons in the experimental condition, many of these increases were small, and only 48% of the untaught target culture metaphors were more widely dispersed compared to the control group input. This can be seen in the results of Mann-Whitney tests provided in Table 6.9. There is no sign of a difference in the dispersion of taught target culture metaphors between the conditions, and in fact a drop in the dispersion of untaught culture metaphors with a moderate effect size.

Table 6.9. Mann-Whitney tests for dispersion of culture metaphors across lessons between the conditions

Metaphor category	*n* [a]	Median % lessons		Mean rank		*U*	*Z*	*p*	*r*
		Con.	Exp.	Con.	Exp.				
Taught	90	10.34%	12.90%	43.51	47.16	923.000	-.663	.508	0.07
Untaught	48	3.45%	3.23%	29.25	21.11	185.00	-2.055	.040[#]	0.30[++]
Combined	138	6.90%	9.68%	71.39	68.00	2233.000	-.499	.618	0.04

a = Culture MRWs with dispersion > 0 - Control: Taught = 41, Untaught = 20; Experimental: Taught = 49, Untaught = 28
# = Non-significant after Bonferroni adjustment
++ = Medium effect size

Another way of visualising the regularity of occurrence is provided by Figure 6.4. Here, the percentage of lessons in which any of the target words (taught and untaught combined) in each metaphor theme appeared is plotted for both conditions. The raw number of lessons is also given in each column (control: 29 class

periods in total; experimental: 31). The assumption in this case is that exposure to target words in a theme has the potential to provoke recall of the theme itself; that is, the teaching of metaphorical expressions in unified themes facilitates recall by strengthening mental links between words. There is some fluctuation in the values, but in most cases, the metaphor themes were brought up in 60 - 70% of class periods, which seems a reasonable rate of re-occurrence given that the study was conducted in a regular class with other learning goals to accommodate. There are, however, drops in the dispersion of the CULTURES ARE CONSTRUCTIONS and CULTURES ARE POSSESSIONS metaphors. CULTURES ARE CONSTRUCTIONS was the least frequent of all metaphor themes for cultures in either condition, and its overall rate of input actually declined in the experimental condition, so it is possible that it was simply overlooked in favour of terms that were more prevalent in the course content. It is less clear, however, why the CULTURES ARE POSSESSIONS metaphors should appear in fewer lessons in the experimental condition. As was shown in Figure 6.2, the total frequency of use for this metaphor increased in line with the other themes.

What these data appear to highlight is the challenge of increasing the regularity with which particular metaphors appear while also respecting the nature of content-based learning. In a purely language-focussed course, it might be possible to plan greater repetition into the lessons, but when the input is to a large degree determined by the subject matter, the natural influence of topic on language cannot easily be overcome. Since the conditions in this study had identical content, the course themes probably occurred at approximately the same rate, and when content rather than language was the focus of a class activity, instructors would likely have used what seemed the most natural language to discuss them.

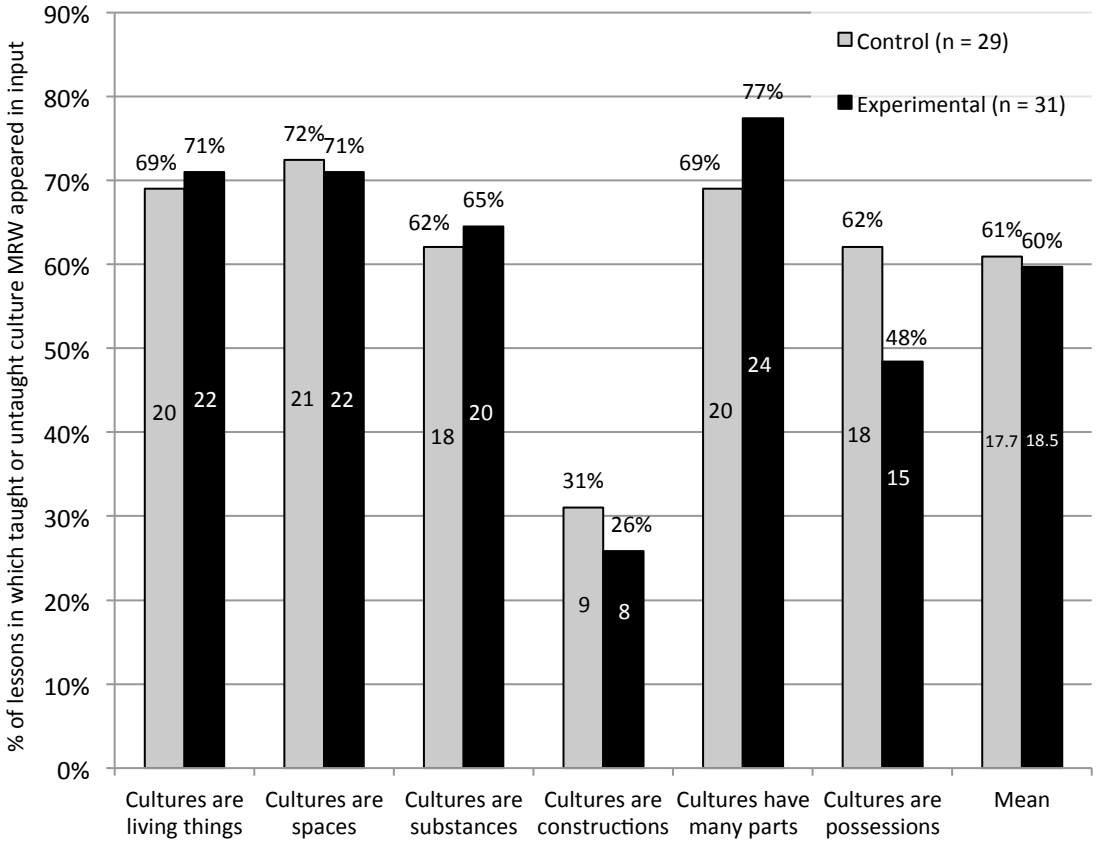


Figure 6.4. Dispersion of culture metaphors across lessons in both conditions

Another possibility is that increasing the range of use of one particular metaphor naturally entails reducing that of another. Culture and cultural change are abstract phenomena, and it is difficult to find literal alternatives to portray them. Assuming that the themes selected here are the most natural ways to describe this topic, then the language environment of the classroom is effectively a zero-sum situation; one cannot use some metaphors more widely without reducing the use of others. The final columns in Figure 6.4 lend support to this conjecture; the average range of use of the combined six metaphor themes in each condition is very similar. The general conclusion is that while raw quantity of metaphor exposure can be manipulated in a classroom intervention, increasing the regularity with which metaphors appear is more difficult, at least when the materials are determined by the course content.

As was mentioned earlier, another factor to consider is the salience of the input. Learners have been shown to acquire greater lexical knowledge from activities that promote noticing and greater depth of processing (Hulstijn & Laufer, 2001), and there is some indication that it is easier to learn from reading than listening for non-advanced learners (Vidal, 2011, p. 242 - 246). To operationalize both input salience and medium, each instance of a culture metaphor appearing in input was classified as either aural input (subdivided into video materials, teacher talk and explicit metaphor explanation) or written input (subdivided into course texts and explicit metaphor learning activities). The aural - video, aural - teacher talk and written - text categories can be considered implicit learning opportunities since the focus in these cases was on the delivery of a content message rather than awareness raising of metaphor. The aural and written explicit metaphor categories account for classroom input where the focus was purely on explaining, discussing, or providing examples of metaphors used to describe culture. At these points in class, there was also an expectation of learner involvement; concept-checking questions were posed, examples and opinions were requested, and learners experienced the metaphor workbook activities described in section 3.5 (page 94).

The distribution of activities over both semesters can be seen in Figure 6.5. It can be seen that the mean number of total exposures increased greatly, and this was primarily due to increases in aural input. Log-likelihood calculations revealed significant differences between the two conditions for total exposures and aural and written explicit input (Table 6.10), although there was a negligible effect size in the last case. As expected, the course materials (in the form of videos and written texts) provided essentially the same amount of input. Incidental input from teacher talk

was also slightly higher, though not significantly so. Since the goal of the study was to raise learners' awareness of metaphor through sustained practice, it was to be expected that there would be some increase in the amount of incidental exposure to metaphor.

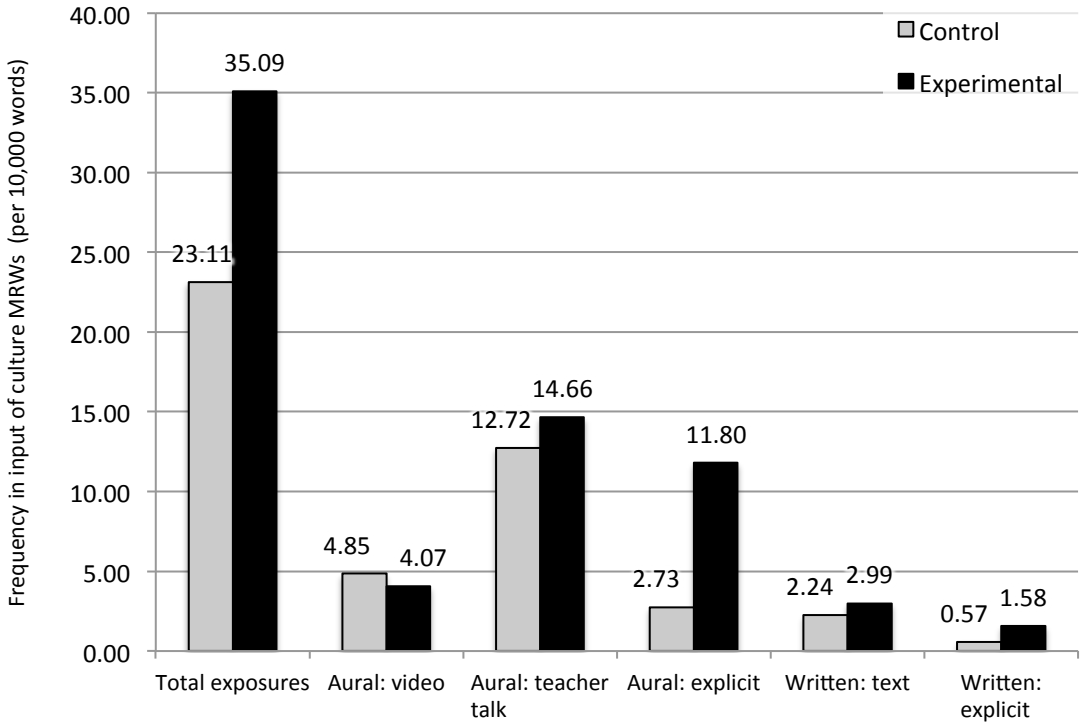


Figure 6.5. Normalised frequencies in input of culture metaphors grouped by input forms

Table 6.10. Log-likelihood calculations for culture metaphors appearing in various input forms between the two conditions

Input form	Frequency		Log-likelihood [a]	Bayes factor
	Con.	Exp.		
Total	567	1044	-66.39****	53.18^{+++}
Aural: video	119	121	1.86	-11.36
Aural: teacher talk	312	436	-3.69	-9.51
Aural: explicit	67	351	-160.66****	147.45^{+++}
Written: text	55	89	-2.89	-10.32
Written: explicit	14	47	-13.05***	-0.15

Note: Control input corpus = 245,363 words; Experimental input corpus = 297,507 words
a = Negative values indicate less representation in control input corpus relative to experimental input
*** = Significant at 0.001 level; **** = 0.0001 level
+++ = Very strong evidence against the null hypothesis

Two other findings are notable here. The first is the degree to which implicit learning opportunities outweigh explicit learning. In the control group, 86% of exposures came when the focus was on the content message rather than the language itself. The experimental treatment had some effect on this, increasing the proportion of explicit learning opportunities to 38% (implicit 62%), largely through aural explanations. This leads to the second point, which is that aural input greatly exceeded written input in both conditions: 88% to 12% in the control and 87% to 13% in the experimental group. It should be noted that the researcher was one of the course teachers, so there is the potential for some influence on the data even though the researcher tried to focus solely on teaching during class time. The intention of the study was to raise awareness of metaphor through providing explicit instruction and increasing incidental exposure. As such, it is no surprise to see that explicit input increased significantly, but if anything, it appears that there was a failure to increase the amount of implicit learning opportunities to any large degree. Whether these findings would hold true for other courses in other contexts cannot of course be confirmed. However, they do provide a vivid illustration of a potential context in which learners might operate, and of the balance of learning opportunities learners may be presented with either in regular content-based instruction or a 'metaphor enhanced' environment.

A final approach to visualising course input was to consider the number of occasions on which instructors dedicated class time to explicit instruction of target metaphors. That is, rather than considering the raw number of occurrences in the input as a measure of salience, this measure assumed that periods of explicit instruction would naturally make target language highly salient for learners. Such

activities would also require learners to apply their own cognition to the target language, which, at least in the case of vocabulary learning (Laufer & Rozovski-Roitblat, 2015, p. 707 - 708), has been identified as a crucial element of language acquisition. During the transcription work to assemble the input corpora, periods of each lesson had been defined by the focus of the class activity at that time (see Appendix E, page 439). For example, class periods were annotated as <ACTIVITY EXPLANATION> or <ANSWER SHARING> in the corpora. The counts for this analysis were the number of separate lesson periods in which target words from each culture metaphor theme were explicitly taught. For example, in the control condition, target words from the CULTURES ARE LIVING THINGS theme were explicitly taught through aural input in six separate class periods throughout the control condition semester, while experimental group learners received explicit aural instruction on words in this theme on 13 occasions. The data for this measure are presented in Table 6.11 and Figure 6.6.

Table 6.11. Number of class periods in which culture metaphors were explicitly taught in the two conditions

Metaphor theme	Aural focus		Written focus	
	Con.	Exp.	Con.	Exp.
CULTURES ARE LIVING THINGS	6	13	2	3
CULTURES ARE SPACES	1	7	2	4
CULTURES ARE SUBSTANCES	1	8	0	3
CULTURES ARE CONSTRUCTIONS	1	3	0	1
CULTURES HAVE MANY PARTS	1	7	1	3
CULTURES ARE POSSESSIONS	1	4	0	1
Total	**11**	**42**	**5**	**15**
Mean (SD)	**1.83 (2.04)**	**7.00 (3.52)**	**0.83 (0.98)**	**2.50 (1.22)**

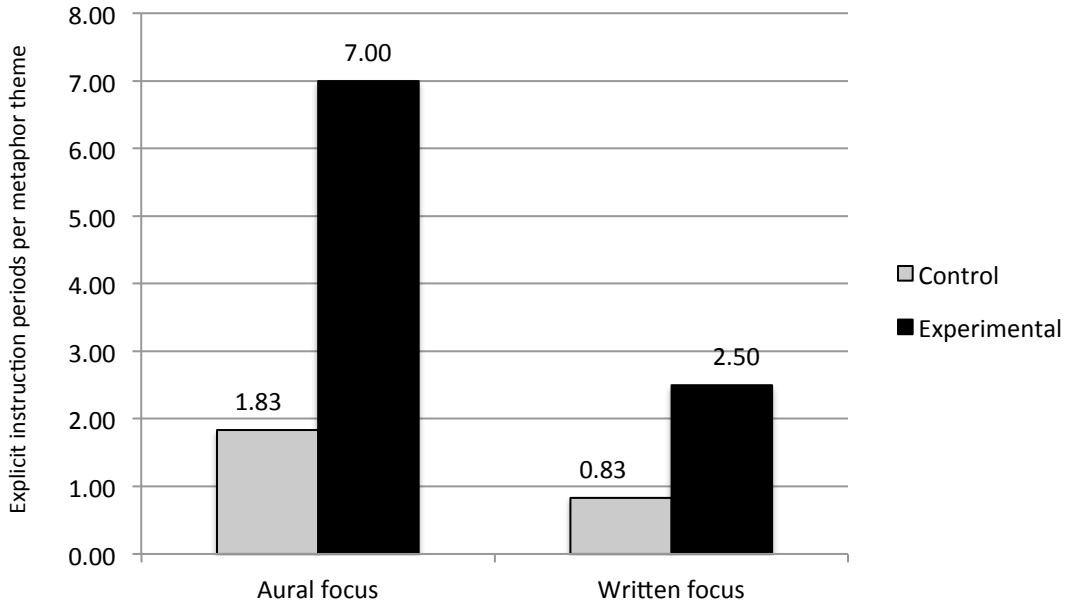


Figure 6.6. Periods of explicit classroom instruction provided for culture metaphors in the two conditions

Since these data are counting periods of class time rather than exposures to words, they cannot be statistically compared with those of Figure 6.5. However, considering the ratios of written explicit input to aural explicit input in the two figures, it does appear that this method of counting gives greater prominence to activities that required learners to engage with the target language in a written medium. Considering only frequency of exposure (i.e., the number of occurrences of target words), written instruction comprised only 17% of all explicit instruction in the control to 83% of occurrences in an aural medium and 12% in the experimental condition to 88% of occurrences of aural input. Regarding periods of class time, class periods which provided an explicit focus on the written form comprised 31% of all periods in the control condition and 26% in the experimental condition respectively. Written activities might not include as many instances of target forms in input as aural explicit instruction, but since learners all need to complete the written tasks (unlike aural instruction, where only some learners may choose to participate) and

often have to relate the language to their own lives, it might be argued that a method of counting that gives greater prominence to activities that require writing is justified. This method also compensates somewhat for the repetition of target words in aural input described in section 6.2.2.

6.4. Metaphors used to describe culture in learner output

Like the balance of metaphor themes across the two conditions in the classroom input, learner production of the themes showed signs of change (Figure 6.7 - 6.8 and Table 6.12). The most obvious of these was the significant increase for metaphors from the CULTURES ARE SPACES theme, with a positive effect size. Metaphors from the CULTURES HAVE MANY PARTS theme were significantly less likely to appear in the experimental corpus than in the control, although the effect size here was negligible.

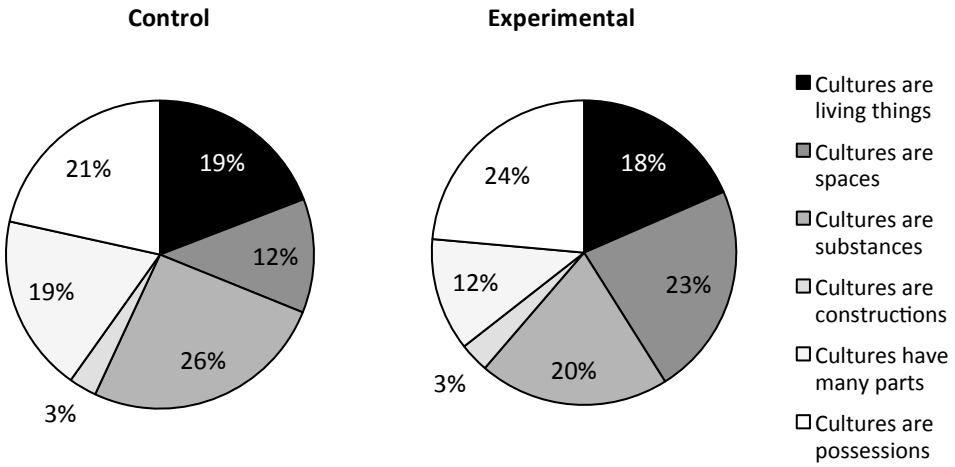


Figure 6.7. Proportions of culture metaphors grouped by theme in course output

Table 6.12. Log-likelihood tests for culture metaphor themes appearing in output between the two conditions

Metaphor theme	*n* Con. [a]	*n* Exp. [a]	Log-likelihood [b]	Bayes factor effect size
CULTURES ARE LIVING THINGS	72	94	0.00	-11.52
CULTURES ARE SPACES	45	115	-16.07****	4.55[+]
CULTURES ARE SUBSTANCES	97	103	2.04	-9.49
CULTURES ARE CONSTRUCTIONS	11	16	-0.08	-11.44
CULTURES HAVE MANY PARTS	70	61	5.22*	-6.30
CULTURES ARE POSSESSIONS	81	120	-0.83	-10.69
Total	376	509	-0.36	-11.16

a = Control output corpus = 43,830 words; Experimental output corpus = 56,972 words
b = Negative values indicate underuse in control condition relative to experimental condition
* = Significant at the 0.05 level; **** = 0.0001 level
+ = Positive evidence against the null hypothesis

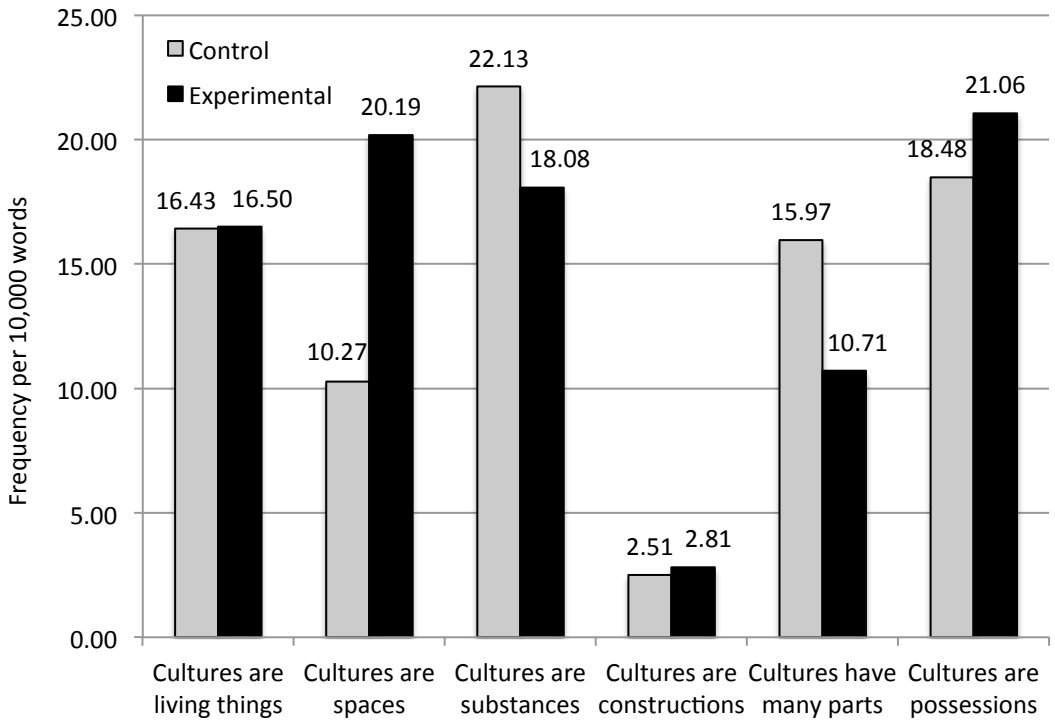


Figure 6.8. Normalised frequencies of culture metaphor themes in output in both conditions

In order to examine the causes of these changes, the frequency and dispersion of the target words across both output corpora were analysed. Frequency was again calculated to a per-10,000-word rate, and dispersion was calculated as both the

percentage of participants who used a given word and the percentage of writing topics in which a given word was used. The results for the 49 taught target culture metaphor word forms and the untaught metaphors produced by learners appear in Tables 6.13 - 6.18. The raw frequencies are given in Appendix L (page 478).

Table 6.13. Log-likelihood tests for output frequency of target metaphors in CULTURES ARE LIVING THINGS theme and dispersions over learners and writing topics

Target word	Frequency per 10,000 words		Log-likelihood [a]	Effect size (Bayes factor)	Dispersion (% of learners)		Dispersion (% of topics)	
	Con.	Exp.			Con.	Exp.	Con.	Exp.
Taught target forms								
alive (adj)	0.00	0.18	/	/	0%	4%	0%	6%
evolution (n)	0.46	0.53	/	/	9%	4%	12%	6%
evolve (v)	2.74	2.28	0.21	-11.31	35%	35%	18%	24%
grow (v)	0.23	1.40	-4.52*	-7.00	4%	22%	6%	35%
hybrid (adj)	2.97	2.46	0.24	-11.28	30%	39%	35%	12%
hybrid (n)	2.97	2.28	0.45	-11.08	35%	43%	18%	18%
Untaught target forms								
ancestor (n)	0.23	0.18	/	/	4%	4%	6%	6%
birth (n)	0.00	0.35	/	/	0%	9%	0%	12%
birthplace (n)	1.37	0.53	1.96	-9.56	22%	13%	12%	6%
born (v)	5.25	4.04	0.79	-10.73	52%	52%	29%	47%
bring up (ph v)	0.00	0.18	/	/	0%	4%	0%	6%
grow up (ph v)	0.00	1.05	-6.85**	-4.67	0%	9%	0%	24%
growth (n)	0.00	0.18	/	/	0%	4%	0%	6%
habitat (n)	0.00	0.18	/	/	0%	4%	0%	6%
recover (v)	0.00	0.35	/	/	0%	9%	0%	12%
revitalize (v)	0.00	0.35	/	/	0%	4%	0%	6%
survive (v)	0.23	0.00	/	/	4%	0%	6%	0%

a = Negative values indicate underuse in control input corpus relative to experimental input corpus. Minimum frequency 5 occurrences in either corpus.

* = Significant at the 0.05 level; ** = 0.01 level

Table 6.14. Log-likelihood tests for output frequency of target metaphors in CULTURES ARE SPACES theme and dispersions over learners and writing topics

Target word	Frequency per 10,000 words		Log-likelihood [a]	Effect size (Bayes factor)	Dispersion (% of learners)		Dispersion (% of topics)	
	Con.	Exp.			Con.	Exp.	Con.	Exp.
Taught target forms								
come into (ph v)	1.14	6.14	-18.13****	6.61++	17%	43%	24%	59%
enter (v)	0.46	3.36	-11.81***	0.28	9%	30%	12%	47%
export (n)	0.00	0.00	/	/	/	/	/	/
export (v)	0.23	0.18	/	/	4%	4%	6%	6%
import (n)	0.00	0.53	/	/	0%	13%	0%	18%
import (v)	0.91	5.79	-18.97****	7.54++	13%	52%	12%	24%
inside (prep)	0.68	0.18	/	/	4%	4%	12%	6%
outside (adv)	0.00	0.00	/	/	/	/	/	/
outside (n)	0.00	0.18	/	/	0%	4%	0%	6%
outside (prep)	0.00	0.00	/	/	/	/	/	/
outsider(n)	0.00	0.00	/	/	/	/	/	/
surround (v)	0.23	0.00	/	/	4%	0%	6%	0%
take in (ph v)	3.88	2.63	1.20	-10.32	35%	30%	47%	29%
transcend (v)	0.00	0.00	/	/	/	/	/	/
Untaught target forms								
access (v)	0.23	0.00	/	/	4%	0%	6%	0%
borderline (n)	0.00	0.18	/	/	0%	4%	0%	6%
bring into (ph v)	0.91	0.00	/	/	4%	0%	18%	0%
contain (v)	1.37	0.00	9.99**	-1.53	4%	0%	29%	0%
deport (v)	0.00	0.18	/	/	0%	4%	0%	6%
go into (ph v)	0.00	0.53	/	/	0%	9%	0%	12%
intake (v)	0.23	0.00	/	/	4%	0%	6%	0%
penetrate (v)	0.00	0.35	/	/	0%	9%	0%	6%

a = Negative values indicate underuse in control input corpus relative to experimental input corpus. Minimum frequency 5 occurrences in either corpus.

** = Significant at the 0.01 level; *** = 0.001 level; **** = 0.0001 level

++ = Strong evidence against the null hypothesis

Table 6.15. Log-likelihood tests for output frequency of target metaphors in CULTURES ARE SUBSTANCES theme and dispersions over learners and writing topics

Target word	Frequency per 10,000 words		Log-likelihood [a]	Effect size (Bayes factor)	Dispersion (% of learners)		Dispersion (% of topics)	
	Con.	Exp.			Con.	Exp.	Con.	Exp.
Taught target forms								
absorb (v)	4.56	0.88	14.00***	2.48[+]	43%	13%	47%	12%
flow (n)	0.23	0.18	/	/	4%	4%	6%	6%
flow (v)	0.00	0.35	/	/	0%	9%	0%	12%
fusion (n)	0.00	0.35	/	/	0%	9%	0%	12%
mix (n)	0.00	0.00	/	/	/	/	/	/
mix (v)	2.05	3.69	-2.30	-9.22	30%	48%	29%	59%
mixture (n)	0.00	1.05	-6.85**	-4.67	0%	17%	0%	24%
shape (n)	0.68	0.53	/	/	13%	9%	18%	6%
shape (v)	0.23	0.53	/	/	4%	9%	6%	12%
spread (v)	7.99	7.02	0.31	-11.21	65%	78%	59%	71%
Untaught target forms								
combine (v)	1.60	0.53	2.87	-8.65	22%	4%	24%	12%
derive (v)	0.46	0.00	/	/	4%	0%	6%	0%
flood (v)	0.00	0.18	/	/	0%	4%	0%	6%
form (n)	2.97	2.11	0.73	-10.79	17%	35%	24%	35%
form (v)	0.23	0.18	/	/	4%	4%	6%	6%
permeate (v)	0.23	0.00	/	/	4%	0%	6%	0%
refine (v)	0.00	0.35	/	/	0%	9%	0%	12%
spread out (ph v)	0.91	0.18	/	/	13%	4%	18%	6%

a = Negative values indicate underuse in control input corpus relative to experimental input corpus. Minimum frequency 5 occurrences in either corpus.

*** = Significant at the 0.01 level; *** = 0.001 level

+ = Positive evidence against the null hypothesis

Table 6.16. Log-likelihood tests for output frequency of target metaphors in CULTURES ARE CONSTRUCTIONS theme and dispersions over learners and writing topics

Target word	Frequency per 10,000 words		Log-likelihood [a]	Effect size (Bayes factor)	Dispersion (% of learners)		Dispersion (% of topics)	
	Con.	Exp.			Con.	Exp.	Con.	Exp.
Taught target forms								
construct (v)	0.91	0.70	/	/	17%	13%	6%	6%
damaged (adj)	0.00	0.18	/	/	0%	4%	0%	6%
fit (v)	0.00	0.00	/	/	/	/	/	/
fit into (ph v)	0.23	0.70	/	/	4%	13%	6%	12%
make up of (ph v)	0.00	0.00	/	/	/	/	/	/
manufacture (v)	0.00	1.05	-6.85**	-4.67	0%	9%	0%	6%
Untaught target forms								
break (v)	0.23	0.18	/	/	4%	4%	6%	6%
destroy (v)	0.23	0.00	/	/	4%	0%	6%	0%
destruction (n)	0.23	0.00	/	/	4%	0%	6%	0%
reconstruct (v)	0.23	0.00	/	/	4%	0%	6%	0%
restoration (n)	0.46	0.00	/	/	9%	0%	12%	0%

a = Negative values indicate underuse in control input corpus relative to experimental input corpus. Minimum frequency 5 occurrences in either corpus.
** = Significant at the 0.01 level

Table 6.17. Log-likelihood tests for output frequency of target metaphors in CULTURES HAVE MANY PARTS theme and dispersions over learners and writing topics

Target word	Frequency per 10,000 words		Log-likelihood [a]	Effect size (Bayes factor)	Dispersion (% of learners)		Dispersion (% of topics)	
	Con.	Exp.			Con.	Exp.	Con.	Exp.
Taught target forms								
aspect (n)	3.42	0.88	8.20**	-3.32	48%	13%	29%	18%
component (n)	0.46	0.70	/	/	9%	17%	12%	24%
divide (v)	0.00	0.18	/	/	0%	4%	0%	6%
element (n)	5.70	2.63	5.83*	-5.69	57%	39%	47%	35%
ingredient (n)	0.00	0.00	/	/	/	/	/	/
part (n)	2.97	1.40	2.87	-8.65	30%	17%	59%	41%
Untaught target forms								
add (v)	3.42	4.56	-0.81	-10.71	35%	48%	24%	29%
piece (n)	0.00	0.35	/	/	0%	4%	0%	6%

a = Negative values indicate underuse in control input corpus relative to experimental input corpus. Minimum frequency 5 occurrences in either corpus.
* = Significant at the 0.05 level; ** = 0.01 level

Table 6.18. Log-likelihood tests for output frequency of target metaphors in CULTURES ARE POSSESSIONS theme and dispersions over learners and writing topics

Target word	Frequency per 10,000 words		Log-likelihood [a]	Effect size (Bayes factor)	Dispersion (% of learners)		Dispersion (% of topics)	
	Con.	Exp.			Con.	Exp.	Con.	Exp.
Taught target forms								
borrow (v)	0.00	0.35	/	/	0%	9%	0%	6%
claim (v)	0.68	0.35	/	/	13%	4%	6%	12%
lose (v)	2.97	5.44	-3.62	-7.90	35%	65%	18%	18%
share (v)	0.00	0.70	-4.56*	-6.96	0%	13%	0%	12%
shared (adj)	0.00	0.00	/	/	/	/	/	/
steal (v)	0.00	0.18	/	/	0%	4%	0%	6%
take (v)	2.51	2.81	-0.08	-11.44	22%	26%	41%	35%
Untaught target forms								
accept (v)	3.65	2.11	2.10	-9.42	35%	35%	53%	41%
adopt (v)	1.60	1.93	-0.16	-11.36	26%	30%	29%	29%
exchange (n)	0.23	0.00	/	/	4%	0%	6%	0%
exchange (v)	0.00	0.18	/	/	0%	4%	0%	6%
give (v)	0.23	0.35	/	/	4%	4%	6%	12%
hand down (ph v)	0.46	0.00	/	/	9%	0%	6%	0%
have (v)	2.51	2.63	-0.01	-11.51	35%	43%	41%	35%
inherit (v)	0.23	0.35	/	/	4%	9%	6%	12%
keep (v)	2.74	1.76	1.08	-10.44	26%	26%	47%	41%
pass (v)	0.00	0.18	/	/	0%	4%	0%	6%
possess (v)	0.00	0.18	/	/	0%	4%	0%	6%
receive (v)	0.68	1.05	-0.39	-11.13	9%	17%	12%	24%
retain (v)	0.00	0.18	/	/	0%	4%	0%	6%
save (v)	0.00	0.35	/	/	0%	4%	0%	6%

a = Negative values indicate underuse in control input corpus relative to experimental input corpus. Minimum frequency 5 occurrences in either corpus.

* = Significant at the 0.05 level

The results here show a mixed pattern of rising and falling frequencies. Seven taught target words (*grow*, *come into*, *enter*, *import* (v), *mixture*, *manufacture* (v), and *share* (v)) were used significantly more frequently in the experimental condition,

but of these, only *come into* and *import* displayed strong effect sizes. In addition, *absorb*, *aspect*, and *element* were all used significantly less often by learners, with a moderate effect size in the case of *absorb*. Overall, 59% of the taught culture metaphors that were actually used by either group had a higher frequency in the experimental condition, while for untaught target culture metaphors this value was 52%.

There appears to be a tendency for taught culture metaphors that appeared at a relatively high frequency in the control condition output to be used less often by experimental group learners. Of the 13 words to be produced at a rate of greater than 1 per 10,000 in the control group, 69% fell in frequency in the treatment condition (although apart from *absorb*, they remained frequently used). Forms that appeared at lower frequencies (less than once per 10,000 words) in the control, however, were likely to be more frequent in the experimental group than to diminish in use. This occurred in 73% (19 of 26) cases.

This pattern seems to support the claim that learners were as much adding to the diversity of their metaphorical output as to the overall rate of metaphorical expression. Words that were commonly used in the control condition were drawn on less often, and other expressions that could fulfil the same semantic function were used instead. In place of *absorb* (88), *come in*/*into* (89) and *enter* (90) were used more frequently to express cultural borrowing, and *mix* (91) became an alternative to *combine* (92) in describing the blending of cultural forms.

(88) The beginning of this style [church weddings] is Europe, and Japan absorbs this style as wedding ceremony element. [CON11]

(89) *I learned that western style came into Japanese culture from Makiko's New world.* [EXP6]

(90) *Today, Japanese maid culture is entering other countries* [EXP2]

(91) *Hybridity means that a thing enters in the other country, and mixes with the other country's culture.* [EXP5]

(92) *However, this dish does not only combine the cultures of India and Japan.* [CON4]

Ten of the taught target culture metaphors did not appear in the output of either condition. For the words *transcend*, *mix* (n), *fit*, *make up of*, *ingredient*, and *shared* (adj), the evidence from Tables 6.2 - 6.7 might suggest that this was simply due to a lack of input. In both input corpora, these words appeared both at low frequencies and in few lessons. Several of these words could also be substituted with a word that did appear frequently in output, such as *mixture* for *mix* (n), *mix* (v) for *made up of*, or *element*/*aspect*/*component*/*part* for *ingredient*, and it may be the case that learners chose to use a smaller number of forms repeatedly rather than to experiment with all of the target expressions. Other unused words represent concepts that are themselves less frequent in the course content, such as transcending a culture or being a cultural outsider.

The words which were most likely to be used frequently in the experimental condition (i.e., at a rate greater than or equal to 1 per 10,000) are those which also appeared in the control condition to some degree (Table 6.19). The use of *hybrid* as a noun and an adjective can be explained by the fact that this term was particularly important in the course content, and it consequently appeared frequently in input

(Table 6.2). Likewise, *element* and *part* also appeared at high frequencies in the control group input. Of the remaining forms, several (*add*, *born*, *enter*, *grow*, *have*, *keep*, *lose*, *mix*, *receive*, and *take*) are conceptually and morphologically simple, high frequency words which most learners in the A2 to B2 CEFR range would certainly be expected to know in their basic senses (Cambridge University Press[39], 2012). Even *import* exists as a loanword in Japanese, increasing the likelihood that it would also be recognised by the learners. Only two of the taught target words that did not appear at all in control group output were produced frequently in the experimental group output (*mixture* and *manufacture*). This seems to suggest that gradual change, rather than widespread adoption of new metaphorical vocabulary, can be expected in language production.

Table 6.19. Target culture metaphors with a frequency ≥1 per 10,000 lexical units in the experimental output corpus grouped by frequency in control output corpus

Culture metaphor category	Normalised frequency in control condition output (per 10,000 lexical units)		
	0	**> 0 - < 1**	**≥ 1**
Taught	mixture, manufacture (v)	grow, enter, import (v)	evolve, hybrid (adj), hybrid (n), come into, take in, mix (v), spread (v), element, part, lose, take
Untaught	grow up	receive	born, form (n), add, accept, adopt, have, keep

To follow up on Ozturk's (2015, p. 106) claim that vocabulary production is influenced more by learner need than by raw frequency as measured in large corpora, correlations were run between the normalised output frequencies of taught target metaphors in each condition, the normalised input frequencies for the same

[39] This thesis has made use of the English Vocabulary Profile. This resource is based on extensive research using the Cambridge Learner Corpus and is part of the English Profile programme, which aims to provide evidence about language use that helps to produce better language teaching materials. See http://www.englishprofile.org for more information.

metaphors in each condition, and the normalised frequencies of those metaphorical words in COCA[40]. Once again, all values were calculated to a 1 per 10,000 rate. Metaphors that did not appear in output were not included in the calculations. The results in Table 6.20 appear to support Ozturk's finding, at least when considering the production of metaphorical vocabulary; learner output showed a closer relationship to the frequencies of the metaphorical words they were exposed to in input than the frequencies of those metaphors in a much broader sample of language.

Table 6.20. Spearman correlations between normalised frequencies of taught target culture metaphors in learner output in both conditions and normalised frequencies of the same metaphors in COCA and the corresponding input corpus

Measure of input	Control			Experimental		
	n [a]	r_s	*p*	*n* [a]	r_s	*p*
Normalised freq. in COCA	25	.216	.278	37	.281	.093
Normalised freq. in input corpus	25	.700	.000**	37	.677	.000**

a = 25 culture metaphor words appeared in output, input, and COCA in the control condition (frequency > 0), while 37 words did so in the experimental condition.

** = Significant at the 0.01 level

6.5. Longitudinal analysis of culture metaphors

The data presented in this chapter have reflected the complexity of the learning environment, with learners' lexical production showing signs of ebb and flow as individuals experiment with new forms and loosen their ties to trusted but overused phrases. Clearly, we are looking at a dynamic system in operation, and because of this, it is important to consider how metaphor production varies over time.

[40] Frequency of metaphors in COCA was extrapolated from 100-line concordance samples, with metaphor being identified using the MIPVU procedure.

One issue with diachronic analyses, however, is that the closer we look at the data, the less distinct any sense of patterning becomes – essentially a 'wood for the trees' problem. For this reason, the temporal aspect of the data will first be examined in blocks of time, allowing for a broader depiction of trends. Later, the analysis will narrow the data down to the finest possible level.

The approach taken here will mirror that of Bell (2009), who examined developments in a learner's use of academic phrases by dividing a period of study into equal parts. Here, the semesters will be divided into three periods, each containing five submissions of written work. Figure 6.9 shows how the normalised frequencies of culture metaphors in learner output (grouped into taught target, untaught target and combined taught and untaught target metaphors) changed over the three time periods. The chart gives a strong impression of growth in metaphor production over time. In order to check for the significance of growth within each condition, Friedman's Two-way Analysis of Variance by Ranks tests were performed (Table 6.21). No significant differences were found in the control condition, but significant differences were found in the experimental condition for taught, untaught, and combined target metaphors. The locations of these significance differences were then checked with Wilcoxon Signed-ranks tests, the results of which are given in Table 6.22. The raw data for this analysis can be found in Appendix M (page 487).

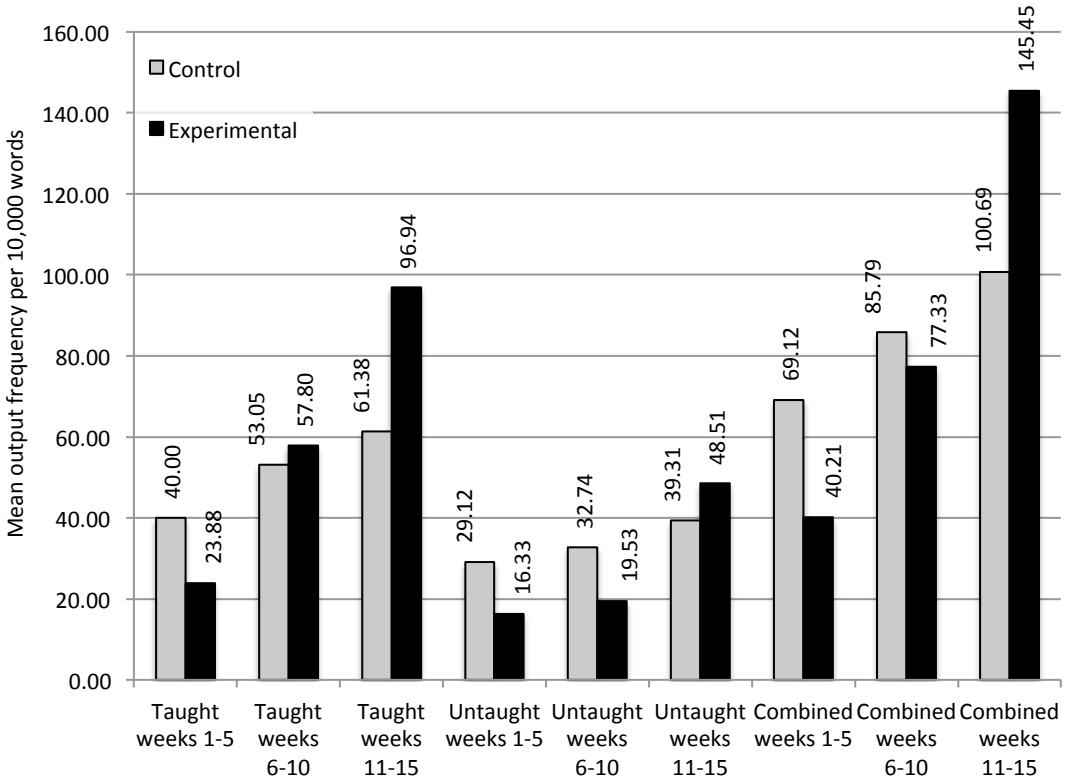


Figure 6.9. Mean normalised output frequencies of target culture metaphors over time

Table 6.21. Friedman tests for normalised output frequencies of target culture metaphors over the three time periods within the two conditions

Condition (category)	Mean ranks			*z*	*p*
	Weeks 1-5	Weeks 6-10	Weeks 11-15		
Con. (taught)	1.70	2.00	2.30	4.356	.113
Con. (untaught)	2.04	1.91	2.04	0.273	.873
Con. (combined)	1.96	1.87	2.17	1.130	.568
Exp. (taught)	1.39	1.96	2.65	18.348	.000*
Exp. (untaught)	1.70	1.74	2.57	11.545	.003*
Exp. (combined)	1.35	1.87	2.78	24.261	.000*

Note: *n* = 23 for all tests

* = Significant difference found

Table 6.22. Wilcoxon Signed Rank tests for normalised output frequencies of target culture metaphors over the three time periods within the experimental condition

Condition	Metaphor type	Time periods (median output frequencies[a])	*n*	*z*	*p*	*r*
Experimental	Taught	Weeks 1-5 / Weeks 6-10 16.31 60.70	23	-3.429	.001**	0.71^{+++}
Experimental	Taught	Weeks 6-10 / Weeks 11-15 60.70 97.48	23	-2.679	.007*	0.56^{+++}
Experimental	Taught	Weeks 1-5 / Weeks 11-15 16.31 97.48	23	-3.902	.000***	0.81^{+++}
Experimental	Untaught	Weeks 1-5 / Weeks 6-10 13.48 15.17	23	-0.945	.344	0.20^{+}
Experimental	Untaught	Weeks 6-10 / Weeks 11-15 15.17 46.89	23	-3.035	.002**	0.63^{+++}
Experimental	Untaught	Weeks 1-5 / Weeks 11-15 13.48 46.89	23	-3.678	.000***	0.77^{+++}
Experimental	Both	Weeks 1-5 / Weeks 6-10 32.36 75.87	23	-3.124	.002**	0.65^{+++}
Experimental	Both	Weeks 6-10 / Weeks 11-15 75.87 149.81	23	-3.429	.001**	0.71^{+++}
Experimental	Both	Weeks 1-5 / Weeks 11-15 32.36 149.81	23	-4.027	.000***	0.84^{+++}

a = Frequencies were normalised to a per 10,000 word rate
* = Significant at the 0.05 level; ** = 0.01 level; *** = 0.001 level
+ = Small effect size; +++ = Large effect size

It can be seen that in most cases, metaphor production increased significantly over each time period in the experimental condition, with medium to large effect sizes present. These results are particularly interesting since it appears that there was a tendency for culture metaphor production to accelerate as the semester progressed. It is also worth noting that while the analyses in chapters four and five found that untaught target metaphors as a whole were less frequent in the experimental condition, this subset of the data suggests that learners were

beginning to reverse that trend by the later stages of the semester. Figure 6.9 shows that output of culture metaphors in the control group also increased over time, but since no significant differences were found, it appears that the experimental treatment may have exacerbated this trend.

The differences between the two conditions at each point in time were also evaluated with Mann-Whitney tests. The results, shown in Table 6.23, give the impression that the experimental group learners were increasing the production of culture metaphors at a more rapid pace than learners in the control condition. For all three metaphor types, experimental group learners initially produced less culture metaphors than the control, significantly so in the case of combined taught and untaught target metaphors. However, they overcame this deficit, and by the final third of the semester were producing more culture metaphors. The Bonferroni-

Table 6.23. Mann-Whitney tests for normalised output frequencies of target culture metaphors across the same time periods between the two conditions

Metaphor type (period)	Median output freq.[a]		Mean ranks		*U*	*z*	*p*	*r*
	Con.	Exp.	Con.	Exp.				
Taught (Weeks 1-5)	31.45	16.31	26.48	20.52	196.000	-1.515	.130	0.22^{+}
Taught (Weeks 6-10)	55.35	60.70	22.35	24.65	238.000	-.582	.560	0.09
Taught (Weeks 11-15)	59.59	97.48	19.22	27.78	166.000	-2.164	.030$^{\#}$	0.32^{++}
Untaught (Weeks 1-5)	30.72	13.48	27.39	19.61	175.000	-1.984	.047$^{\#}$	0.29^{+}
Untaught (Weeks 6-10)	26.25	15.17	26.43	20.57	197.000	-1.493	.135	0.22^{+}
Untaught (Weeks 11-15)	23.75	46.89	21.13	25.87	210.000	-1.199	.230	0.18^{+}
Combined (Weeks 1-5)	62.89	32.36	28.74	18.26	144.000	-2.648	.008*	0.39^{++}
Combined (Weeks 6-10)	87.98	75.87	24.91	22.09	232.000	-.714	.475	0.11^{+}
Combined (Weeks 11-15)	89.29	149.81	19.13	27.87	164.000	-2.208	.027$^{\#}$	0.33^{++}

Note: *n* = 46 for all tests
a = Frequencies normalised to a per 10,000 word rate
# = Non-significant after Bonferroni adjustment
* = Significant at the 0.05 level
+ = Small effect size; ++ = Medium effect size

adjusted *p*-values for the final third of the semester in both the taught and combined categories do not reach significance, but the effect sizes do imply that there has been some change in performance.

Although the data presented in Figure 6.9 show output of metaphors per 10,000 words in order to account for differences in text length, it may be possible that longer texts nevertheless allow for greater metaphor use simply because learners discuss abstract notions in greater depth. The output corpus data presented in Appendix I (page 455) show that experimental group learners produced more words than their control group counterparts, so this was a point worthy of investigation. Friedman tests revealed that there were significant differences in the amount of lexical units produced within both conditions over the three five-week periods (Table 6.24). Wilcoxon Signed Ranks tests using a Bonferroni-adjusted value for alpha revealed that in both conditions, there were significant differences between the number of lexical units produced in the first five-week period and the number produced in both the second and third five-week periods of the semesters. Effect sizes were large in all cases. There were no significant differences between the second and third period in either condition (Table 6.25). Table 6.26 shows the results of Mann-Whitney tests (with Bonferroni adjustment applied) for the number of lexical units produced at the three time periods between the conditions. In all cases, the experimental group produced significantly more lexical units in their writing, although the effect sizes show this difference was large at the beginning of the semester and moderate thereafter.

Table 6.24. Friedman tests for lexical units produced over the three time periods within the two conditions

Condition	*n*	Chi-square (df = 2)	Mean ranks Weeks 1-5	Weeks 6-10	Weeks 11-15	*p*
Control	23	21.130	1.22	2.39	2.39	.000*
Experimental	23	11.217	1.43	2.35	2.22	.004*

Table 6.25. Wilcoxon Signed Rank tests for lexical units produced over the three time periods within the two conditions

Condition	Negative ranks (mean ranks)	Positive ranks (mean ranks)	Ties	Sums of ranks	*z*	*p*	*r*
Control condition							
Weeks 1-5 / 6-10	4 (4.00)	19 (13.68)	0	16.00 / 260.00	-3.711	.000***	0.77^{+++}
Weeks 1-5 / 11-15	1 (2.00)	22 (12.45)	0	2.00 / 274.00	-4.137	.000***	0.86^{+++}
Weeks 6-10 / 11-15	13 (10.77)	10 (13.60)	0	140.00 / 136.00	-0.061	.951	0.01
Experimental condition							
Weeks 1-5 / 6-10	6 (4.92)	17 (14.50)	0	29.50 / 246.50	-3.300	.001**	0.69^{+++}
Weeks 1-5 / 11-15	4 (7.25)	19 (13.00)	0	29.00 / 247.00	-3.315	.001**	0.69^{+++}
Weeks 6-10 / 11-15	14 (13.82)	9 (9.17)	0	193.50 / 82.50	-1.688	.091	0.35^{++}

** = Significant at the 0.01 level; *** = 0.001 level
++ = Medium effect size; +++ = Large effect size

Table 6.26. Mann-Whitney tests for lexical units produced across the same time periods between conditions

Period	Median lexical units		Mean ranks		*U*	*z*	*p*	*r*
	Con.	Exp.	Con.	Exp.				
Weeks 1-5	522	747	14.80	32.20	64.500	-4.394	.000***	0.65^{+++}
Weeks 6-10	737	882	18.61	28.39	152.000	-2.472	.013*	0.36^{++}
Weeks 11-15	672	843	18.74	28.26	155.000	-2.406	.016*	0.35^{++}

Note: *n* = 46 for all tests.
* = Significant at the 0.05 level; *** = 0.001 level
++ = Medium effect size; +++ = Large effect size

Therefore, although a relationship between greater quantity of output and increased metaphor production cannot be completely ruled out, the evidence suggests that text length alone cannot explain the increased levels of target

metaphor production over time. While the experimental group did produce more lexical units than the control, this difference was decreasing over the semester (Table 6.26), yet the rate of target metaphor production in the experimental group was increasing relative to the control (Figure 6.9 and Table 6.22). In the control condition, text lengths increased significantly between the first five-week period and the second and third periods of both semesters (Table 6.25), but metaphor production rose only slowly and with no significant differences at any stage (Table 6.21). In experimental group, there was the same pattern of increases in text length, but as Table 6.22 shows, target metaphor production increased significantly in all categories and between almost all time periods. In fact, the only occasion when a significant increase was not found (untaught target metaphor production between weeks 1 - 5 and 6 - 10) occurred when text length did increase significantly.

This analysis has shown a tendency for the production of metaphors relating to culture to increase over time. This seems to occur in both conditions, with the effect being much more pronounced in the experimental group than the control. In order to consider the causes of this apparent change in behaviour, it is necessary to examine both input and output in as fine a level of detail as possible.

As was explained in chapter three, learners in both conditions took two classes of the Japanese Popular Culture course every week for a whole semester and were asked to complete a piece of written reflective writing for homework at the end of each week. Thus, there were two periods of classroom input for every written assignment. In order to simplify the presentation of data somewhat, Figures 6.10 - 6.15 have been constructed with the input from the two classroom periods each week combined. The horizontal axes show the 15 weeks of the semester and each

interval corresponds to one written assignment being completed. The output from each writing topic (shown by a line chart) has been plotted against the classroom input (shown as a column chart) from the previous week, so that output is displayed in the same position as the input that informed it. In other words, although the output in week 1 was actually submitted at the beginning of the second week, it is plotted here in the same position as the input from the first week, since that contained the concepts and themes the learners were reflecting on. The vertical axes show the frequency of output (taught and untaught forms combined) and input in each corpus, and the column shading indicates the medium of input and form of learning opportunity (implicit or explicit) afforded by each instance of classroom input.

Considering input in the control condition first, it is clear that the six metaphor themes describing culture appeared in class during most weeks of the semester. The one exception to this is the CULTURES ARE CONSTRUCTIONS theme, which had the lowest rates of input overall. Thus, it can be claimed that while control group learners did not receive sustained explicit instruction in metaphor, they were exposed to such language regularly. Rather than showing any particular trend, input in each metaphor theme fluctuated, with peaks caused by the chosen course topics. It can be seen that in most cases, the peaks in the control condition are formed of incidental learning opportunities, although periods in which explicit instruction was provided for individual words caused spikes in input. For example, instruction provided for the word *hybrid*, which was part of the CULTURES ARE LIVING THINGS metaphor theme, in week six required a lot of explanation and review, since this was a key course concept.

Turning to the experimental group, the most obvious difference between the two conditions are the huge spikes in input that appear when each metaphor theme was taught explicitly (Table 6.27). These spikes in the data demonstrate the impact that explicit instruction has in terms of frequency of exposure to target forms. Other spikes are also present, either because a particular word was taught in its metaphorical sense, because a review activity was carried out later in the course, or because a theme was particularly prominent at one time in the course materials. Few of these spikes come close to matching the initial period of instruction for sheer exposure, however. The one exception to this is the use of the CULTURES ARE LIVING THINGS metaphor in the sixth week of the experimental condition when the term hybrid was extensively used in class.

Table 6.27. Explicit teaching of culture metaphors in the experimental condition

Metaphor theme	Teaching in experimental condition
Cultures are living things	Week 1
Cultures are spaces	Week 5
Cultures are substances	Week 12
Cultures are constructions	Week 12
Cultures have many parts	Week 12
Cultures are possessions	Week 13

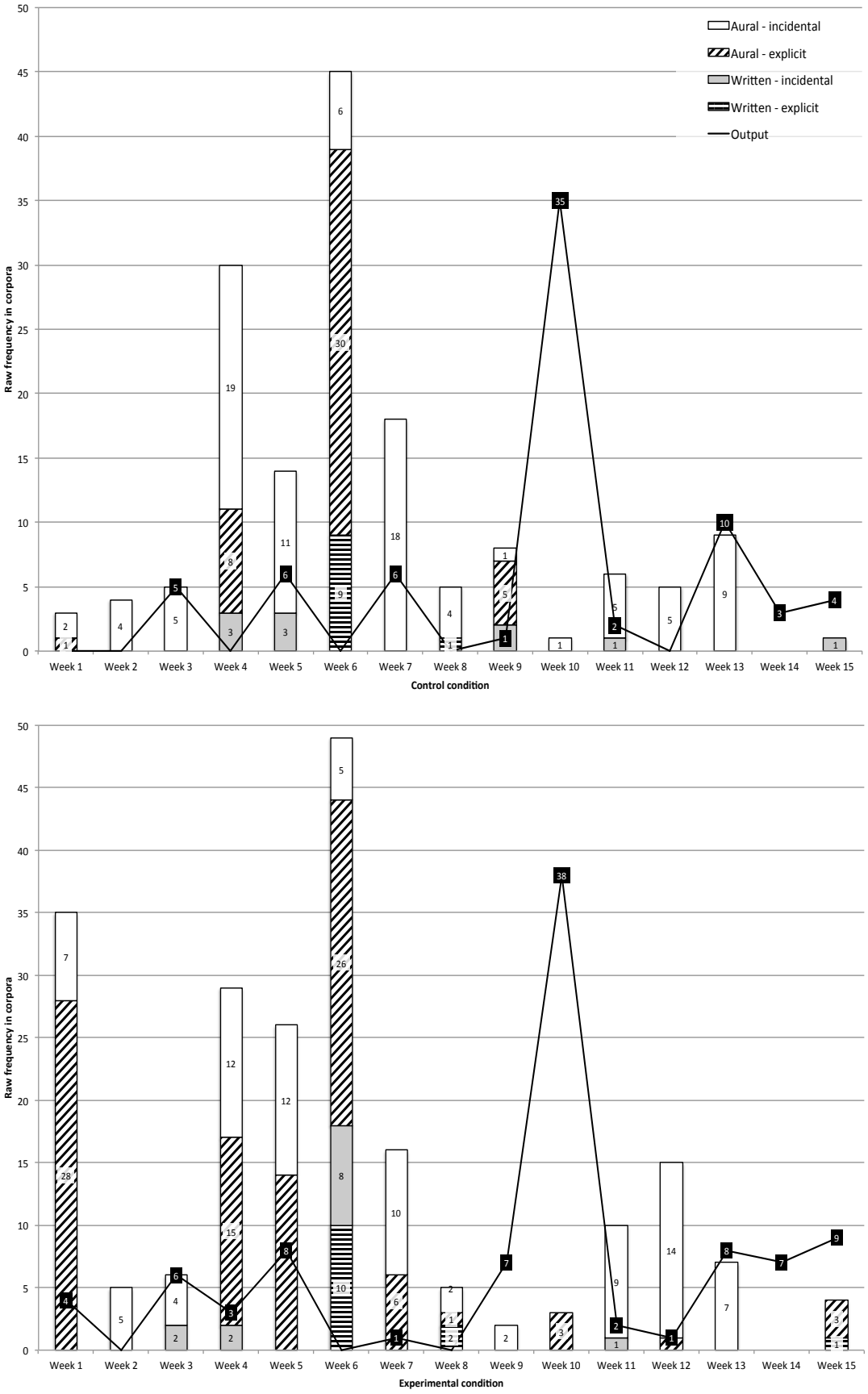


Figure 6.10. Input and output of the CULTURES ARE LIVING THINGS metaphor over time

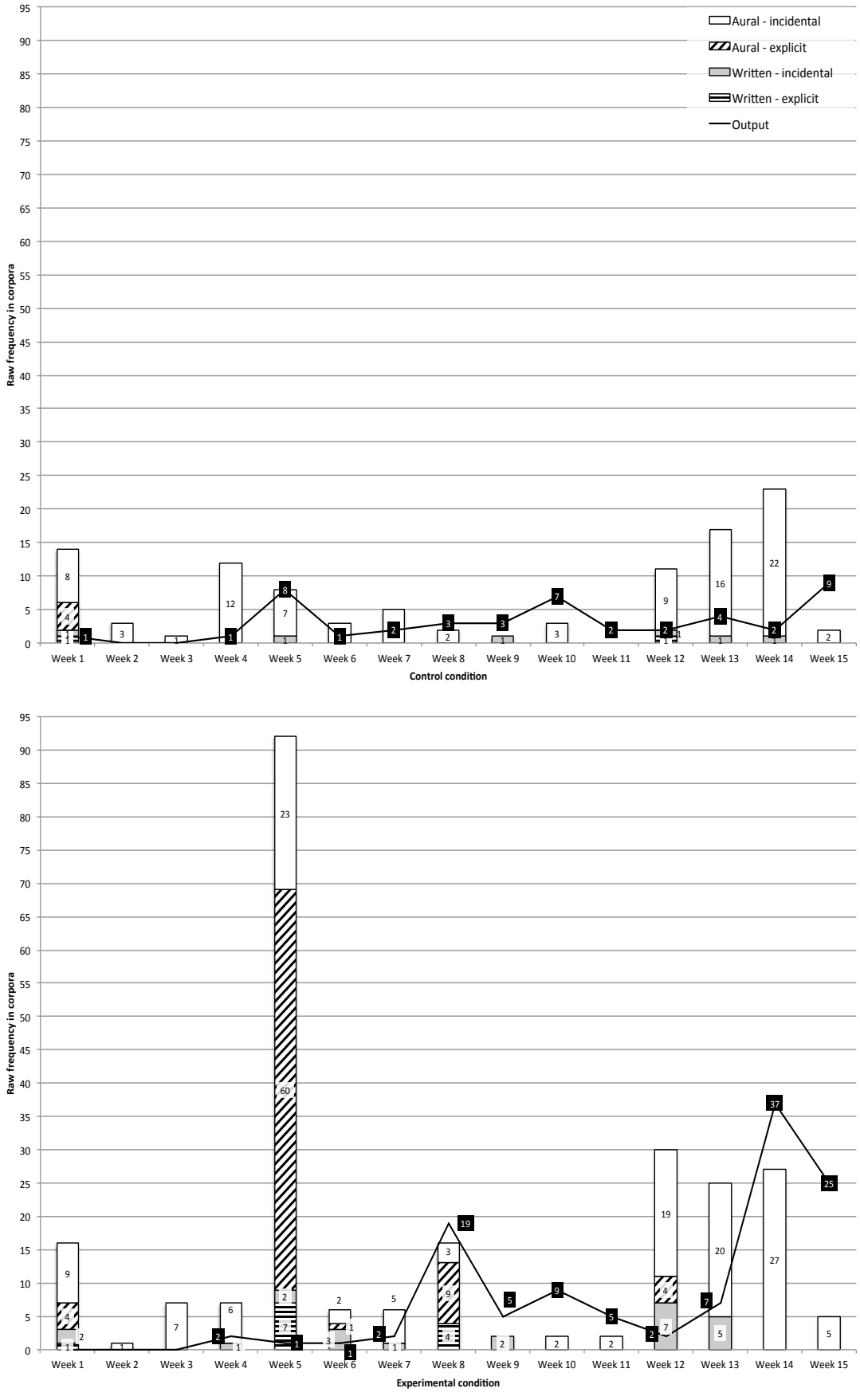


Figure 6.11. Input and output of the CULTURES ARE SPACES metaphor over time

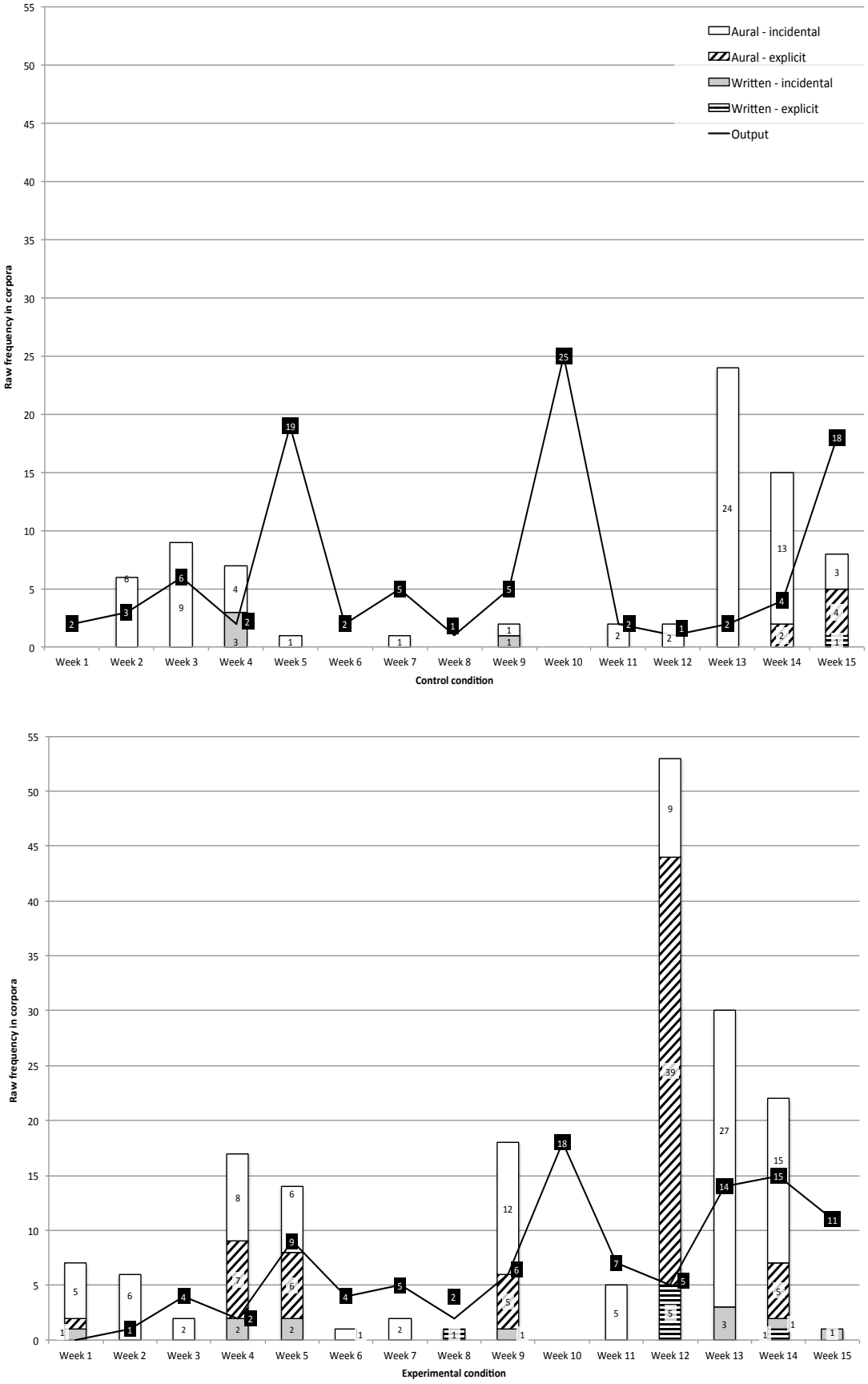


Figure 6.12. Input and output of the CULTURES ARE SUBSTANCES metaphor over time

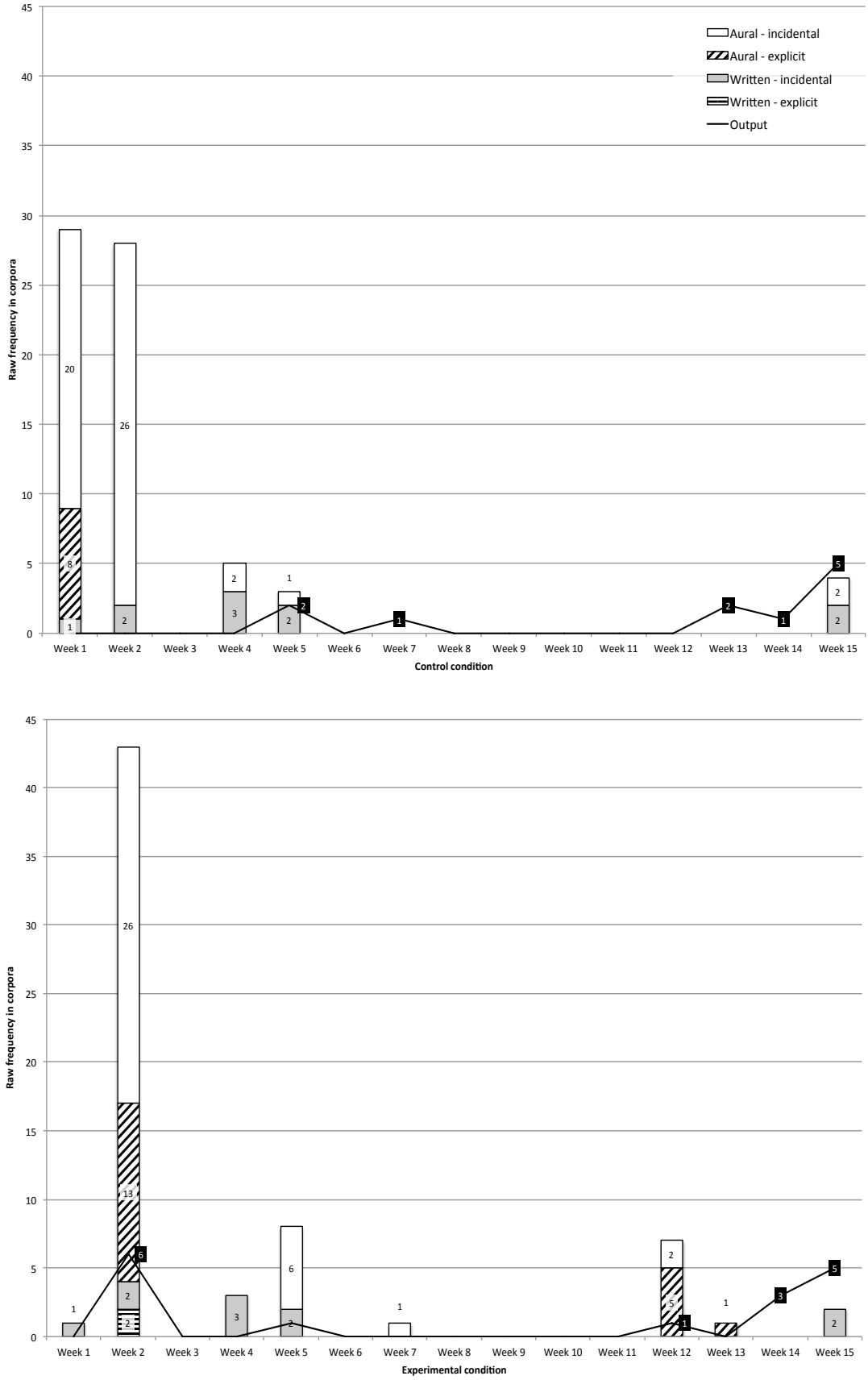


Figure 6.13. Input and output of the CULTURES ARE CONSTRUCTIONS metaphor over time

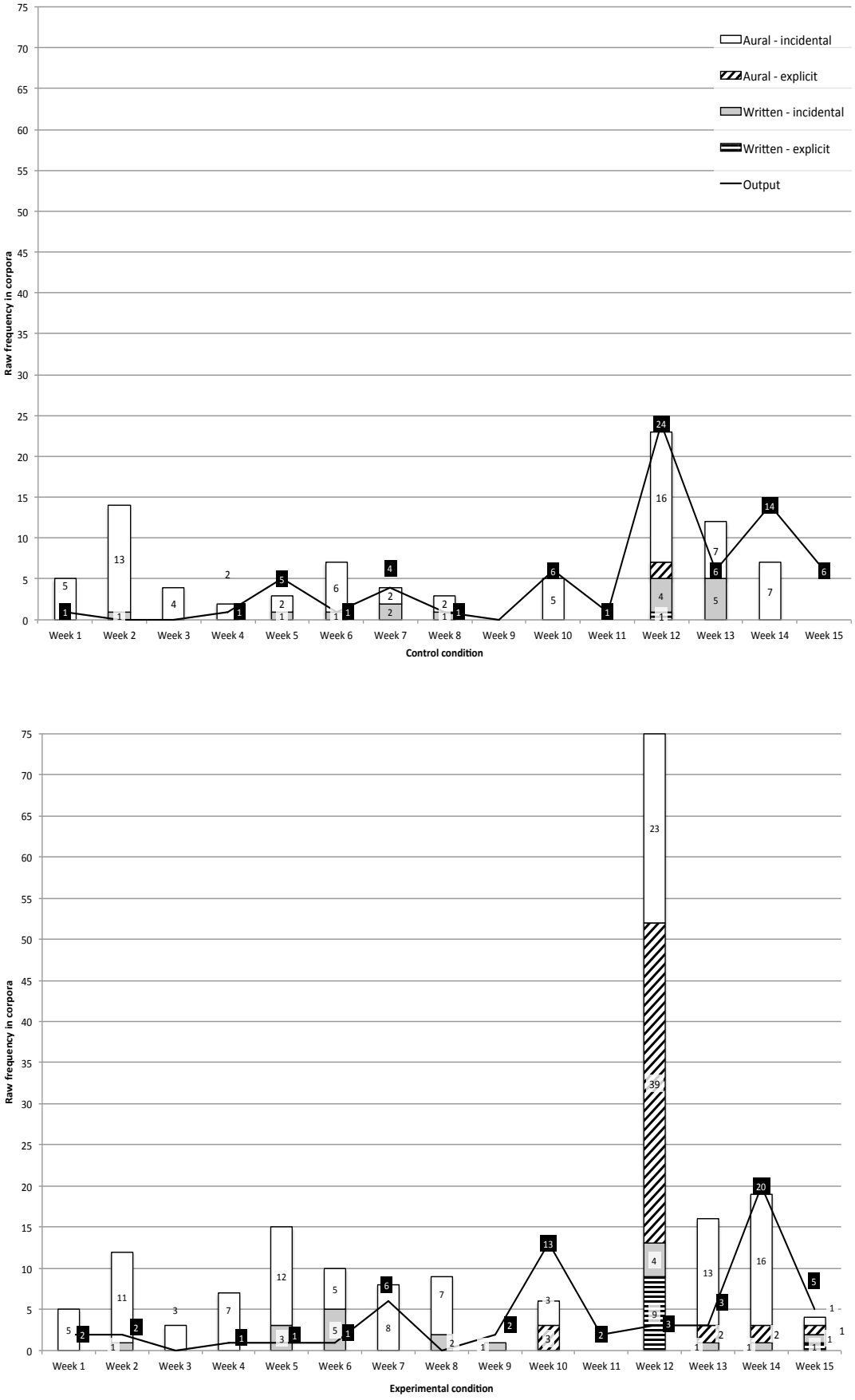
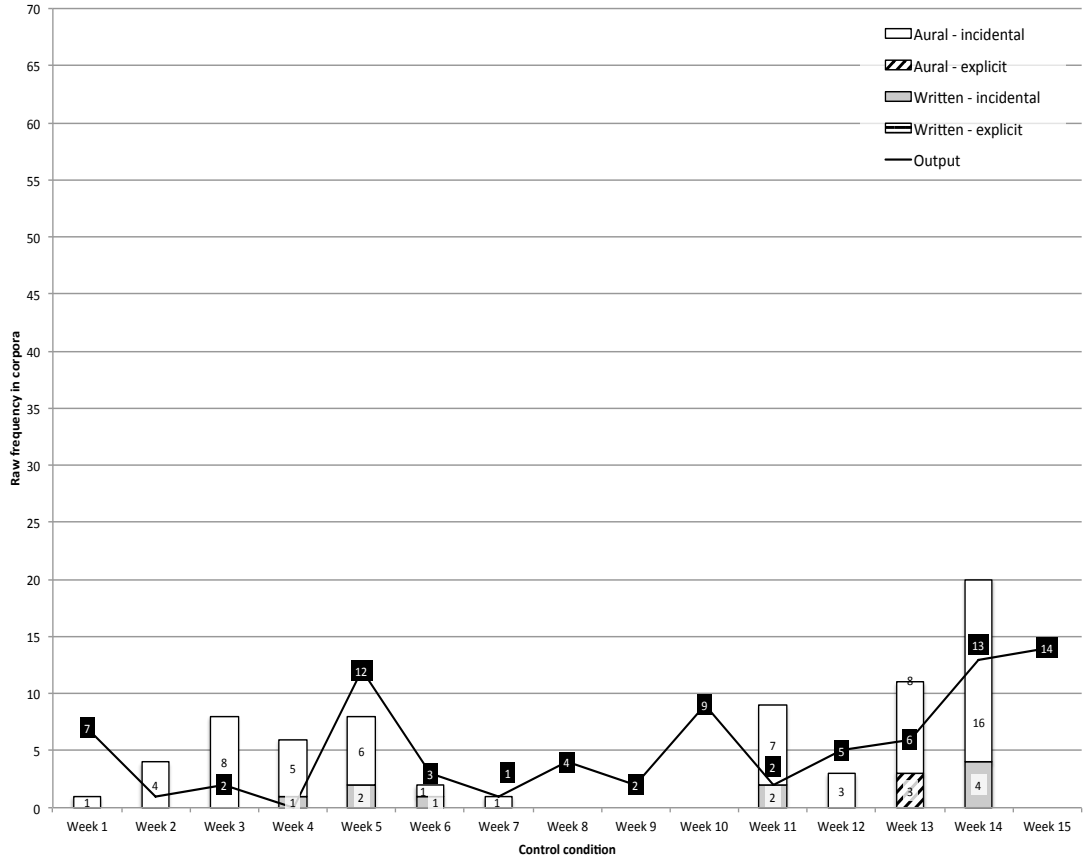


Figure 6.14. Input and output of the CULTURES HAVE MANY PARTS metaphor over time



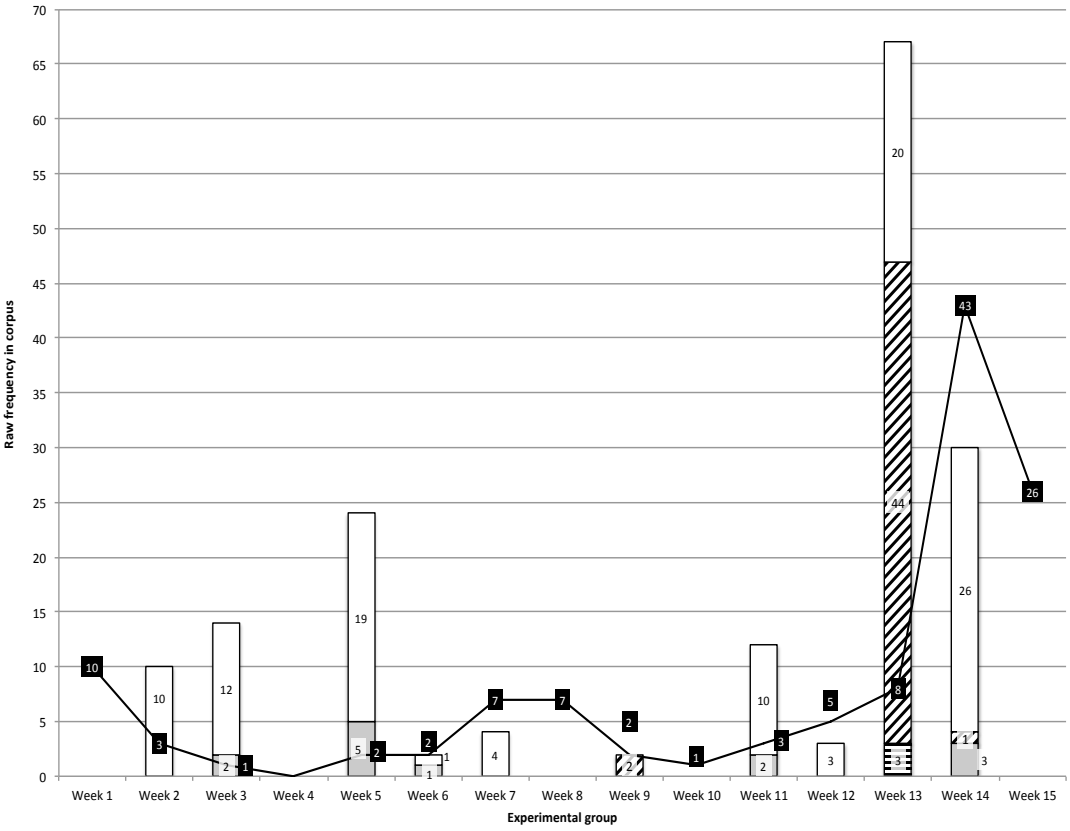


Figure 6.15. Input and output of the CULTURES ARE POSSESSIONS metaphor over time

Despite the clear differences in input caused by explicit instruction, there are also signs that the sequence of course topics has led to similarities in metaphor exposure. The CULTURES ARE LIVING THINGS theme was prominent in input during weeks four to seven of both conditions, for example, and a cluster of spikes can be seen in both conditions for several of the themes around weeks 12 to 14.

Looking at the changes in frequency of output over time, we can see the value of considering the data at different levels of granularity. Learner output, like classroom input, is shown to fluctuate greatly, with peaks on certain topics. What appeared a clear trend in Table 6.9 now appears much more chaotic, with wild oscillations from week to week. This is not to suggest, however, that the trend is no longer present; in most of the five themes, it can be seen that the higher peaks are skewed towards the final weeks of the semester, which is the cause of the larger fluctuations. Broadly speaking, learners did produce more culture metaphors as they progressed through the course, but it should be remembered that this trend was far from linear. Again, there is also a hint of similarity between the two conditions, with peaks appearing quite often at the same stage of the semesters.

It is unfortunate that the effects of topic cannot easily be separated from developments over time with this data set. Without running the study again with course topics introduced in a different order, this would be hard to achieve. It does appear that the topics and writing prompts are influencing production to some degree, as the peaks in output for most of the themes occur for the same topics. Week 14 in particular, saw high levels of output of culture metaphors. An examination of the data reveals that 28 of the 39 metaphors produced in week 14 for the CULTURES ARE SPACES metaphor were forms of *import*, and 37 of the 56

metaphors in the same week for CULTURES ARE POSSESSIONS were forms of *lose*. These appeared in response to the following writing prompt:

> *"Are the Japanese losing their culture? Is there too much American influence? Are the Japanese westernized or modernized?"*

Certainly, the use of the word *losing* in the prompt invites learners to respond with the same term. The word *import*, while not appearing in the prompt, also lends itself very naturally to discussing this topic. That being said, 25 of the 49 CULTURES ARE POSSESSIONS metaphors that learners were exposed to in that week were also a form of *lose*, and *import* had appeared 15 times out of 50 CULTURES ARE SPACES metaphors in the same period, so it could be argued that these forms were particularly prominent in the course at that time.

6.6. Concluding comments

This chapter has closely examined how a small group of metaphors appeared in classroom input and learner writing across the two conditions. It has revealed some clear trends but also tremendous complexity in the learning environment.

Considering classroom input, learners in the experimental treatment were exposed to nearly twice as many culture metaphors as their control group counterparts. Most of the culture metaphor themes saw increased levels of input, but there was some variation as some themes became much more prominent in the input. However, while raising the frequency of exposure was unproblematic, it was much more difficult to increase the spread of metaphors throughout the course; although taught target metaphors did appear in a greater proportion of lessons in

the experimental condition, the difference was not particularly large. This may be a valuable point to consider for future classroom investigations of metaphor. To some degree, the content of a course determines the language that will be used, so studies that focus on a large number of metaphors may well encounter the same issue. There would be merit in conducting similar studies in academic environments but limiting the range of metaphorical expressions taught so that learners encounter the same metaphors more often and are more likely to have repeated opportunities to add them to their productive lexicons.

Other insights from the analysis of course input were firstly, the degree to which spoken discourse dominated over written text as a source of input, and secondly, the greater amount of implicit learning opportunities compared with explicit instruction. It must be conceded that the researcher was also one of the class instructors, but these trends are apparent in both conditions, so it does appear that at least in this context, learners are more likely to encounter culture metaphors aurally and in situations in which the focus is on the content message rather than metaphor itself. Clearly, this needs to be compared with the findings of other studies as each course environment is unique, but it does have potential ramifications for establishing best teaching practices. It is increasingly claimed that explicit attention to vocabulary is necessary to boost learner uptake, and that a reliance on incidental learning is simply too slow, inefficient or unreliable to fully meet learners' needs (Hulstijn, 2001; Laufer, 2005, 2006, p. 161; Nation, 2013, p. 357). It must be conceded, however, that practical concerns will limit the amount of vocabulary that can be given explicit attention in class. This is the basic conundrum for vocabulary building; the most efficient method for learning is also the most time consuming. What the

findings of this chapter suggest is that even in a course with a strong focus on explicit metaphor awareness raising, the majority of input still came through implicit learning opportunities. This seems to suggest that explicit instruction should focus on lexis both at the word level (teaching particular forms) and also at the broader level of vocabulary awareness (examining lexical phenomena such as collocation, multi-word expressions and metaphorical patterning). The hope would be that training learners to be more aware of vocabulary patterning would allow them to make better gains from incidental learning.

What is clear from the analysis of learner output is that changes in productive language use are far from linear. The data showed mixed patterns of growth and decline of individual word forms, and there is a possibility that the treatment had the effect of encouraging learners to produce partially known word forms in new senses. This fits in with the view of vocabulary learning as an incremental process in which aspects of knowledge appear at different stages of development. Obtaining data on these developmental stages in lexical growth would be a huge challenge, but studies in this area would be tremendously valuable.

The data point to the need for consideration to be given to context in assessing productive language use. As might be expected, the frequencies of metaphorical vocabulary used in the classroom in which the study took place were shown to be much more closely related to learner output than those of large monitor corpora. This reinforces the view that in academic contexts there is merit in considering the language at the level of individual disciplines.

The longitudinal analysis of metaphor production revealed the strongest evidence of increased output over time. This analysis also suggested that untaught

target metaphor use was increasing as the semester progressed, which is a positive sign as the results of chapters four and five had indicated that this was not the case for the entire range of themes in the metaphor workbook. However, it also suggested that experimental group output was actually lower than that of the control in the early weeks of the course, which may indicate avoidance of language that was initially perceived as difficult on the learners' part.

CHAPTER 7. INFLUENCES ON METAPHOR PRODUCTION: WORD FREQUENCY, PART OF SPEECH, PHRASEOLOGY AND THE L1

7.1. Introduction and rationale for the investigation

Up to this point, the present study has considered the interactions between learner production of metaphorical words and various aspects of the classroom environment: treatment condition, learner ability, topic effects, language input, and time. This chapter will broaden the scope of inquiry to consider some factors that relate more to the influence of language itself.

The first of these is a comparison of metaphorical vocabulary usage at various frequency bands between the participants in this study and a comparable group of native speakers of English writing on similar topics. The frequency of a word has long been known to have a strong relationship with the likelihood of its being recognised by language learners (Palmer, 1917, p. 123), and several studies have demonstrated that this holds true for learners in different contexts and for words down to relatively low frequencies (Aizawa, 2006, p. 111 - 113, Beglar, 2009, p. 10 - 12; Brown, 2012; Milton, 2009, p. 25 - 42). There is some evidence that this relationship between receptive knowledge and frequency follows through into vocabulary production, with Laufer's (1998) study of lexical frequency profiles suggesting that learners gradually begin to increase the proportion of lower frequency words they use in written compositions as they gain in proficiency. Thus, it might be expected

that language learners would show a greater ability or willingness to use high frequency words in their metaphorical senses than words of lower frequency. Firstly, there is a greater likelihood that they would already be aware of metaphorical senses for these words, and even if only the basic sense were known, the learning burden of adding a new meaning to a known word may be lighter than for that of a less well-known or even totally new word. As Ozturk (2015, p. 106) and the findings of chapter six (Table 6.20, page 267) suggest, owing to contextual needs, productive use of vocabulary may not be particularly closely related to broad-scale estimates of word frequency. However, in this chapter, what was investigated was production of metaphorical vocabulary at a range of frequency bands by both groups of learners in this study and a group of anthropology students whose native language is English. The purpose of this was to obtain evidence on whether these groups of language learners were under- or overusing metaphorical vocabulary at any particular frequency range, and in the case of underuse, to identify particular word forms that language learners make less use of than native speakers. In other words, this process was intended to identify potential gaps in learners' productive use of metaphorical vocabulary.

There is also some evidence that word form itself has some influence on the learning of vocabulary. Nouns, particularly concrete nouns, tend to evoke a clearer mental image in the minds of language learners (Paivio, 1969), and this might make them easier to learn in some cases (e.g., Crossley et al., 2009, p. 328; Salsbury, Crossley & McNamara, 2011, p. 355). However, Na and Nation's (1985) study into guessing vocabulary from context found that verbs were easier to guess than nouns, so it is possible that the medium of interaction with words has some influence too.

How and to what extent this translates into productive usage, however, remains unclear. Complexity has been posited as one cause of linguistic avoidance in language learners, yet Laufer and Eliasson's (1993) study into avoidance of phrasal verbs found no evidence that learners avoided using verbs with more figurative (hypothesised as complex) meanings. But this is just one rather narrow example, and it does not negate the possibility that grammatical, rather than conceptual, complexity influences learners' lexical choices. The influence of word form is clearly an area in need of further investigation.

Thirdly, the interaction between learners' use of metaphor and phraseological aspects of language will be considered. Given that word senses are to a large degree determined by the lexico-grammatical surroundings in which they appear (Deignan, 2005, p.157 - 162; 2016, p. 114 - 118; Firth, 1957, p. 11 - 13), metaphor and phraseology can be seen as intimately linked, but this presents a problem for language learners, who tend to process meaning at the micro-, rather than macro-, level (Conklin & Schmitt, 2012; Wray, 2002, p. 206 - 208). The importance of this point for teachers attempting to raise awareness of content-appropriate metaphor cannot be overstated. The data obtained during this study offered several examples of how learners can show knowledge of the extended meaning of a lexical item yet still lack the phraseological awareness to use it without creating ambiguity or unintended comic effects. In examples (93) and (94), learners had been taught the metaphorical use of *explode* as part of the metaphor theme MORE OF AN ABSTRACT THING IS AN INCREASE IN SIZE OR HEIGHT, but had not appreciated the need for an abstract subject collocate to avoid a literal interpretation. The error in (95) is more related to colligation, with the adjectival form of *absorb* being required to activate the sense of

being totally focused on something. (96) is presented as an example of perhaps a more borderline case. This use of *spread* to describe cultural flow was very common in both conditions, but here, the concrete object and lack of an adverbial phrase blurs the distinction between literal and metaphorical interpretations. Examples with concrete referents could be found in large corpora, as in (97), but not in a transitive pattern.

(93) She was born in Japan and exploded in Japan. [EXP22]

(94) When American militaries came to Okinawa, Tacos was brought to Okinawa by American because Tacos was exploded in America. [EXP19]

(95) Sumo makes old people absorb. [CON23]

(96) I think that it is good things for many people to spread Japanese food. [EXP21]

(97) And that food is spread throughout the world. [COCA]

Yet rather than viewing this as problematic, it should probably be seen as a developmental stage. As learners begin to stretch the meaning potentials of known lexical items, they will inevitably make errors such as these. What is needed then is for researchers to examine how learners begin to add formulaic language to their output. The data obtained in this study offer an opportunity to consider whether experimental group output differs from that of the control in terms of the quantity or variety of collocation and colligation features used and whether learners who receive instruction on metaphor are able to produce it in ways that would be recognised as idiomatic by native speakers.

This raises the question of how recognition might best be determined. Non-native language use has frequently been compared to the norms of native-speaker

production, whether that be a corpus or a more traditional codification of language such as a dictionary, grammar book, or pronunciation manual (Bamgbose, 1998, p. 2). In this case, native speaker often refers to speakers of standard varieties of English from an 'inner circle' country (Kachru, 1991, p. 186). This has been justified on the grounds that the ultimate goal of instruction is to raise proficiency levels and that this entails a closer approximation to native-speaker norms (Granger, 2002, p. 13), but it must also be conceded that a sole emphasis on the supposed authenticity of one or other varieties of English overlooks the reality that learners in many contexts are studying the language in order to use it as a lingua franca or as a means of furthering the international development of their own society (Granger, 2002, p. 29; Widdowson, 1997, p. 142 - 145). The Japanese context is described by Kachru as part of the 'expanding circle' – nations whose use of English is primarily for international or specific purposes rather than serving an internal function in government, education, or mainstream culture (Kachru, 1991, p.179, 1992, p. 1, Kachru & Nelson, 2001, p. 13). This remains an accurate description of the role of English for the participants in this study, who were being taught primarily in English, were due to spend around three months on a study abroad program in one of Kachru's 'inner circle' countries the following semester, and whose use of English post-graduation was, owing to Japan's economic links with the US, most likely to be as an international language or occasional lingua franca. How then, might the realities of this context be considered in making decisions on whether language use is idiomatic, innovative, or erroneous? Bamgbose (1998, p. 2 - 5) offers a useful distinction between external criteria, which typically take the form of native-speaker norms, and internal criteria, which take account of local factors such as demography and

codification or acceptability in that particular context. Since the learners in this study were of relatively low proficiency and were likely to need English for communication with native speakers of that language, an argument can be made for reference to native speaker norms (Bamgbose, 1998, p. 12 - 13), particularly with an emphasis on providing supportive, rather than prescriptive, feedback (section 3.6.4, page 119). In considering internal factors, the language of the classroom itself was felt to be the best guide, since English has no official role in Japan and, as has been mentioned in previous chapters, individual academic disciplines and topics lend themselves to the use of particular language forms.

A further point to consider is the impact of learners' first language on their L2 production. Even in tertiary level educational settings, where learners might be expected to have amassed considerable experience of an L2, this will be dwarfed by their exposure to L1 forms and norms. This base of linguistic knowledge can play a facilitative role in learning new words (Laufer & Girsai, 2008), but in other respects, the L1 may be less helpful. This is often the case if learners are required to produce L2 speech or writing, when the fact that they are attuned to the usage patterns of their native language leads them to replicate this in L2 output. Lexical errors are sometimes ascribed to L1 influence, and this is particularly the case with regard to collocations (Hemchua & Schmitt, 2006, p. 20; Nesselhauf, 2003, p. 234 - 237). Since metaphor and collocation are so closely entwined, the potential for the L1 to influence this element of learner output is particularly high, and research has begun to argue that L2 collocational errors can be traced back to either disparities in phraseological patterning (Philip, 2006, p. 898; 2007; 2010) or variations in how shared metaphor themes are realised across languages (Danesi, 1994, 1995).

Findings in applied and cognitive linguistics as well as psycholinguistics strongly suggest that the four factors described here are likely to exert a considerable influence on learners' metaphor production, and it would be remiss not to give them due consideration. The research questions that will guide this chapter are as follows:

a) To what extent did learner metaphor production follow the pattern predicted by word frequency? Did learners show signs of being overly dependent on high frequency words to create metaphor?
b) Did learners appear to favour or avoid any particular part of speech in the production of metaphor?
c) Did the language used by either group of learners follow typical English phraseological patterning? Did the explicit focus on metaphor in the experimental condition lead to greater or more varied production of phraseologically well-formed metaphors?
d) Can evidence be found of learners' first language influencing their L2 metaphor production? Does this have a facilitative or debilitative effect?

7.2. Methodology

7.2.1. Datasets used in the study

This section of the study was based on the same datasets as were used in chapter five, section 5.6 and chapter six. That is, polywords were identified by their presence in either the Macmillan or Longman dictionaries, and a small number of target words were coded as metaphorical despite their being recorded in separate entries in the Macmillan dictionary.

Unlike chapter six, which only analysed metaphor themes to express culture, this chapter drew on the entire data set. However, due to the prominence of culture as a course theme, many of the findings reported here involve metaphors used to describe culture. Sections 7.5 - 7.6 deal with collocations, colligations, and L1 influence. Owing to the complexity of these analyses, smaller subsets of the data were used to identify patterns. These subsets will be explained in sections 7.2.3.3 - 7.2.3.5.

7.2.2. Data sources

This chapter considers the influence of external factors on learner output of metaphorical vocabulary. For this reason, a number of reference corpora were used to provide information on word frequencies in other contexts, and the feedback form described in section 3.6.4 (page 119) was used to obtain data on the sources that learners drew on for their metaphorical output.

7.2.2.1. Reference corpora

In order to compare the output of the participants in this study with that of native speakers of English while minimising the effects of writing topic, it was necessary to obtain a reference corpora of anthropological writing, and the British Academic Written English Corpus (BAWE) was selected for this purpose. The BAWE corpus is a collection of academic writing from a range of disciplines produced by undergraduate and Master's degree students at four universities in the UK. The assignments were all of at least a satisfactory standard, as determined by the students' instructors (Alsop & Nesi, 2009, p. 71 - 73). Alternatives to BAWE were the

Michigan Corpus of Upper-Level Student Papers (MICUSP) or the Louvain Corpus of Native English Essays (LOCNESS), but these were rejected since neither contains a section of papers from anthropology courses.

The entire BAWE corpus contains over 6.5 million words of student writing from around 30 disciplines (BAWE, n.d). Although the corpus contains 49 anthropology texts, six are by students whose L1 is either Slovenian or Swedish, and many are from courses that are in other branches of anthropology to the Japanese Popular Culture course that was the context for this study, such as Primate Conservation and Archaeology of Human Origins. For this study, a small sample of texts were selected that were as similar in topic as possible to those produced by learners in this study. All came from modules related to Social Anthropology, and four were from modules that focussed on Japan. In order to match the ages of the participants as closely as possible, it was originally intended to select only texts from first- or second-year students, but one text from a third-year was added because the topic, Japanese anime, was closely related to the themes covered in this study. All students listed English as their first language. In total, 18 texts were selected, comprising 28,689 words[41]. The full details of the selected BAWE texts can be found in Appendix N (page 489). Metaphor-related language in these texts was then identified using the variant MIPVU procedure described in section 3.4.2.2 (page 68).

A second rater, also trained in the use of MIPVU to identify linguistic metaphors, then coded 10% of the data (2,475 lexical units). The results of the two separate analyses were compared and found to be in agreement (coding items as indirect, direct, or implicit MRW or non-MRW) in 92.97% of cases, which yielded an adjusted

[41] The MIPVU lexical unit count differs from this figure.

for chance Cohen's Kappa value of 0.70. While this indicates good agreement, instances of disagreement were discussed in order to identify errors and to improve consistency in coding. Over 55% of cases of disagreement in the first round of coding had been prepositions, determiners, or pronouns, so discussion focussed on these forms initially. After these cases had been resolved, agreement reached 97.17%, with a Kappa value of 0.89. Attention then turned to open-class forms. Once these had been discussed, agreement was 99.31%, and Kappa 0.97. As with the analysis of learner data, the results of coding decisions were re-applied to the whole sample of BAWE texts to improve consistency and accuracy.

The second reference corpus used in this chapter was COCA (Davies, 2008-). This is a balanced corpus of American English that contains over 520 million words of text. The corpus was used to obtain estimates of the frequency with which particular words appear in English discourse (section 7.3) and to estimate the strength or frequency of the collocations and colligations found in participants' written output (section 7.5).

The final section of this chapter considers the influence of the first language on learner metaphor production. A reference corpus was used to provide evidence for the existence of metaphor themes in Japanese similar to those taught in this study and to determine the frequency of particular metaphorical collocations in the learners' L1. The jpTenTen11 corpus, an 8.4 billion-word web-crawled corpus of Japanese, was selected for this purpose as its size allows for searches of infrequent specialised word combinations and because the Sketch Engine (https://the.sketchengine.co.uk/open/) website on which it is stored offers a wide

array of search capabilities that facilitate pattern recognition (Srdanović, Suchomel, Ogiso & Kilgarriff, 2013).

Finally, the input corpora described in section 3.6.2 (page 106) were again drawn on in order to provide evidence of more localised discourse norms. This was typically the case when there was a need to verify that the specialised nature of classroom discourse led to the use of language that was infrequent or unverified in larger corpora.

7.2.2.2. Sources of metaphor forms

The forms described in section 3.6.4 (page 119) were used to obtain information from participants on sources of metaphorical vocabulary that they had used in their writing. Learners could indicate whether they had used monolingual or bilingual dictionaries, had translated directly from the L1, or had used language they had learned either in that course or elsewhere. This provided an estimation of how much learners were consciously drawing on their first language in comparison with other resources available to them as well as allowing for both successful and unsuccessful metaphor use to be related to each source of information. In cases where learners had translated, the Japanese translations provided evidence of the metaphorical influences from their L1.

7.2.3. Data coding

7.2.3.1. Norms

As was mentioned in 7.1 above, both external and internal factors were considered in the determination of norms of production. External evidence of linguistic norms

was obtained from the COCA and BAWE. In these cases, this evidence was used not for prescriptive purposes, but to draw comparisons between learner and native-speaker output so as to be able to describe learner writing at this stage of development and to identify language features that might be suitable for teaching in further studies. The data from these analyses could also act as a benchmark against which to compare the performance of other learners at different levels of proficiency or in other contexts. One of the counter arguments to the criticisms of using native speaker norms to judge non-native language performance is that there is first a need to describe international English in all its varieties (Granger, 2002, p. 28), and it was hoped that the results of this study can play a small part in this.

Since English has no official role in Japan but does function as an important medium through which the country carries out its interactions with the world, the internal factors that were most relevant to this context were those of the classroom itself. As will be seen in the following analyses, certain language, while not particularly prevalent in large corpora, could be shown to feature more prominently in classroom discourse, and this highlights the need to consider more localised norms.

7.2.3.2. Error coding

As well as giving consideration to usage norms, it was also necessary to address the ways in which learners' metaphor production deviated from standard usage. As examples (93 - 96) demonstrated, adding a particular language item to the range of expressions available for productive use requires more than an understanding of meaning and form; each item needs to be successfully integrated with other

linguistic elements or there is a risk that communication will be impeded. Such changes are likely to take time, and it is recognised that interlanguage is an important feature of learner development (Selinker, 1972). Accordingly, the ways in which learners' written production showed signs of error were examined.

In order to codify examples of erroneous usage, the error tagging system developed for the National Institute of Information and Communications Technology Japanese Learner of English Corpus (NICT JLE) was used. This tagset was designed to be relatively simple to use by focusing on error at the linguistic level rather than broader features of discourse, and by having a relatively limited number of tags (Izumi, Uchimoto & Isahara, 2005, p. 75). In total, 47 tags are available, classified into 11 part-of-speech groups and an 'others' group. The part-of-speech groupings and classifications of error by type, rather than offering varying levels of specificity, as in other tagsets (Díaz-Negrillo & Fernández-Domínguez, 2006, p. 90 - 91), were felt to be consistent with the data obtained from MIPVU.

As the focus of this investigation was on open-class metaphor use, only the error tags for nouns, verbs, adjectives, and adverbs, as well as the 'others' group, were used in the analyses in this chapter (see Appendix O, page 491). To consider in broad terms the issues learners face when trying to use metaphorical language in writing, the error tags were divided into three main categories: grammatical, phraseological, and lexical. Grammatical errors included inflected forms, countability of nouns, tense, aspect, voice, and other language features that ground an expression in terms of its relation to time and place, and to its participants and their interactions (Langacker, 2008b, Chapter 9). Phraseological errors involved the relationship between the metaphorical word and those in its proximity. These included complements,

collocations, and cases of word misordering. The category of lexical errors contained cases in which the metaphorical word itself was an inappropriate semantic choice.

One further point to discuss regarding coding is the distinction between the lexis and collocation tags. Clearly, there is no clear demarcation point between an error that resides in a single lexical item and one that involves its immediate neighbours, and nor is there a clear line between collocates and non-collocates. However, the distinction is useful in the examples below. Example (98) contains a collocation that is found in COCA (although the use of *culture* in the subject position here makes it somewhat marked), but in this case *come into* would seem the more appropriate choice. Example (99), meanwhile, contains a collocation that is well evidenced in COCA, but as a whole unit does not convey the intended meaning, which was presumably closer to 'most common opinions.' These errors seem to be more of a semantic nature than (100), which uses *big* to modify *culture* (a rarer combination of words) and cannot easily be rectified without introducing other structures (e.g., 'biggest elements of Japanese popular culture,' 'most popular Japanese cultural exports'). Thus, (98) and (99) were given the code 'Collocation: lexis' and (100) was coded as a 'Collocation: phrase'. These separate codes were not in the NICT JLE tag set, but were felt to be useful here in distinguishing between semantic and combinatorial issues in producing collocations.

(98) Until Meiji era, foreign culture cannot go into the Japanese culture because of the national isolation. [EXP22]

(99) One of the strongest opinions is that pasta was originated in China as early as 1700 B.C.E. [CON14]

(100) Nowadays Japanese Anime culture can be one of the biggest Japanese popular culture... [EXP7]

Finally, two codings were retained that fell outside this categorization system; *unknown type errors* were those that could not be classified under any of the other codes, and *Japanese English* was a code reserved for cases where error may have been caused by *wasei-eigo* (see section 2.6.1, page 34).

7.2.4. Analytical procedures

The analyses performed in this chapter used the data sources outlined above to investigate the effects of frequency, phraseology, and the first language on learners' use of metaphor. Because the methods of analysis differed from those of previous sections, each will be described below.

7.2.4.1. Frequency analysis

In order to investigate the words that had been used metaphorically by both learners in this study and native-speakers in the BAWE corpus, the data-sorting function of Excel was used to produce lists of lexical units that had been tagged as MRW for each part of speech during the MIPVU procedure. The counts for inflected forms were then combined with the head word. The frequencies of each metaphorical group were calculated and then normalised to a per-10,000-word rate to allow for comparisons to be drawn between conditions.

In order to compare metaphor use and word frequency between both study conditions and native speakers, an experimental design similar to that used by Durrant and Schmitt (2009, p. 168) was employed. Rather than making a blunt and

somewhat arbitrary distinction between frequent and non-frequent words, COCA was used to provide reliable estimates of word frequency across a range of values. The proportion of MRWs in the learner corpora and BAWE could then be compared against these values, and finer distinctions could be drawn for metaphor use across a range of frequencies. It should be noted that the frequencies used from COCA were for all instances of that word, not just metaphorical uses. This was because the hypothesis being tested was that learners would be more willing or able to use metaphors for words that were generally more frequent, and thus more likely to be known. The data for the frequency and part of speech analyses is provided in Appendix P (page 493).

7.2.4.2. Parts of speech

The comparison of metaphor use by part of speech between language learners and native speakers drew on the same lists of MRWs described in 7.2.4.1. To investigate whether a particular group made greater use of metaphorical language from a given part of speech, log-likelihood ratio calculations were made to compare the relative use of nouns, verbs, phrasal verbs, adjectives, and adverbs by both the learners in this study and the native speakers who produced the texts sampled from the BAWE corpus. The rates of use for all instances of usage were calculated, as were instances of metaphorical usage. This allowed for investigation of whether learners under- or overused these parts of speech in general or just in terms of metaphorical use.

7.2.4.3. Collocations

For judgments about phraseological patterns to be made, it was necessary to identify constructions that occurred frequently enough for patterns to be apparent.

For this reason, broad categories of analysis were selected based on an inspection of the part of speech lists. To investigate collocation, metaphorical adjective/noun combinations and metaphorical verb collocates of the noun *culture(s)* were chosen. The selected metaphorical adjective/noun collocations appeared 87 times in the control and 128 times in the experimental condition, and metaphorical verb collocates of *culture* appeared 118 times in the control and 185 times in the experimental group. Collocations produced by learners in both conditions were analysed for both the strength and the frequency with mutual information (MI) scores and t-scores respectively. Collocation and colligation data is provided in Appendix Q (page 501).

7.2.4.4. Colligations

Colligation was investigated for both frequently occurring target metaphorical verbs and nouns that belonged to the CULTURES ARE CONSTRUCTIONS metaphor theme. Each verb and all but one of the nouns appeared at least 20 times across the study. These phraseological constructions were compared with data from COCA to ascertain whether learners in either condition were under- or overusing particular patterns. The error tags described in section 7.2.3.2 were also applied to learner output of these forms to consider which factors influenced accuracy of output and whether accurate usage differed between either ability groupings or treatment conditions.

7.2.4.5. L1 influence

The influence of the L1 on learner production of metaphor was investigated with two sources of data. Firstly, feedback was obtained directly from learners on the sources of the metaphorical language they produced (as described in section 3.6.4,

page 119). The second approach involved using corpus data to estimate the effect of L1 frequency on L2 output. Owing to the challenges of making judgments about translations and transfer effects, this part of the analysis was limited to metaphorical verb collocates of *culture(s)*, or 文化 (*bunka*) in Japanese. The jpTenTen11 corpus, a very large corpus compiled from Internet sources and available on the Sketch Engine website, was used as a reliable source of L1 data. As well as functioning as a concordancer, Sketch Engine offers the ability to produce 'word sketches', or easily obtained lists of the collocational and colligational behaviour of target words. This function was used to understand how *bunka* interacts with other words in Japanese. Once data from the L1 had been obtained, inferences could be made for how the how the first language might have affected specific cases of metaphor production, and the competing influences of L1 and L2 frequency on learner writing could be compared. Data on feedback forms and L1 influence is provided in Appendix R (page 509) and Appendix S (page 511).

7.3. Learner metaphor production and word frequency

In order to visualise the relationship between frequency and metaphor use in the two conditions and how it might differ from that of native speakers, two charts were plotted. In the first, the percentages of all open-class metaphorical tokens that fell into each of a range of COCA frequency bands for each corpus was calculated, and the results are shown in Figure 7.1. At first glance, the distribution of metaphors may appear somewhat unusual, with peaks at either end of the scale. However, this is due to the competing influences of a small number of extremely high frequency words on the right side of the scale and a huge number of low frequency words on

the left. The figure shows that between 30% and 40% of the tokens in the two learner corpora were highly frequent words in COCA, occurring at least five times per 10,000 words. The value for the BAWE data is somewhat lower at around 20%, but this still comprises a relatively large proportion of that corpus. This is entirely predictable, as other studies into word frequency have shown that a very large proportion of text is covered by a relatively small group of high frequency words (Nation, 2006, p. 64; Schmitt, 2010, p. 68 - 71). It is also recognised that polysemy, much of which might be explained as metaphorical extensions of meaning, is very common amongst high frequency words, so we should expect a lot of metaphors in this frequency band. Log-likelihood tests revealed that these highly frequent words comprised a significantly greater proportion of both the control and the experimental group output corpora than the BAWE corpus, with positive effect sizes present (Tables 7.1 - 7.2). This is a similar finding to Durrant and Schmitt (2009, p. 170 - 172), who also noted learners' tendency to overuse highly frequent lexical patterns. There is also some suggestion that experimental group learners produced more of these high-frequency open-class metaphors than learners in the control group, although the effect size is not large in this case.

Appendix P (page 493) lists the particular metaphors that were found to significantly favour one corpus over another in this analysis. The section for words occurring at least five times per 10,000 words reveals that despite the overall strong preference for words in this band to favour the learner corpora, several of the individual metaphorical words were much more frequent in the BAWE corpus. Examples include *give*, *show*, and *see* (101 - 106), all of which were used by learners, but not to the degree of the native speaker output.

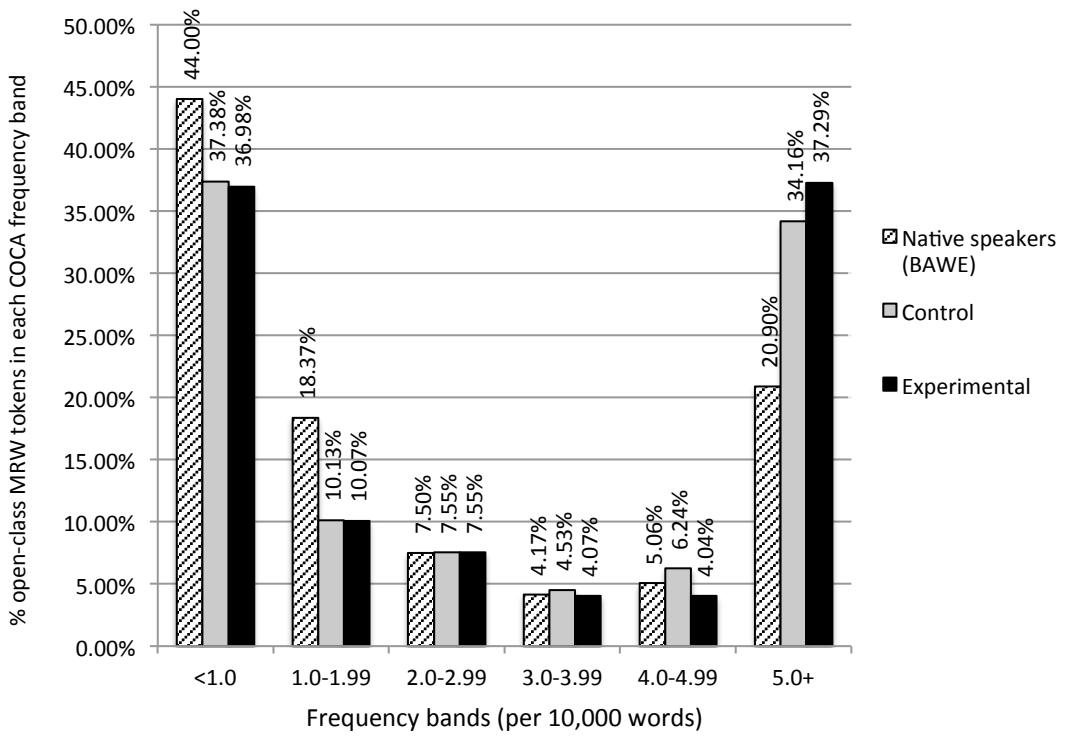


Figure 7.1. Percentage of open-class metaphor tokens in the output corpora and BAWE sample falling into a range of COCA frequency bands

Table 7.1. Log-likelihood tests for proportions of open-class metaphor tokens in each COCA frequency band between the output corpora and BAWE sample

		1. Control	2. BAWE	1. Experimental	2. BAWE	1. Control	2. Experimental
		Log-likelihood [a]	Bayes factor effect size	Log-likelihood	Bayes factor effect size	Log-likelihood	Bayes factor effect size
COCA frequency band (per 10,000 words)	< 1.0	-94.08****	82.96^{+++}	-106.94****	95.64^{+++}	0.04	-11.48
	1.0-1.99	-128.68****	117.56^{+++}	-143.66****	132.36^{+++}	0.00	-11.52
	2.0-2.99	-6.40*	-4.72	-6.91**	-4.39	0.00	-11.52
	3.0-3.99	-1.87	-9.25	-4.75*	-6.55	0.80	-10.72
	4.0-4.99	-0.26	-10.87	-14.95***	3.65^{+}	15.94****	4.42^{+}
	5.0+	15.48****	4.35^{+}	34.64****	23.34^{+++}	-4.88*	-6.64

a = Negative values indicate underuse in corpus 1 relative to corpus 2 (positive values overuse in corpus 1)

* = Significant at the 0.05 level; ** = 0.01 level; *** = 0.001 level; **** = 0.0001 level

+ = Positive evidence against the null hypothesis; +++ = Very strong evidence

(101) *Cultural hybridity and hybridism give Japan opportunities to exchange idea with other cultures.* [CON14]

(102) *Aspects of society are presented as they are 'classified and ordered by the Japanese people,' giving the reader insight into their individual culture.* [BAWE3135b]

(103) *These component of Taco rice's history show that even though foreign country dominates another country, it may be good for food or something like that because the dominated country can get other country's culture.* [EXP19]

(104) *Through his comparison he wanted to show that different cultural areas with no historical contact, show similar phenomenon in culture...* [BAWE3014a]

(105) *The habit is seen as a sacred practice in Japan...* [EXP1]

(106) *Conversely to non-gypsy opinion this behaviour is seen as respectful of gypsy tradition...* [BAWE3055a]

At the next most-frequent level (4.0 - 4.99 per 10,000 words), both the BAWE writers and the control group produced more metaphor than learners in the experimental condition, but otherwise the mid-frequency vocabulary showed few differences between the three corpora. At the level of individual words, several nouns appeared to be more prominent in the writing of native speakers than language learners. *System*, *area*, and *power*, and to a lesser extent, *sense* and *field*, all favoured the BAWE corpus over the learner corpora, while there were no nouns which favoured the language learners. Three words (*long*, *lose*, and *develop*) were consistently used more often in the learner corpora than the BAWE sample. As was mentioned in section 5.6.2.1. (page 206), 78% of the learner uses of metaphorical

long immediately preceded the nouns *history* or *time*, which suggests that these collocations are safe choices in learners' productive vocabularies, but also that there may be some over-dependence on these patterns. The verbs *lose* and *develop* were both used frequently to describe cultural change. *Lose* was a taught target metaphor, while *develop* was used significantly less often in the experimental condition than the control, perhaps due to the teaching of the near synonyms *grow*, *evolve*, *spread*, and *advance*.

At the less frequent end of the scale in Figure 7.1, there is evidence of the presumably greater vocabulary size of native speakers. Both words in the 1.0 - 1.99 range and those occurring less than once per 10,000 words in COCA were very strongly favoured by writers in the BAWE corpus than either language learner group (Table 7.1). However, as with the highly frequent words, some individual metaphors were still found to favour the learner corpora. These included adverbial phrases (*in addition*, *at first*, *on the other hand*, *and so on*), target metaphors (*born*, *spread*, *absorb*, *import*), and vocabulary that was related to the course topics (*idol*, *hybrid*, *gothic*, *lolita*). Metaphorical nouns (*view*, *form*, *approach*, *order*, *structure*, *element*, *context*, *perception*, *boundary*) were again shown to strongly favour the native speaker corpus. Only the last three of these were completely absent from learner output, so some degree of knowledge of these words can be assumed. However, apart from *order*, these words all appear on published lists of academic vocabulary (Coxhead, 2000; Gardner & Davies, 2014), and it may be that the learners' relative lack of familiarity with this genre meant such expressions were less activated in many individual's productive vocabularies. The other group of word forms that was shown to be more prominent in native speaker writing was verbs. In particular,

metaphorical verbs that are common in academic discourse (Coxhead, 2000; Gardner & Davies, 2014) such as *present*, *apply*, *maintain*, *reflect*, *demonstrate*, *link*, *illustrate*, and *highlight* appeared significantly less often in learner writing. These verbs are particularly useful for moving beyond the simple description of a topic towards the relating of ideas, drawing of conclusions, or adding further comment that is typically expected of academic writing (107 - 109), and may therefore be a suitable focus for further studies on developing learner use of academic vocabulary in writing.

(107) This only goes to demonstrate that our perception of religion can be clouded by classification. [BAWE 3088b]

(108) I have highlighted the most prominent elements of Japanese society that Allison has presented... [3135b]

(109) The amount of time that employees devote to their jobs also reflects part of the Japanese work ethic... [3135b]

The second chart (Figure 7.2) plotted types of metaphors used by learners at each of the COCA frequency bands. The values are the mean proportion of metaphor types that fell into each frequency band for the participants in each condition, presented as a box and whisker diagram as in chapter five. Kruskal-Wallis tests revealed significant differences in types of metaphor produced in the three corpora in the 1.0 - 1.99, 4.0 - 4.99, and 5.0+ frequency bands (Table 7.2). Post-hoc Mann-Whitney tests showed that the BAWE corpus writers produced a significantly greater

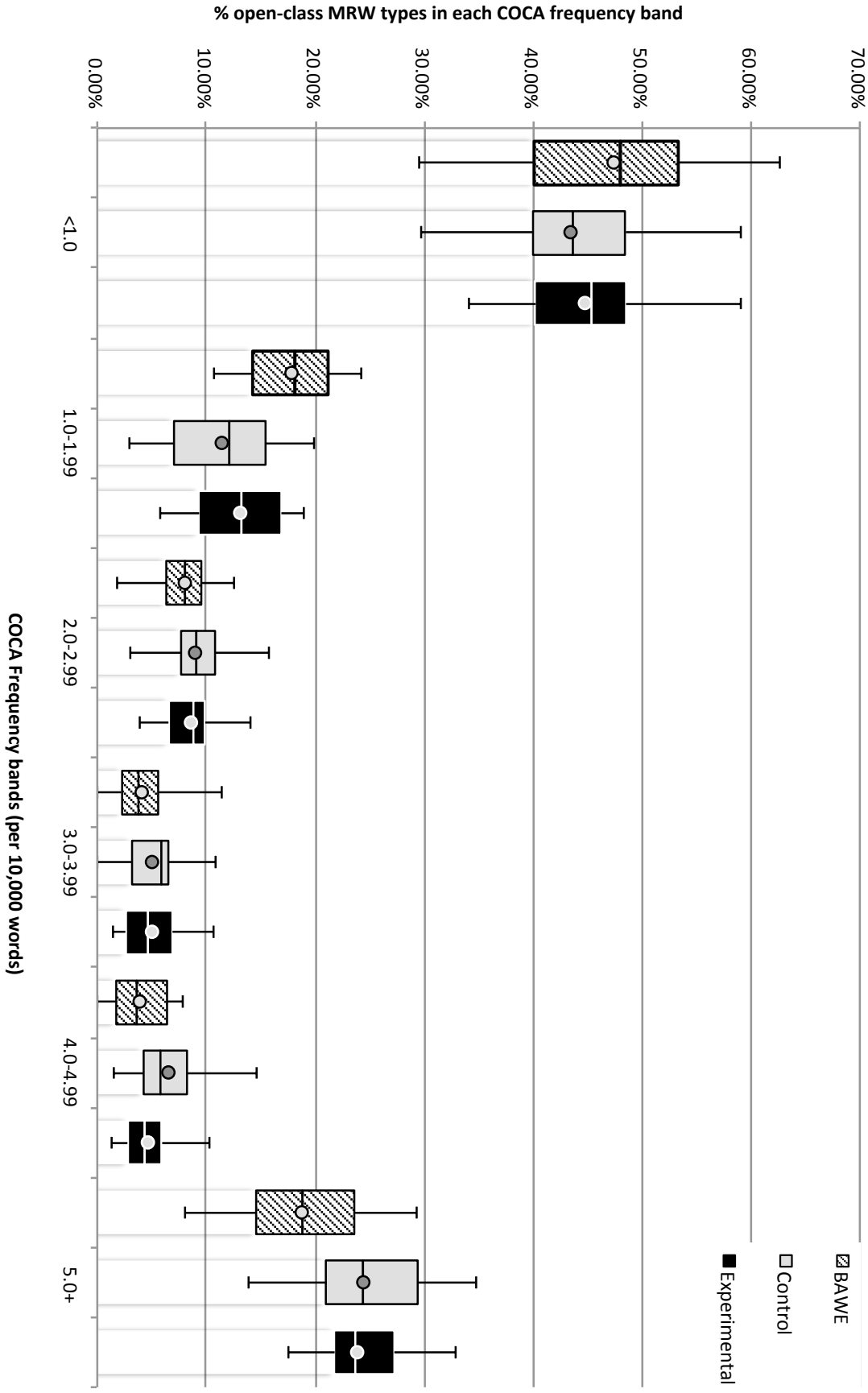


Figure 7.2. Distribution of open-class metaphor types in the output corpora and BAWE sample falling into a range of COCA frequency bands

Table 7.2. Kruskal-Wallis tests for proportions of open-class metaphor types in each COCA frequency band between the output corpora and BAWE sample

	Frequency band	*n* [a]	Mean rank			Chi-square	*df*	*p*
			Con.	Exp.	BAWE			
COCA frequency per 10,000 words	<1.0	64	28.96	31.91	37.78	2.302	2	.316
	1.0-1.99	64	23.98	30.09	46.47	15.343	2	.000*
	2.0-2.99	64	35.98	32.20	28.44	1.664	2	.435
	3.0-3.99	64	35.93	33.74	26.53	2.737	2	.254
	4.0-4.99	64	40.93	29.41	25.67	7.779	2	.020*
	5.0+	64	38.04	35.78	21.22	9.359	2	.009*

a = Control = 23, Experimental = 23, BAWE = 18
* = Significant difference found

variety of metaphors from the 1.0 - 1.99 band than either of the learner groups, the control group used a greater variety of metaphors from the 4.0 - 4.99 band than the BAWE writers, and both learner corpora contained a wider variety of metaphors from the 5.0+ band than the BAWE corpus (Table 7.3).

It is interesting that the native speakers did not produce a significantly greater variety of metaphor types from the <1.0 band than the learner groups, although the sheer amount of words that are available in this frequency range may have influenced this result. The native speakers' range of vocabulary use can be more easily seen in the 1.0 - 1.99 band, however, where there is strong evidence of their drawing on a much greater variety of metaphorical vocabulary than either learner group. For high frequency words, it seems probable that these learners, owing to their comparative lack of alternative lexical choices, are forced to keep returning to particularly well-known words, such as *have*, *big*, *feel*, and *thing* in order to express themselves. The general trend across the range of frequency bands appears to be that native speakers use a greater variety of metaphors from less frequent bands,

while language learners draw on both greater quantities and varieties of high frequency words.

Table 7.3. Mann-Whitney tests for proportions of open-class metaphor types in each COCA frequency band between the output corpora and the BAWE sample

		Corpora	*n*	Mean ranks			*U*	*z*	*p*	*r*
				BAWE	Con.	Exp.				
COCA frequency band (per 10,000 words)	<1.0	BAWE - Control	41	23.81	18.80	/	156.500	-1.327	.185	0.28^{+}
		BAWE - Experimental	41	23.47	/	19.07	162.500	-1.169	.242	0.24^{+}
		Control - Experimental	46	/	22.25	24.85	233.500	-0.681	.496	0.14^{+}
	1.0-1.99	BAWE - Control	41	28.56	15.09	/	71.000	-3.573	.000**	0.75^{+++}
		BAWE - Experimental	41	27.42	/	15.98	91.500	-3.034	.002**	0.63^{+++}
		Control - Experimental	46	/	20.89	26.11	204.500	-1.318	.187	0.27^{+}
	2.0-2.99	BAWE - Control	41	18.25	23.15	/	157.500	-1.301	.193	0.27^{+}
		BAWE - Experimental	41	19.69	/	22.02	183.500	-0.617	.537	0.13^{+}
		Control - Experimental	46	/	24.83	22.17	234.000	-0.670	.503	0.14^{+}
	3.0-3.99	BAWE - Control	41	17.36	23.85	/	141.500	-1.721	.085	0.36^{++}
		BAWE - Experimental	41	18.67	/	22.83	165.000	-1.103	.270	0.23^{+}
		Control - Experimental	46	/	24.09	22.91	251.000	-0.297	.767	0.06
	4.0-4.99	BAWE - Control	41	15.92	24.98	/	115.500	-2.404	.016*	0.50^{+++}
		BAWE - Experimental	41	19.25	/	22.37	175.500	-0.828	.408	0.17^{+}
		Control - Experimental	46	/	27.96	19.04	162.000	-2.252	.024$^{\#}$	0.47^{++}
	5.0+	BAWE - Control	41	15.25	25.50	/	103.500	-2.719	.007*	0.57^{+++}
		BAWE - Experimental	41	15.47	/	25.33	107.500	-2.614	.009*	0.55^{+++}
		Control - Experimental	46	/	24.54	22.46	240.500	-0.527	.598	0.11^{+}

= Non-significant after Bonferroni adjustment

* = Significant at the 0.05 level; ** = 0.01 level

+ = Small effect size; ++ = Medium effect size; +++ = Large effect size

Taken in combination, these results suggest that both groups of learners in this study are capable of using metaphors from a range of frequencies, but the factors that influence this differ across the frequency spectrum. There is evidence of words at the highest frequency levels being relatively overused by language learners in comparison with native speakers writing on similar topics, which implies a limited awareness of more precise or colourful synonyms. Exceptions to this were also present, however, with the words *see*, *give*, and *show* being more prominent in the native speakers' writing. These findings may be of use for further studies on Japanese learners of similar ability, with target words or expressions selected on the basis of learner over- or underuse in previous studies.

7.4. Learner metaphor production and part of speech

To investigate learners' favouring or disfavouring of particular parts of speech in the production of metaphor, the number of lexical units from each open-class part of speech in the three corpora from section 7.3 was identified. A series of log-likelihood calculations was then performed to establish whether there were significant differences between metaphorical usage of individual parts of speech across corpora.

The results of these comparisons between the three corpora are given in Table 7.4. It appears that the only part of speech that these non-native writers make significantly greater use of when compared with native speaker writers is adverbs, although even this trend is noticeably weaker in the experimental condition. There was a similar finding in chapter five, where adverbs, especially conjunctive adverbs, were shown to be much more frequent in the control than the experimental group. Metaphorical adjectives, nouns, and verbs, meanwhile, all appear to be more

common in the BAWE corpus than in learner writing, but in the case of adjectives and verbs this skewedness becomes weaker when comparing the experimental condition writers with the native speakers. Metaphorical adjectives were one of the target themes, and many of the taught target metaphors were verbs, so there is a possibility that this was a result of explicit instruction. However, the comparison of metaphorical parts of speech between the control and experimental groups in the right-hand columns of Table 7.4 show a significant difference for the distribution of adjectives, but not verbs, between the two learner corpora, and even here, the effect size is weak.

Table 7.4. Log-likelihood tests for the proportions of metaphorical parts of between the output corpora and BAWE sample

	1. Control 2. BAWE		1. Experimental 2. BAWE		1. Control 2. Experimental	
Part of speech	**Log-likelihood[a]**	**Bayes factor effect size**	**Log-likelihood**	**Bayes factor effect size**	**Log-likelihood**	**Bayes factor effect size**
Adjectives	-27.11****	15.99^{+++}	-9.77**	-1.53	-7.62**	-3.90
Adverbs	55.94****	44.82^{+++}	6.53*	-4.77	44.86****	33.35^{+++}
Nouns	-127.38****	116.26^{+++}	-128.35****	117.05^{+++}	-0.37	-11.15
Phrasal verbs	-2.26	-8.86	-1.82	-9.49	-0.07	-11.45
Verbs	-27.52****	16.40^{+++}	-18.71****	7.41^{++}	-1.93	-9.59

a = Negative values indicate underuse in corpus 1 relative to corpus 2 (positive values overuse in corpus 1)

* = Significant at the 0.05 level; ** = 0.01 level; **** = 0.0001 level

++ = Strong evidence against the null hypothesis; +++ = Very strong evidence

These results imply that the written output of language learners who are in the relatively early stages of acquiring a language will differ markedly in its use of open-class parts of speech in their metaphorical senses. However, what is not clear is whether this is a feature solely of metaphorical words or of the language system as a whole. Accordingly, a second series of log-likelihood calculations was made, this time

for all uses of open-class parts of speech, metaphorical or otherwise (Table 7.5). These values were then plotted against those from Table 7.4 in order to discern whether there was any difference between overall usage and metaphorical usage of individual word classes. Figures 7.3 and 7.4 show the plots for the BAWE corpus against the control and experimental groups respectively.

These charts suggest that while language learners in both conditions were much more likely to use nouns in general, metaphorical noun use was strongly skewed towards the native speaker writers. Thus, it might be said that the tendency to overuse nouns is reversed in the case of metaphorical nouns. The same pattern can be seen for verbs, with an overuse of verbs in general contrasted with a strong tendency for greater usage of metaphorical verbs by native speakers. While greater use of verbs by language learners may be an expected outcome, it is something of a surprise that learners produced a greater proportion of nouns. Previous studies have found a general pattern for learner writing, particularly that of less experienced writers, to bear more similarity to spoken discourse than academic text (Shaw & Ting-Kun Liu, 1998, p. 245 - 248). Granger and Rayson (1998, p. 127 - 128), too, found that advanced French learners of English adopted a more style more typical of oral communication, with both verbs and adverbs being overused at the expense of nouns.

For phrasal verbs, the category as a whole appears to be skewed toward the native speaker corpus, but there is no strong evidence of overuse in either corpus for metaphorical senses of these forms. Adjectives in general are relatively overused by the experimental group learners, but there is no clear trend in the control condition. In both cases, the learners appear to significantly underuse metaphorical adjectives

relative to the BAWE writers, although the effect size is weak in the experimental group. The final category of adverbs repeats the finding of chapter five, as well as that of Granger and Rayson mentioned above, that learners appear to overuse this form, strongly so in the case of the control group.

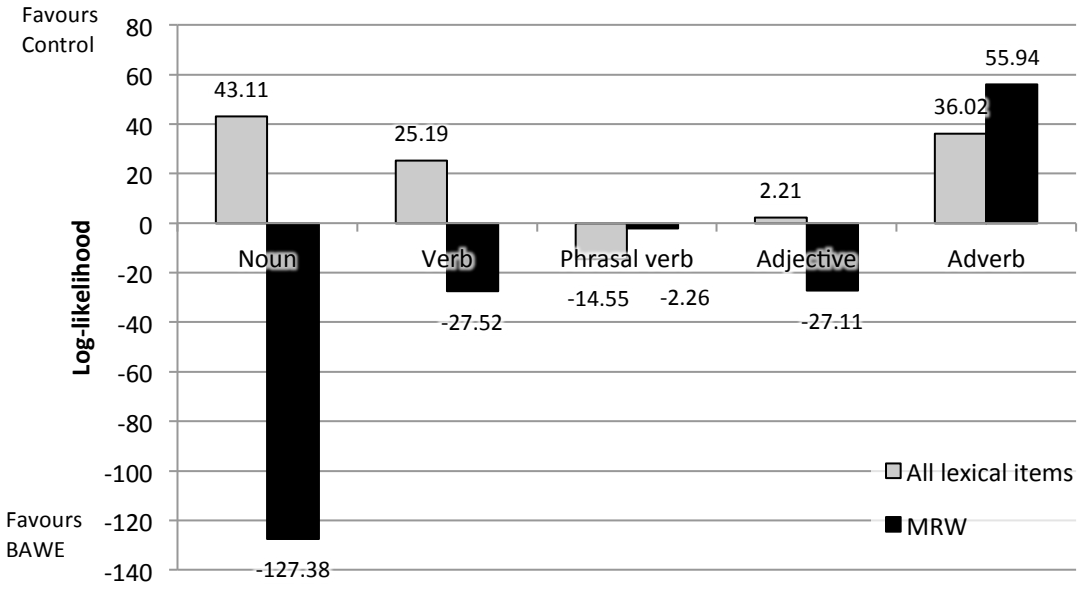


Figure 7.3. Log-likelihood values to show the relative frequency of metaphors and all lexical items between the BAWE sample and the control output for open-class parts of speech

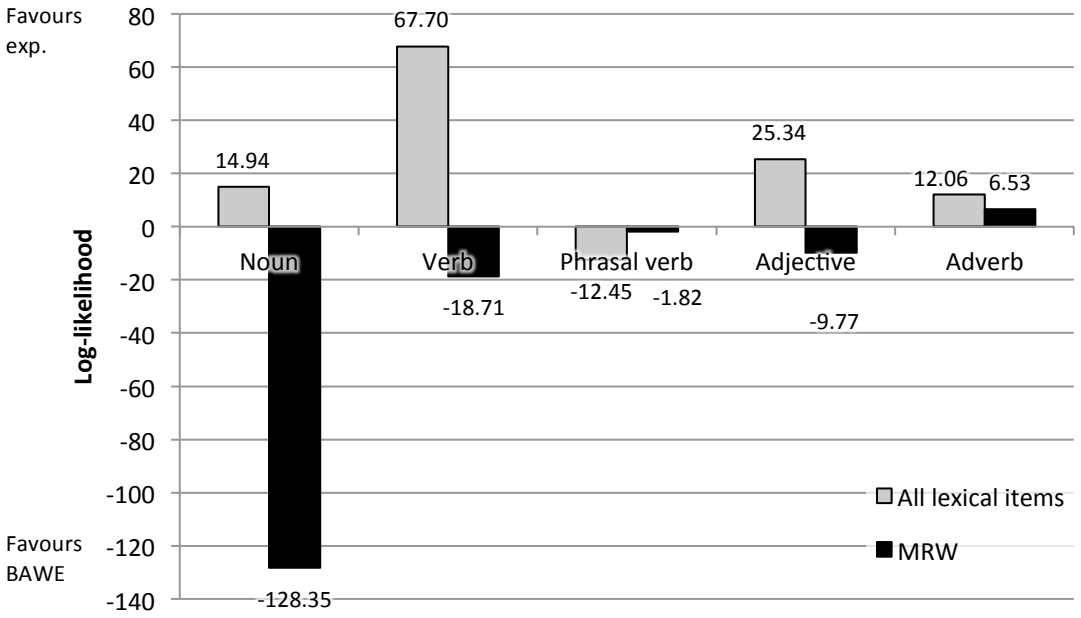


Figure 7.4. Log-likelihood values to show the relative frequency of metaphors and all lexical items between the BAWE sample and the experimental output for open-class parts of speech

Table 7.5. Log-likelihood tests for the relative frequency of all lexical units (MRW & non-MRW) from open-class parts of speech between the output corpora and the BAWE sample

Part of speech	1. Control 2. BAWE		1. Experimental 2. BAWE		1. Control 2. Experimental	
	Log-likelihood[a]	Bayes factor effect size	Log-likelihood	Bayes factor effect size	Log-likelihood	Bayes factor effect size
Adjectives	2.21	-8.91	25.34****	14.04^{+++}	-17.99****	6.47^{++}
Adverbs	36.02****	24.89^{+++}	12.06***	0.76	11.62***	0.10
Nouns	43.11****	31.99^{+++}	14.94***	3.64^{+}	13.21***	1.69
Phrasal verbs	-14.55***	3.42^{+}	-12.45***	1.15	-0.31	-11.21
Verbs	25.19****	14.07^{+++}	67.70****	56.40^{+++}	-13.43***	1.91

a = Negative values indicate underuse in corpus 1 relative to corpus 2 (positive values overuse in corpus 1)

*** = Significant at the 0.001 level; **** = 0.0001 level

+ = Positive evidence against the null hypothesis; ++ = Strong evidence; +++ = Very strong evidence

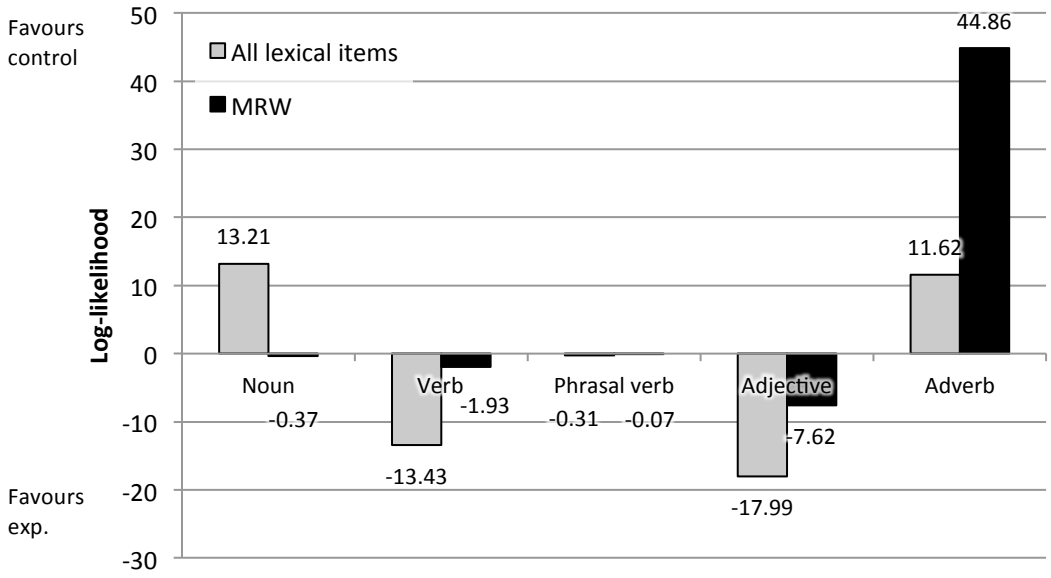


Figure 7.5. Log-likelihood values to show the relative frequency of metaphors and all lexical items between the control and experimental output corpora for open-class parts of speech

When the two groups of learners are compared using the same method (Figure 7.5), the large change in production of adverbs, both metaphorical and in general, becomes evident. This is almost certainly a reflection of the overuse of fixed adverbial phrases that happened to be metaphorical (e.g., *and so on*, *in addition*) that was shown in Figure 5.27 (page 217). For verbs and adjectives, it can be seen

that experimental group learners produced a greater proportion of these forms overall (with a strong effect size in the case of adjectives), but the distribution of metaphorical uses was only slightly tilted towards the experimental output corpus. It is likely that learners in both conditions were already familiar with metaphorical uses of highly frequent adjectives, and this may have served to disguise developments learners in the experimental condition made with this word form. Verbs constituted the largest group of taught target words, and the fact that there is only a small preference towards the experimental group may also reflect learners' already being familiar with many metaphorical verb senses. Other possibilities are that some experimental group learners avoided using these word forms in metaphorical senses because of a perceived difficulty with producing metaphor, or that the prompts for the writing topics encouraged learners in both conditions to engage with abstract subject matter to an approximately equal degree, and this masked the effects of instrument. It should also be remembered that chapter five found that the experimental treatment had a greater effect on variety than on quantity of metaphorical output, and such effects would not be revealed in this analysis. Phrasal verbs were the least frequent of the open-class parts of speech, so it is perhaps it is to be expected that their usage should show the least signs of change between the conditions. Metaphorical nouns also displayed no sign of a change in output between the two learner groups.

The results shown in Figures 7.3 and 7.4 as well as in Appendix P (page 493) support the view that learner output most clearly differs from that of native speakers with regard to use of metaphorical nouns. Whether this is due to a perceived difficulty, perhaps of phraseology, or the greater challenge of visualising

abstract entities rather than abstract processes (metaphorical verbs showed no signs of avoidance) would be a topic worthy of further study.

7.5. Learner metaphor production and phraseology

Phraseology, as has already been noted, is a vastly complex feature of language, and one that learners only gradually become aware of. More advanced learners, such as the Korean student whose output was the focus of Bell's (2009) study into lexical phrases, show signs of 'chunking chunks', or combining multiple phrases into coherent text. The learners in this study, however, had not yet reached this level of complexity in their writing. Such lexical chunks that could be found consisted of smaller fragments of text, typically collocations, simple colligations and fixed adverbial phrases. Learners' overuse of adverbial phrases was discussed in chapter five, and so the focus of this section will be on collocation and colligation.

7.5.1. Metaphorical adjective / noun collocations

In the experimental condition, learners completed activities in the metaphor workbook on the metaphorical meanings of five pairs of adjectives: high/low, big/small, dark/bright, strong/weak and deep/shallow. All instances of these adjectives appearing in an attributive or predicative relationship within two spaces of a noun in learner output were listed and then categorised as literal or metaphorical. There were no uses of *shallow* in literal or metaphorical senses, but the remaining nine adjectives and their categorisations are given in Table 7.6.

The data in the table give the impression that there was a shift towards greater production of metaphorical adjective/noun collocations in the experimental

condition, mostly due to increased metaphorical use of *high, big*, *dark*, and *strong*. However, Mann-Whitney tests found no significant differences in the ratio of literal adjective/noun collocations to lexical units (Control *Md* = 0.07%, *n* = 23, Experimental *Md* = 0.04%, *n* = 23, *U* = 184.000, *z* = -1.815, *p* = .070, *r* = 0.27) or metaphorical adjective/noun collocations to lexical units (Control *Md* = 0.10%, *n* = 23, Experimental *Md* = 0.19%, *n* = 23, *U* = 181.500, *z* = -1.824, *p* = .068, *r* = 0.27) between the two conditions.

Table 7.6. Adjectives appearing in literal and metaphorical adjective/noun collocations in the two conditions

Adjective	Control		Experimental	
	Literal	Metaphorical	Literal	Metaphorical
High	0	27	1	37
Low	1	8	0	3
Big	12	22	12	47
Small	18	2	7	3
Dark	1	3	0	9
Bright	0	0	0	3
Strong	8	15	5	19
Weak	0	1	0	1
Deep	0	4	0	1
Total (% [a])	**40 (32.79%)**	**82 (67.21%)**	**25 (16.89%)**	**123 (83.11%)**
Error (% [b])	**2 (5.00%)**	**9 (10.98%)**	**2 (8.00%)**	**31 (25.20%)**

a = % of all adjective/noun collocations

b = % of all metaphorical adjective/noun collocations

It is also interesting to note that across the nine adjectives, similarities can be seen between the two conditions both in terms of overall frequency and in the balance of literal and metaphorical usage. The largest increases in metaphor production came from just two adjectives, *big* and *high*, which are by some distance the most frequent of the nine overall in the language. This echoes Ellis and Ferreira-

Junior's (2009) finding that the most frequent members of a construction act as pathbreakers, from which further experimentation typically flows.

Errors in adjective/noun collocations were identified using the NICT JLE tagset (Appendix O, page 491). Generally, error rates appeared to be higher in the experimental condition than the control, and higher for metaphorical rather than literal senses. However, since many of the learners did not produce any metaphorical collocations containing errors (Appendix Q, page 501), it was not possible to analyse the error rates between the two conditions statistically. The distribution of errors across particular tags and broader categories is provided in Table 7.7. The most noticeable difference between the conditions is the higher rate of errors related to phraseology in the experimental condition. Since there is a possibility that the production of recently studied language features will lead to increased error rates, the data here point to the need for a future investigation that might overcome the limitation of the limited sample size in this study.

In order to examine collocation strengths, MI and t-scores were calculated based on COCA for each metaphorical collocation produced by learners[42]. Then, taking significant values as 3 for MI and 2 for t-scores, the percentage of collocations that were significant in each condition could be calculated. Tables 7.8 and 7.9 show that whichever measure of association was applied, and whether the collocations were counted as tokens or types, the control group produced a greater proportion of significant collocations. However, when log-likelihood measures were calculated for

[42] The calculations were based on the number of times the collocate appeared within two spaces either side of the node word.

Table 7.7. Breakdown of error categories by TOEIC ability group for metaphorical adjective/noun collocations

Condition & level		Uses	Total errors (% of total)	Grammatical error (% of total	Lexical error (% of total	Phraseological error (% of total
Control	Low	12	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
	Mid	25	3 (12.00%)	2 (8.00%)	0 (0.00%)	1 (4.00%)
	High	45	6 (13.33%)	2 (4.44%)	4 (8.89%)	0 (0.00%)
Experimental	Low	45	12 (26.67%)	4 (8.89%)	1 (2.22%)	7 (15.56%)
	Mid	34	8 (23.53%)	1 (2.94%)	2 (5.88%)	5 (14.71%)
	High	44	11 (25.00%)	4 (9.09%)	1 (2.27%)	6 (13.64%)
Total	Low	57	12 (21.05%)	4 (7.02%)	1 (1.75%)	7 (12.28%)
	Mid	59	11 (18.64%)	3 (5.08%)	2 (3.39%)	6 (10.17%)
	High	89	17 (19.10%)	6 (6.74%)	5 (5.62%)	6 (6.74%)

the number of significant collocations as a proportion of each learner corpus[43], the results for both MI values (log-likelihood = 0.35, Bayes Factor effect size = -11.17) and t-scores (log-likelihood = -0.04, effect size = -11.48) show no sign of significant overuse in either corpus. Table 7.9 also reinforces the impression that the experimental group's gains in metaphorical adjective/noun collocation usage were more due to increased production of particular metaphorical adjectives than to any great change in the number of types produced; given the respective sizes of the two corpora, the number of types is actually less in the experimental condition in relative terms (control type/token ratio: 0.62, experimental: 0.49).

Table 7.8. Metaphorical adjective/noun collocation tokens and COCA significance measures

Condition	Total tokens	Significant MI (% of total)	Significant t-score (% of total)	Non-significant (% of total)
Control	82	59 (71.95%)	80 (97.56%)	2 (2.44%)
Experimental	123	69 (56.10%)	107 (86.99%)	16 (13.01%)

[43] Control output corpus: 43,490 lexical units; Experimental output corpus: 56,972 lexical units

Table 7.9. Metaphorical adjective/noun collocation types and COCA significance measures

Condition	Total types	Significant MI (% of total)	Significant t-score (% of total)	Non-significant (% of total)
Control	51	33 (64.71%)	49 (96.08%)	2 (3.92%)
Experimental	60	30 (50.00%)	48 (80.00%)	12 (20.00%)

Table 7.10 lists the ten highest and lowest significance values for each measure in both conditions. Some of the highly significant collocations came directly from the course materials; *dark side* and *big hit* both received explicit classroom attention after appearing in texts, and *big game*, *big problem*, *high quality* and *bright future* were taught as part of the metaphor workbook activities.

The right hand side of the table is perhaps the more interesting though, as it highlights some of the issues learners may face when attempting to use adjective/noun collocations. In many cases, such as *strong manner*/*level*, *high technique* and *big*/*small popularity*, it appears that learners were simply using an atypical adjective to complete the collocation. In other cases, the choice of noun also creates problems. In (110), the learner was describing the growth and diversification of a particular sub-culture and probably intended a meaning closer to *phenomenon*, while in (111) *outbursts of emotion* might be more appropriate. Both of these cases were coded as 'Collocation: lexis' errors.

(110) Moreover, Otaku has been becoming <u>bigger category</u>. [CON4]

(111) For instance, people get crazy when Japan wins in a soccer team but there are no <u>big movement</u> in case of Sumo. [CON8]

Table 7.10. Highest and lowest COCA significance measures for metaphorical adjective/noun collocations in each condition

Top 10 by MI				Bottom 10 by MI			
Control		Experimental		Control		Experimental	
big hit	7.02	high level	7.31	high technique	-0.09	strong manner	-1.37
low cost	6.77	bright future	7.18	big incident	0.03	big meaning	-0.68
dark side	6.20	low cost	6.77	strong level	0.62	big culture	-0.63
high quality	5.96	dark side	6.20	low age	1.20	dark life	-0.28
high cost	5.86	strong bond	5.97	big category	1.22	big mobilization	-0.16
strong influence	5.84	high quality	5.96	big movement	1.28	strong oil	-0.13
strong desire	5.82	high cost	5.86	deep culture	1.41	high technique	-0.09
big difference	5.63	big difference	5.63	deep history	1.77	small popularity	0.60
strong opinion	5.51	small group	5.54	dark part	1.90	big popularity	0.83
high rank	5.50	big corporation	5.47	high nutrition	1.98	bright world	1.10

Top 10 by t-score				Bottom 10 by t-score			
Control		Experimental		Control		Experimental	
big problem	72.77	high level	129.76	high technique	-0.33	big culture	-3.10
small number	65.16	small group	78.80	big incident	0.07	strong manner	-2.23
high cost	58.24	big problem	72.77	strong level	2.55	big meaning	-2.01
big difference	54.77	high cost	58.24	deep culture	2.99	dark life	-1.42
low cost	52.11	big difference	54.77	strong curiosity	3.17	strong oil	-0.40
high quality	51.64	low cost	52.11	high nutrition	3.66	high technique	-0.33
big company	48.01	high quality	51.64	big category	4.00	big mobilization	-0.11
dark side	41.68	dark side	41.68	deep history	5.09	small popularity	0.83
big change	40.63	big game	41.46	low age	5.31	strong cigar	1.15
big hit	39.64	big change	40.63	big movement	5.96	big popularity	1.32

As was mentioned in section 7.1, however, it should be remembered that the norms of native speaker usage based on a large corpus might not always be the most appropriate for judging learner output. As evidence of internal factors, the two input corpora were also examined for the same collocations that learners had produced and MI and t-scores calculated as before. However, only 30.77% of collocation types in the control and 27.12% of types in the experimental condition were present in

input as measured by the corpora. Of the low-ranked collocations listed in Table 7.10, only *big movement* in the control and *big popularity* in the experimental group appeared in input, and in both cases just a single time each. It appears then, that the majority of these weakly-related collocations were not influenced by classroom input and may instead have been the learners' own creations, with the possibility of influence from the L1.

Interestingly, however, the proportion of collocation tokens that were significant under either measure was higher in the experimental condition than the control when the collocation measures were determined based on the input corpora (Table 7.11). Log-likelihood tests found a significant difference with a low effect size for the proportion of collocations that were significant under the t-score measure but no significant differences for the proportions of collocations that were significant under the MI measure or that were non-significant under either measure. Thus it appears that a large number of learners' metaphorical adjective/noun collocations were not reflecting classroom input; they may have been based on prior knowledge or created ad hoc to meet an immediate communicative need. Many of the weakly related collocations from Table 7.10 belong in this group. However, there is a possibility that the treatment encouraged learners to use more metaphorical collocations that they had heard in class, particularly those expressions which appeared in explicit instruction, as was the case for *high quality*, *dark side*, and *big game*/*problem*/*difference*.

The results in Tables 7.9 and 7.11 indicate that the choice of comparison corpus has the potential to influence analyses of language production. Clearly, the limited number of participants in this study and narrow focus on a single course restricts the

conclusions that can be drawn, but again this may prove to be a worthwhile area for a future investigation.

Table 7.11. Log-likelihood tests for the relative frequency of metaphorical adjective/noun collocations grouped by input corpus significance measure

Condition	Total tokens	Significant MI (% of total)	Significant t-score (% of total)	Non-significant (% of total)
Control	82	25 (30.49%)	8 (9.76%)	56 (68.29%)
Experimental	123	49 (39.84%)	34 (27.64%)	66 (53.66%)
Log-likelihood tests based on totals as proportions of output corpora [a]				
Log-likelihood [b]		-2.90	-11.23***	0.29
Bayes factor effect size		-8.62	-0.30	-11.23

a = Control output corpus size = 43,830, Experimental output corpus size = 56,972

b = Negative values indicate underuse in control corpus relative to experimental corpus (positive values overuse in control corpus)

*** = Significant at the 0.001 level

Overall, it should be noted that the vast majority of metaphorical collocations were testified in COCA data. Just 2.44% and 13.01% of control and experimental group metaphorical collocations respectively were non-significant with either measure. If a criticism could be made, it is that learners were usually very dependent on these high-frequency adjectives to form collocations. There was a *big incident*, but no *serious* or *major incidents*, and several uses of *high popularity*, but no *growing*, *great* or *enormous popularity*. This would be a harsh criticism to make, however, this finding merely suggests that the learners were at a point in their development where they could often apply basic adjectives in collocations to express abstract concepts, but had not yet reached the point of being able to add variety to their output.

7.5.2. Metaphorical verb / noun collocations

Culture was by far the most frequent noun produced by learners, and in metaphorical collocations, it appeared as the subject or object of 71 different metaphorical verbs across the two corpora. Table 7.12 shows the total number of metaphorical verb + *culture* collocations in each condition, the number that reached significance under the MI and t-score measures in COCA, and the number that did not reach significance under either measure. As with the analysis of adjective/noun collocations, the proportion of collocations that were significant under either measure was slightly lower for experimental group learners. However, log-likelihood tests showed that the production of collocations with significant MI or t-scores was actually very evenly balanced between the two conditions.

Table 7.12. Log-likelihood tests for the relative frequency of metaphorical verb + *culture* collocations grouped by COCA significance measures

Condition	Total (% of all lexical units)	Significant MI (% of total)	Significant t-score (% of total)	Non-significant (% of total)
Control	118	41 (34.75%)	73 (61.86%)	45 (38.14%)
Experimental	187	52 (27.81%)	98 (52.41%)	89 (47.59%)
Log-likelihood tests based on totals as proportions of output corpora [a]				
Log-likelihood [b]	-2.88	0.01	-0.04	-5.48*
Bayes factor effect size	-8.64	-11.51	-11.48	-6.04

a = Control output corpus size = 43,830, Experimental output corpus size = 56,972

b = Negative values indicate underuse in control corpus relative to experimental corpus (positive values overuse in control corpus)

* = Significant at the 0.05 level

If we turn to the more localised context of the classroom, Table 7.13 reveals that, as with the adjective/noun collocations, when the input corpora are used for comparison, the experimental group produced higher proportions of collocations

that were significant under either the MI or t-score measures. Log-likelihood tests found that the experimental group learners produced significantly more collocations that were significant under the MI and t-score measures, although the effect sizes were negligible, while there was no difference in the proportion of collocations that did not reach either significance measure. It should be noted that 40.48% of learner-produced collocation types were present in the input corpus for the control group, and 53.45% in the experimental condition. Considering tokens, 50.00% of control group collocations were present in input, while 74.33% of experimental group collocations also appeared in class. This implies that the classroom may have been an important source for learners' collocation choices in both conditions, particularly the experimental group. Although it is highly probable that learners had already known some of this language before, their output was nevertheless reflecting the language of the classroom.

Table 7.13. Log-likelihood tests for the relative frequency of metaphorical verb + *culture* collocations grouped by input corpora significance measures

	Total	Significant MI (% of total)	Significant t-score (% of total)	Non-significant (% of total)
Control	118	36 (30.51%)	11 (9.32%)	80 (67.80%)
Experimental	187	92 (49.20%)	40 (21.39%)	93 (49.73%)
Log-likelihood tests based on totals as proportions of output corpora [a]				
Log-likelihood [b]		-12.86**	-10.79**	0.53
Bayes factor effect size		1.34	-0.73	-10.99

a = Control output corpus size = 43,830, Experimental output corpus size = 56,972

b = Negative values indicate underuse in control corpus relative to experimental corpus (positive values overuse in control corpus)

* = Significant at the 0.05 level; ** = 0.01 level

In Table 7.14, the ten highest and lowest collocations for each measure when judged against COCA are listed. It can be seen that all of the collocations with the highest MI scores in both conditions were significant by this measure (≥ 3), and the same is true for the collocations with high t-scores, for which a value of 2 or higher is significant. Many of these were target metaphors in the experimental condition. Looking at the verb collocates with low scores, it seems that almost all appear in the lists for both measures. However, it would be inaccurate to say that these are incorrect choices on the learners' part as collocations; several (e.g., *take*, *have*, *make*, and *come*) do appear as collocates of *culture* dozens or even hundreds of times in COCA, but not enough to score highly with either measure of collocation due to their extremely high overall frequency in language. In other cases, the specific nature of the course needs to be taken into account. *Come into a culture* was a taught target expression and this collocation reached a significant MI measure in both input corpora. Likewise, the verb *add* was treated as an untaught target metaphor, and its collocation with *culture* was significant by both measures in the control input corpus. Other low-scoring verbs on the list do appear marked (112 - 115), at least to this writer, and are extremely rare but attested in COCA (116 - 119).

(112) However, I understood that much life did not have a change after Western culture spread out in the Meiji era. [CON12]

(113) [Japanese people] try to mix foreign culture to bring up their culture. [EXP1]

(114) This is why, Japanese cultures are rising in this country. [EXP20]

(115) ...I think that Japanese culture could grow up by accepted foreign culture. [EXP20]

(116) At the end of history, all of past culture is spread out and equally available to contemporary artists... [COCA]

(117) ...we have to be very careful when we seek to make positive statements in a culture that has been brought up to see Latinos otherwise. [COCA]

(118) ... is this the best that the culture can rise itself to? [COCA]

(119) There is a culture that has grown up over a long period of time... [COCA]

Table 7.14. Highest and lowest COCA significance measures for metaphorical verb + *culture* collocations in each condition

Top 10 by MI				Bottom 10 by MI			
Control		Experimental		Control		Experimental	
shape	6.16	shape	6.16	come into	-1.49	realize	-2.71
absorb	4.53	preserve	5.98	show	-1.24	feel	-1.71
export	4.28	revitalize	5.22	add	-0.99	come into	-1.49
interact	4.25	evolve	4.85	give	-0.84	show	-1.24
expose	4.04	absorb	4.53	take	-0.83	bring up	-1.08
derive	3.97	erase	4.12	have	-0.51	add	-0.99
emerge	3.89	expose	4.04	make	-0.44	give	-0.84
adopt	3.86	share	3.94	come	-0.34	take	-0.83
adapt	3.79	retain	3.91	grow up	0.55	rise	-0.70
relate	3.75	adopt	3.86	spread out	0.57	have	-0.51

Top 10 by t-score				Bottom 10 by t-score			
Control		Experimental		Control		Experimental	
shape	14.32	shape	14.32	have	-11.40	have	-11.40
produce	11.53	preserve	13.92	take	-6.73	feel	-9.62
relate	9.97	share	12.75	show	-5.95	realize	-7.85
bring	9.54	produce	11.53	give	-5.27	take	-6.73
emerge	8.95	relate	9.97	make	-4.07	show	-5.95
reflect	8.18	reflect	8.18	add	-3.55	give	-5.27
adopt	7.74	evolve	8.13	come	-2.73	make	-4.07
connect	7.69	adopt	7.74	come into	-2.55	add	-3.55
develop	7.56	fit	7.72	spread out	0.33	come	-2.73
introduce	7.51	connect	7.69	grow up	0.95	come into	-2.55

As shown in Table 7.13, a considerable proportion of the metaphorical verb collocates of *culture* produced by learners were non-significant by either collocation measure when judged against COCA. However, the lists in Table 7.14 suggest that many of these might be explained as a dependence on high frequency verbs or the influence of classroom input. In fact, only four verbs were entirely uncorroborated by COCA data and as such did not have MI or t-scores. Of these, the first three are comprehensible if some allowance is made for grammar (120 - 122), but might be considered phraseological errors, since there is no evidence for these collocations in larger corpora. The remaining case (a WIDLII coding) (123) has been retained in the data (rather than coded as DFMA) since there is some similarity to the second sense in Macmillan: 'to improve a skill by practising.'

(120) However, they use spoon or fork or knife in modern, so Japanese people put in other culture from other countries. [EXP21]

(121) However, demerits are Japanese traditional cultures may be deleted, because foreign culture is imported a lot. [EXP14]

(122) Makiko got Western culture positively, but on the other hands, Mr. Moriguchi didn't feel good for Western culture. [EXP5]

(123) This is why I think that Japanese culture, for example ceremony of Japan or annual event, are polishing toward future... [EXP20]

Errors were classified with the NICT JLE tagset, and the results were quite similar between the two conditions (Table 7.15). The overall rate of error declined with ability in both groups, and the error rate in general appeared higher than that for the metaphorical adjective collocations. This might not be a surprise, since use of verbs

requires consideration of tense, aspect, voice, and inflection, so there are more possibilities for error. A Mann-Whitney test found no significant difference in the amount of errors as a proportion of all lexical units between the control (*Md.* = 0.15%, *n* = 21[44]) and experimental (*Md.* = 0.09%, *n* = 22) groups (*U* = 219.000, *z* = -0.293, *p* = .770, *r* = 0.04). Kruskal-Wallis tests also found no significant differences in the error rates of the ability groups in the control (Low: *Mean rank* = 15.29, *n* = 7; Mid: *Mean rank* = 10.07, *n* = 7; High: *Mean rank* = 7.64, *n* = 7; χ^2 (2, n = 21) = 5.607, *p* = .061) or experimental conditions (Low: *Mean rank* = 13.21, *n* = 7; Mid: *Mean rank* = 12.57, *n* = 7; High: *Mean rank* = 9.06, *n* = 8; χ^2 (2, n = 22) = 1.816, *p* = .403). When the results of the two groups were combined, however, a Kruskal-Wallis test revealed a significant difference in the error rate between ability groups (Low: *Mean*

Table 7.15. Breakdown of error categories by TOEIC ability group for metaphorical verb + *culture* collocations

Condition & level		Total uses	Total errors (% of total)	Grammatical error (% of total)	Lexical error (% of total)	Phraseological error (% of total)
Control	Low	33	19 (57.58%)	11 (33.33%)	2 (6.06%)	6 (18.18%)
	Mid	47	17 (36.17%)	13 (27.66%)	0 (0.00%)	4 (8.51%)
	High	38	12 (31.58%)	8 (21.05%)	1 (2.63%)	3 (7.89%)
Experimental	Low	50	26 (52.00%)	11 (22.00%)	1 (2.00%)	14 (28.00%)
	Mid	70	26 (37.14%)	16 (22.86%)	3 (4.29%)	7 (10.00%)
	High	67	20 (29.85%)	8 (11.94%)	3 (4.48%)	9 (13.43%)
Total	Low	83	45 (54.22%)	22 (26.51%)	3 (3.61%)	20 (24.10%)
	Mid	117	43 (36.75%)	29 (24.79%)	3 (2.56%)	11 (9.40%)
	High	105	32 (30.48%)	16 (15.24%)	4 (3.81%)	12 (11.43%)

[44] Two participants in the control group and one in the experimental group did not produce any of the metaphorical verbs in their written output.

rank = 28.25, *n* = 14; Mid: *Mean rank* = 22.07, *n* = 14; High: *Mean rank* = 16.10, *n* = 15; χ^2 (2, n = 43) = 6.837, *p* = .033). Post hoc Mann-Whitney tests revealed that when the two sets of participants were combined, high ability learners (*Md.* = 0.04%) produced significantly less errors in metaphorical verb + *culture* patterns than low level learners (*Md.* = 0.16%) (*U* = 50.500, *z* = -2.395, *p* = .017, *r* = 0.44).

7.5.3. Colligations with common verbs

The six most frequent taught target metaphorical verbs over both conditions were *evolve*, *take in*, *absorb*, *mix*, *spread* and *lose*. Each usage of these words was classified as a passive, transitive or intransitive construction, and the distribution of these groupings was compared with that of the same words in COCA.

The distribution of constructions for each metaphorical verb is shown in Figures 7.6 - 7.11. On the whole, it appears that the balance of colligation patterns is similar across the three conditions. In most cases, the three constructions are ranked in the same order of frequency for output in the three corpora. The exceptions to this are *mix*, for which both transitive and intransitive usage are relatively frequent, and *evolve*, which attracted several uses of the passive construction in the control condition, although this pattern is actually very rare in COCA. Apart from these two verbs, the two learner groups make more use of the most frequent pattern for each verb than is the case in COCA. A dependence on the prototype form is in accordance with models of language development that predict learners will initially make use of only the most frequent form of a language feature before later diversifying their range of production as proficiency increases (Ellis & Ferreira-Junior, 2009). Most words have multiple usage patterns, but language learners take time to gain an

understanding of these possibilities and to become aware of the need to move out of their comfort zones.

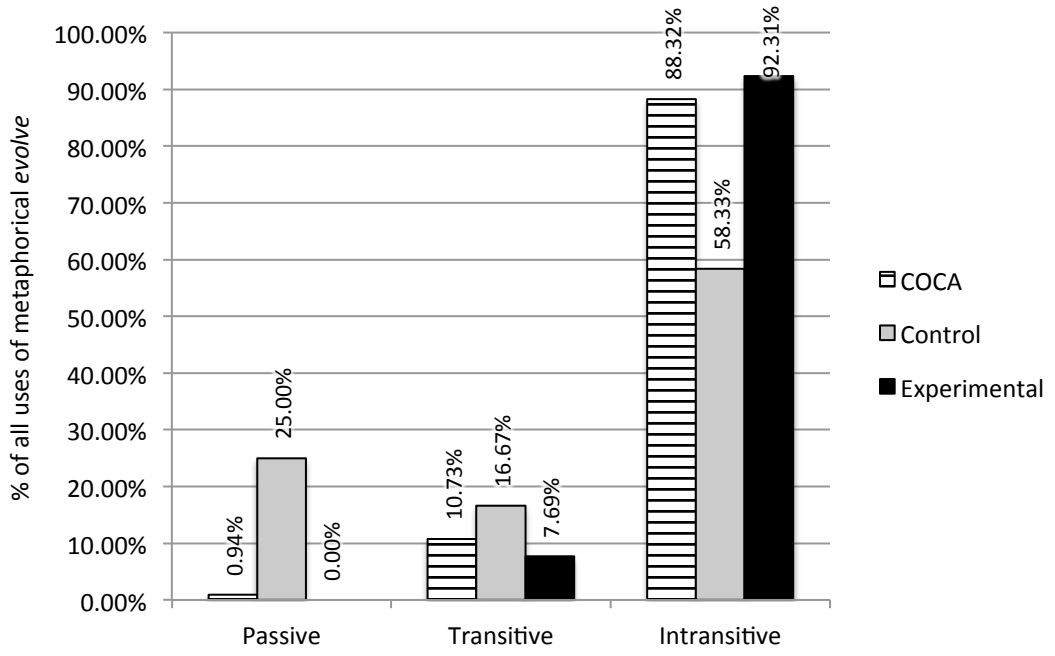


Figure 7.6. Colligation patterns for metaphorical *evolve* in the output corpora and all instances of *evolve* in COCA

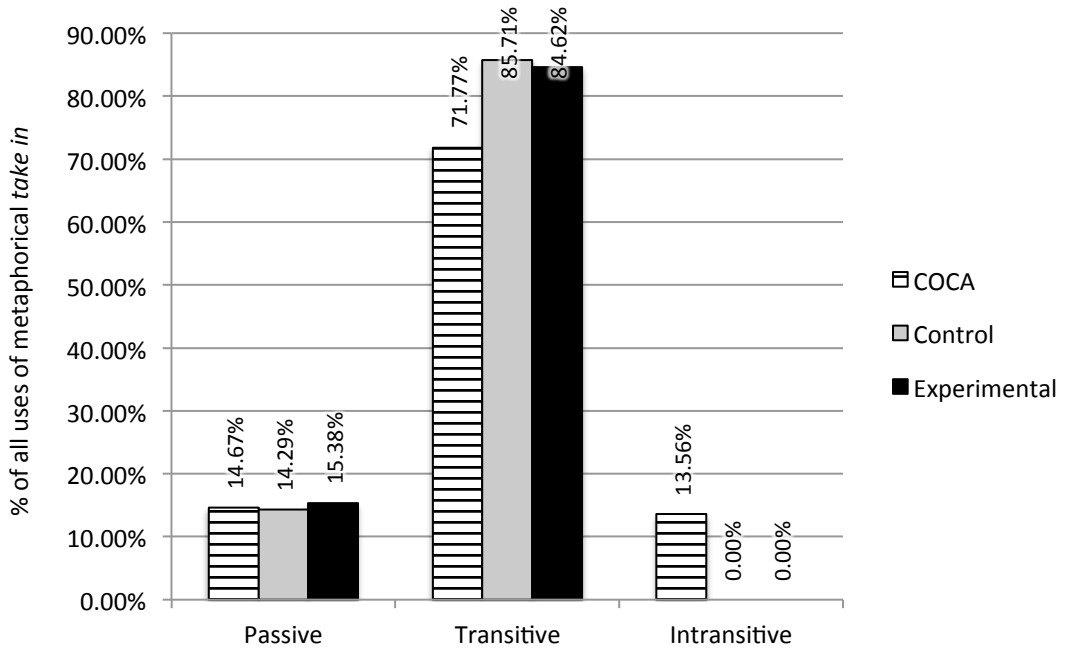


Figure 7.7. Colligation patterns for metaphorical *take in* in the output corpora and all instances of *take in* in COCA

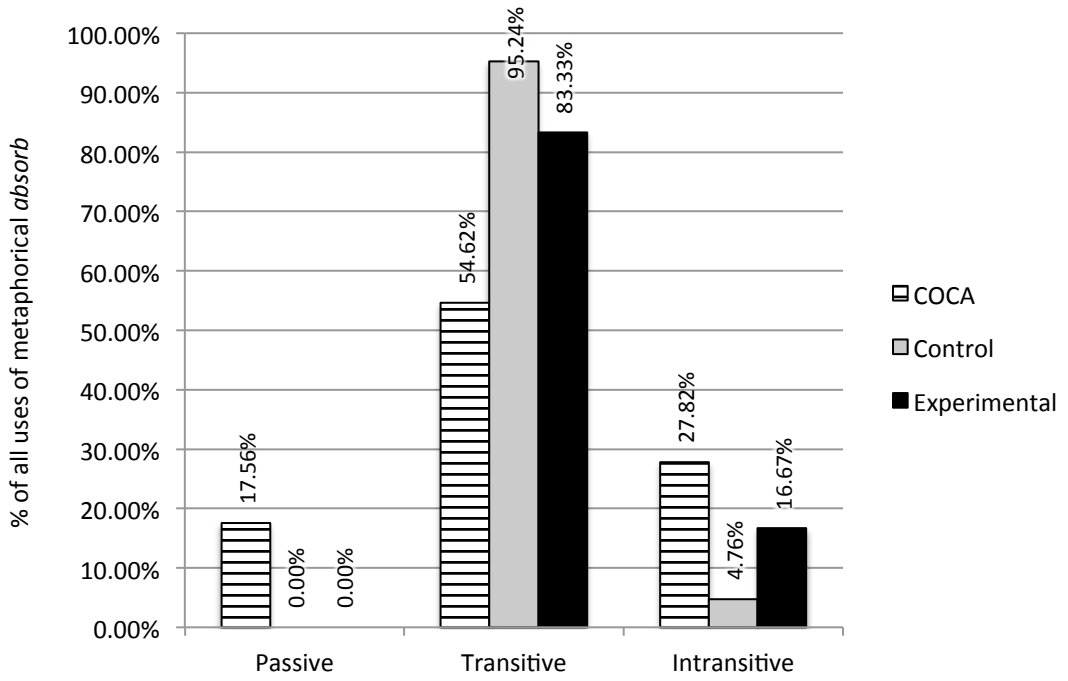


Figure 7.8. Colligation patterns for metaphorical *absorb* in the output corpora and all instances of *absorb* in COCA

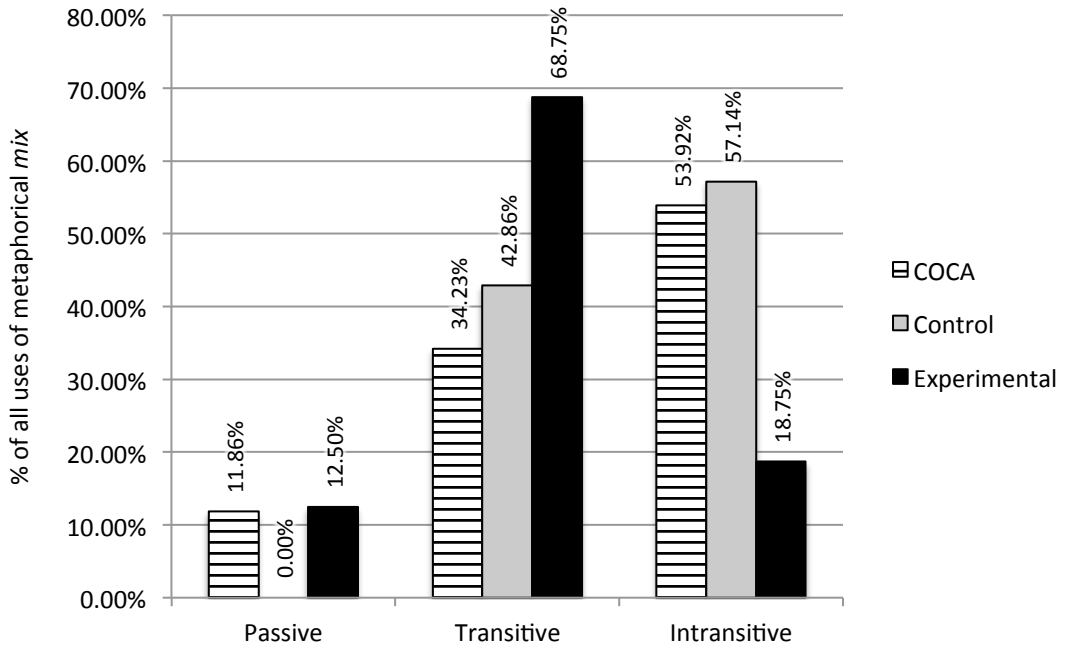


Figure 7.9. Colligation patterns for metaphorical *mix* in the output corpora and all instances of *mix* in COCA

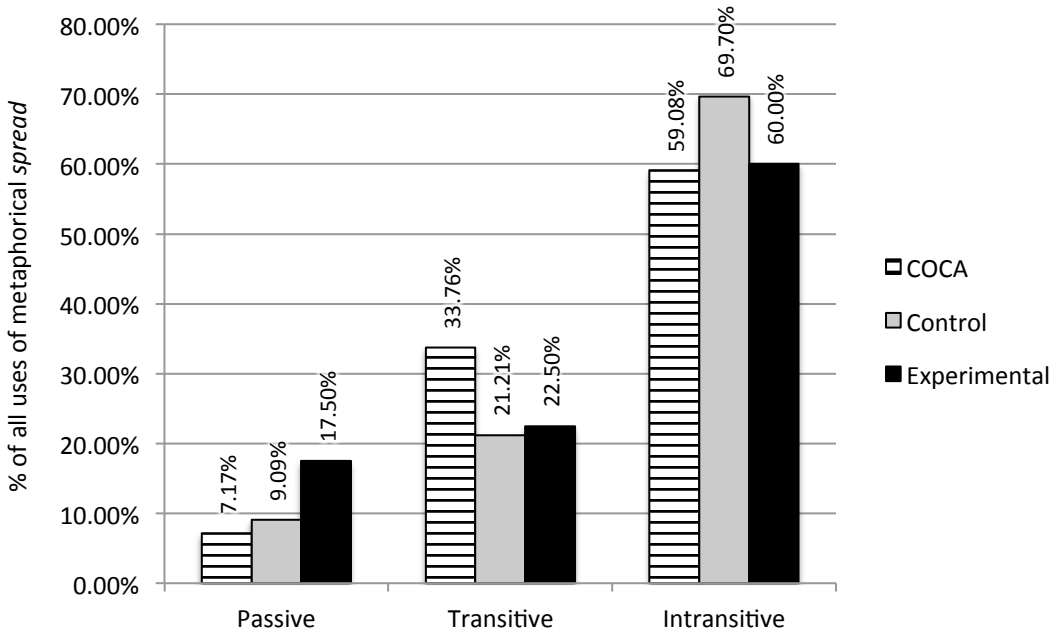


Figure 7.10. Colligation patterns for metaphorical *spread* in the output corpora and all instances of *spread* COCA

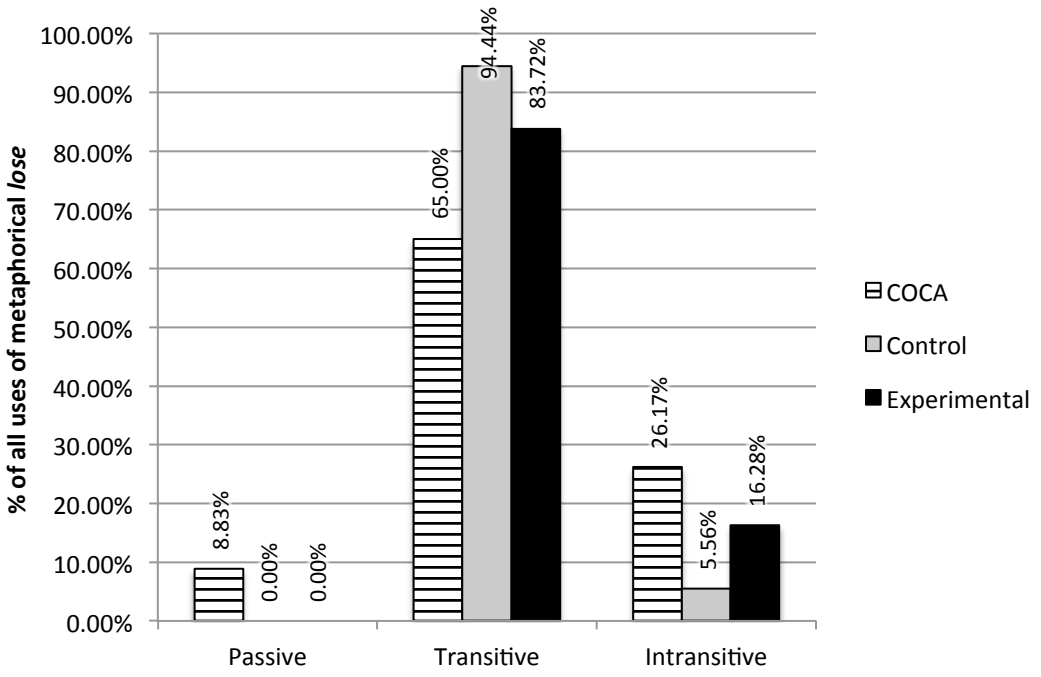


Figure 7.11. Colligation patterns for metaphorical *lose* in the output corpora and all instances of *lose* in COCA

All instances of the six metaphorical verbs appearing in learner output were also coded using the NICT JLE tag list (section 7.2.3.2). It was noticeable that many of the errors made by learners were in the less frequent patterns for each verb, often

because another pattern would have been the more suitable choice. This was the case for *evolve* (Figure 7.6), where the control group learners used a passive form of the verb in 25% of cases. This is actually a rare construction for *evolve* to appear in, and an examination of the learners' output suggests that these individuals did not yet have full understanding of the typical usage patterns for this word (124 - 126). Similar cases can be seen in (127), where the lack of a preposition (*into*) after the verb creates an unintended transitive pattern, and (128), in which either a passive construction or a transitive use with an object such as *popularity* would have been more appropriate.

(124) Japanese wedding has been evolved, and adapted many other wedding style. [CON6]

(125) There are two elements which wedding can be evolved from now on. [CON6]

(126) Japanese wedding in the future will be more evolve. [CON22]

(127) And then, food culture of the Udon entered and spread the life of the common people over the Edo period. [EXP10]

(128) Furthermore, Japanese things such as geta, kanzashi, Japanese sword and Japanese umbrella are losing now. [EXP23]

A log-likelihood test was applied to the combined uses of the six metaphorical verbs in both conditions, and it was found that the less frequent patterns were significantly more likely to contain errors than the most frequent pattern of use for each verb (as determined by COCA usage) (Table 7.16 and Appendix Q, page 501).

Table 7.16. Log-likelihood test for errors in most frequent and less-frequent colligation patterns for common metaphorical verbs

	Total instances	Errors (% of total)	Log-likelihood [a]	Bayes Factor effect size
Most frequent patterns	174	63 (36.21%)	-7.73**	2.26[+]
Less-frequent patterns	62	40 (64.52%)		

Note: Log-likelihood value calculated from the number of errors as a proportion of the number of instances.
a = Negative value indicates a lower proportion of error among most frequent patterns
** = Significant at the 0.01 level
+ = Positive evidence against the null hypothesis

All instances of error in the production of the six verbs were classified using the set of tags provided in Appendix O (page 491) and then grouped into the three categories outlined in section 7.2.3.2. The results, presented in Table 7.17, appear to show fluctuating error rates across the categories of error. Log-likelihood tests were used to compare the error rates between the two conditions and found no significant differences in any category of error.

Table 7.17. Log-likelihood tests for errors in usage of six common metaphorical verbs

Condition	Total uses	Total errors (% of total)	Grammatical error (% of total)	Lexical error (% of total)	Phraseological error (% of total)
Control	105	49 (46.67%)	32 (30.48%)	4 (3.81%)	13 (13.38%)
Experimental	131	54 (41.22%)	24 (18.32%)	11 (8.40%)	19 (14.50%)
Log-likelihood [a]		0.39	3.60	-3.26	-0.20
Bayes factor effect size		-5.07	-1.86	-2.20	-5.27

a = Negative values indicate a lower proportion of error in the control corpus

The errors in production of the six common metaphorical verbs were also then broken down by ability level (Table 7.18). The general impression of the error rates for the sub-categories of error suggest that there was a reduction in error with increased ability in the experimental condition, but that there was no clear trend for the control group. The low error counts for some categories prevented a statistical

analysis of error at the sub-category level, but it was possible to analyse the total error rate over the ability groups in each condition. An independent samples Kruskal-Wallis test found no significant difference in the total error rate across the ability groups in the control condition (Low: *n* = 8, *Mean rank* = 10.81, Mid: *n* = 7, *Mean rank* = 11.14, High: *n* = 6, *Mean rank* = 11.08; χ^2 (2, n = 21) = 0.012, *p* = .994). For the experimental group, a Kruskal-Wallis test revealed a significant difference in the total error rate across ability groups (Low: *n* = 8, *Mean rank* = 15.50, Mid: *n* = 7, *Mean rank* = 14.21, High: *n* = 8, *Mean rank* = 6.56; χ^2 (2, n = 23) = 8.143, *p* = .017). Post-hoc Mann-Whitney tests between the experimental condition ability groups found no significant difference between the low (*Md.* = 65.00%) and mid (*Md.* = 50.00%) groups (*U* = 21.500, *z* = -0.758, *p* = .448, *r* = 0.20), but significant differences were found between the low (*Md.* = 65.00%) and high (*Md.* = 12.70%) groups (*U* = 10.500, *z* = -2.278, *p* = .023, *r* = 0.57) and mid (*Md.* = 50.00%) and high (*Md.* = 12.70%) groups (*U* = 6.000, *z* = -2.574, *p* = .010, *r* = 0.66).

It is interesting that error rates appear to behave differently across the two conditions, although when the results across both conditions are tallied, a more predictable pattern of increased accuracy with ability emerges. Gathering more data might well allow for more definitive statements to be made, but based on the present evidence, it appears that low level learners are those most likely to make errors in the production of these metaphorical verbs, and that this is most evident with grammatical errors. However, the chosen method of coding errors allots more tags to the grammar category than lexis or phraseology, so the higher rates of grammatical error may simply reflect the fact that there were more ways of making errors, rather than anything meaningful about learner development.

Table 7.18. Breakdown of error categories by TOEIC ability group for colligation patterns with metaphorical verbs

Condition & level		Total uses	Total errors (% of total)	Grammatical error (% of total)	Lexical error (% of total)	Phraseological error (% of total)
Control	Low	28	16 (57.14%)	11 (39.29%)	0 (0.00%)	5 (17.86%)
	Mid	53	21 (39.62%)	14 (26.42%)	2 (3.77%)	5 (9.43%)
	High	24	12 (50.00%)	7 (29.17%)	2 (8.33%)	3 (12.50%)
Experimental	Low	48	33 (68.75%)	17 (35.42%)	7 (14.58%)	9 (18.75%)
	Mid	37	16 (43.24%)	6 (16.22%)	3 (8.11%)	7 (18.92%)
	High	46	5 (10.87%)	1 (2.17%)	1 (2.17%)	3 (6.52%)
Combined	Low	76	49 (64.47%)	28 (36.84%)	7 (9.21%)	14 (18.42%)
	Mid	90	37 (41.11%)	20 (22.22%)	5 (5.56%)	12 (13.33%)
	High	70	17 (24.29%)	8 (11.43%)	3 (4.29%)	6 (8.57%)

7.5.4. Colligations with common nouns

Other than *culture*, few individual nouns appeared frequently across both conditions. However, the set of nouns including *part*, *element*, *component* and *aspect* (hereafter referred to as [part] nouns) was used by both learner groups (although more in the control than the experimental condition) and was dispersed across a range of topics. Instances of these words being used were classified according to whether they contained a pre-modifying adjective, acted as the head of a prepositional phrase, or were acting as the subject or object of a verb other than *be*. The rate of usage for these patterns was again compared with data from COCA. Unlike the analysis of verbal colligation patterns, these categories were not mutually exclusive, and learners often combined them, as in examples (129 - 130). For this reason, the percentages reported below frequently sum to more than 100%. Also, the words *component* and *aspect* were not as frequently used as the other members of the

group, which led to difficulties in obtaining a sense of their usage patterns. Therefore, the results will discuss *element* and *part* separately and then the group of [part] nouns as a whole.

(129) Sumo is related to other aspects of Japanese culture. [EXP17]

(130) Sumo is national sports and it has existed for hundreds years. In the long history, it has changed in several point, but still maintains its essential element. [CON1]

The distributions of colligation patterns for these words are shown in Figures 7.12 - 7.14. As with verbal colligations, we see the prototypical construction (the head of a prepositional phrase) is most frequent across all three corpora. Compared with verbal colligations, learners also appeared to be making frequent use of the secondary constructions. Log-likelihood calculations between the COCA patterns and individual learner corpora for all [part] nouns found that both learner groups were significantly more likely to use pre-modifying adjectives in this construction, but the effect sizes were negligible. When the learner data was combined for comparison with COCA, however, a positive effect size was present. No significant differences were found for the other patterns (Table 7.19).

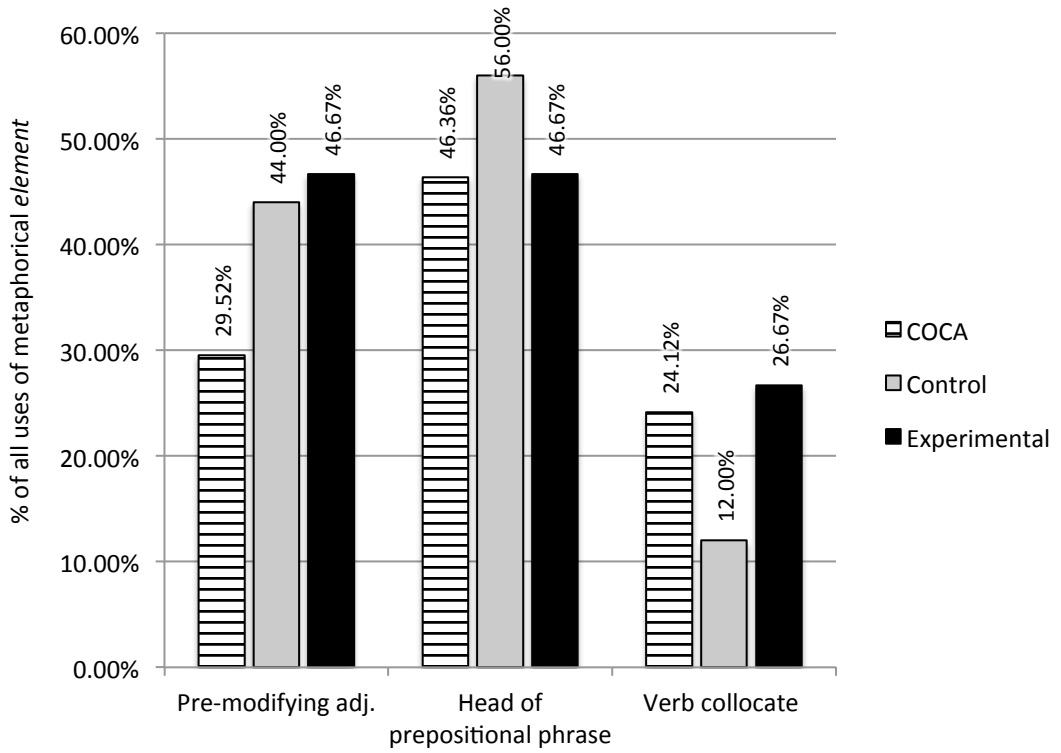


Figure 7.12. Colligation patterns for metaphorical *element* in the output corpora and all instances of *element* in COCA

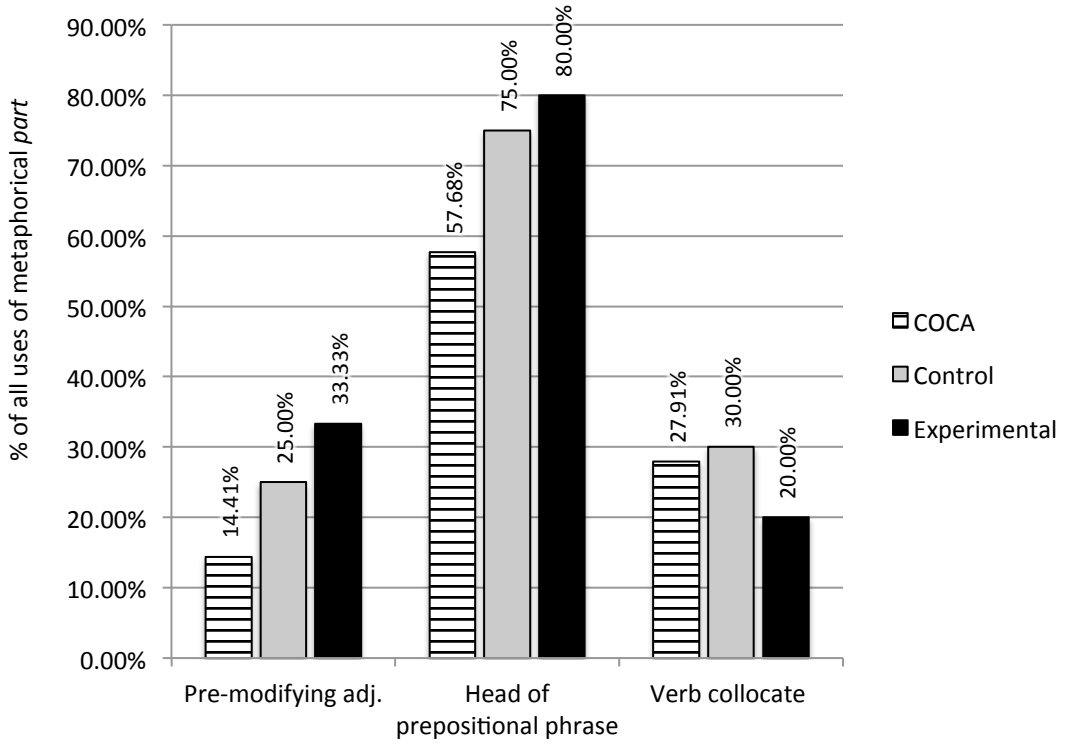


Figure 7.13. Colligation patterns for metaphorical *part* in the output corpora and all instances of *part* in COCA

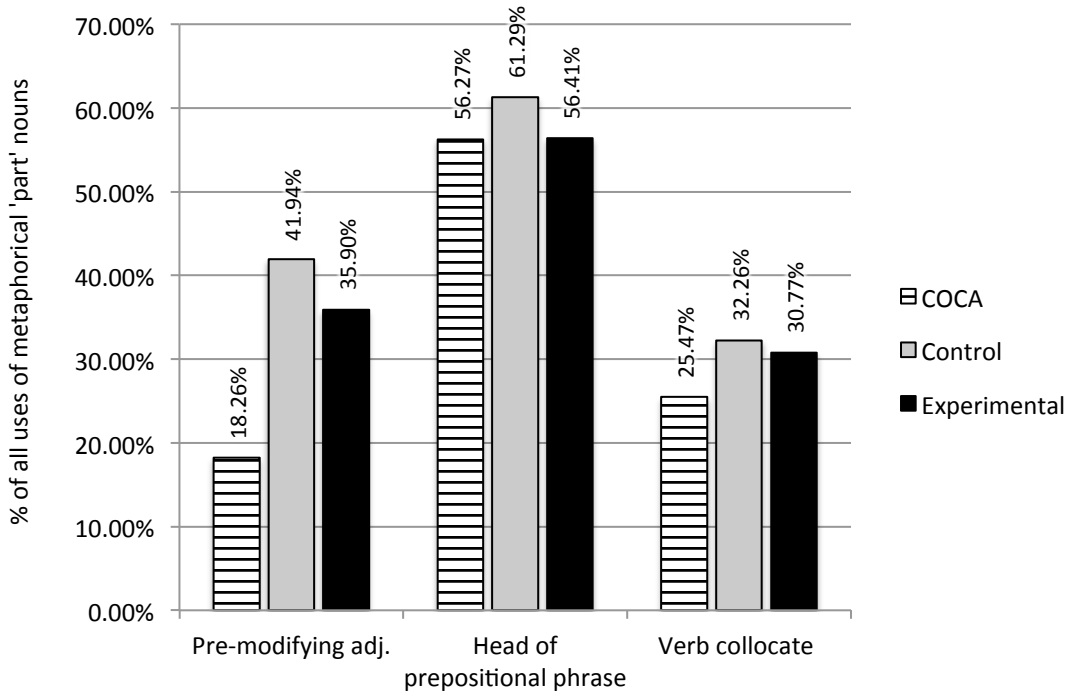


Figure 7.14. Colligation patterns for all metaphorical [part] nouns in the output corpora and all instances of these nouns in COCA

Table 7.19. Log-likelihood tests for relative frequency of colligation patterns for [part] nouns between COCA and learner output corpora

	Pre-modifying adjective		Head of prepositional phrase		With verb collocate	
	Log-likelihood [a]	Bayes factor effect size	Log-likelihood [a]	Bayes factor effect size	Log-likelihood [a]	Bayes factor effect size
Control	-13.85***	0.75	-0.29	-12.81	-1.03	-12.08
Experimental	-5.16*	-7.94	0.00	-13.11	-0.40	-12.71
Combined learners	-18.81****	5.71[+]	-0.14	-12.96	-1.43	-11.68

Note: Log-likelihood values calculated from the number of uses of each pattern as a proportion of the number of instances of all patterns (see Appendix Q, page 501, for raw counts).
a = Negative values indicate underuse in COCA relative to learner corpora
* = Significant at the 0.05 level; *** = 0.001 level; **** = 0.0001 level
+ = Positive evidence against the null hypothesis; Strong evidence

Learner errors were again classed with the NICT JLE coding scheme and error groups tabulated with learner proficiency (Table 7.20). With these data, there was less of a clear trend for error to decrease with proficiency. While high-level learners produced the lowest levels of error in both conditions, it was those in the mid-level group who made the most errors. Unfortunately, owing to the low overall

production of this language feature, it was not possible to analyse the rates of error between the two conditions. Instead, the number of errors with metaphorical [part] nouns as a proportion of all lexical units was calculated for all participants, and a Kruskal-Wallis test was run across all participants in both conditions. No significant difference was found in the error rate across the ability groups (Low: *n* = 8, *Mean rank* = 18.13, Mid: *n* = 12, *Mean rank* = 19.75, High: *n* = 13, *Mean rank* = 13.77; χ^2 (2, n = 33) = 2.627, *p* = .269). It was also noticeable that grammatical error as a whole was less of a factor than had been the case with verbal colligation patterns, and lexical errors were more prominent. As was mentioned in the previous section, this may have been a factor of the coding scheme itself, with less grammatical error tags available than for verbs. However, examples (131 - 134) show that learners from a range of proficiencies made incorrect lexical selections using these words.

Table 7.20. Breakdown of error categories by TOEIC ability group for colligation patterns with metaphorical [part] nouns

Condition & group		Total uses	Total errors (% of total)	Grammatical errors (% of total)	Lexical errors (% of total)	Phraseological errors (% of total)
Control	Low	14	6 (42.86%)	1 (7.14%)	1 (7.14%)	4 (28.57%)
	Mid	23	10 (43.48%)	1 (4.35%)	7 (30.43%)	2 (8.70%)
	High	25	5 (20.00%)	2 (8.00%)	0 (0.00%)	3 (12.00%)
Experimental	Low	7	3 (42.86%)	0 (0.00%)	2 (28.57%)	1 (14.29%)
	Mid	14	10 (71.43%)	4 (28.57%)	4 (28.57%)	2 (14.29%)
	High	18	6 (33.33%)	2 (11.11%)	1 (5.56%)	3 (16.67%)
Combined	Low	21	9 (42.86%)	1 (4.76%)	3 (14.29%)	5 (23.81%)
	Mid	37	20 (54.05%)	5 (13.51%)	11 (29.73%)	4 (10.81%)
	High	43	11 (25.58%)	4 (9.30%)	1 (2.33%)	6 (13.95%)

(131) *I think it is very good that having a hobby and devoted to a favorite things. So, I think "otaku" is <u>part</u> of hobby.* [CON19] [Low]

(132) *Not all Japanese musicians can speak English, so it is precious that part of them sing a song in English or give an interview in English...* [EXP20] [Low]

(133) *Japanese wedding is unusual elements in the world.* [EXP18] [Mid]

(134) *Because of the difference of language, [Japanese singers] can't express what they really want to say in the component.* [EXP11] [High]

The results in this section suggest that learners' language performance may vary depending on the particular language feature at hand as well as with proficiency. Error rates did not always decline among more proficient learners; for collocations with metaphorical adjectives, there was little difference in error rate across the ability groups, while the error rate for colligation patterns with nouns fluctuated over the ability levels. The two cases where low ability learners produced significantly more errors than higher ability students both involved the use of verbs, and it is possible that a relative lack of control in the use of tense, aspect, voice, and subject-verb agreement is influencing this result. By comparison, lexical choice or phraseological issues appeared to cause greater problems for learners with the [part] noun construction. It was also notable that few low level learners (50.00%) attempted to produce this structure across both conditions, compared with 85.71% of mid-level learners and 81.25% of the higher ability participants, so it is also possible that avoidance is playing a part, or that these learners are not ready to incorporate structures based around abstract nouns into their stock of productive language. It appears that analyses of learner language use at different stages of development can be expected to reveal considerable flux in levels of both output and error.

7.6. L1 influence on metaphor production

7.6.1. Sources of metaphors

The first approach to investigating the influence of the L1 was through the feedback the learners themselves provided on the sources of metaphors that they produced. As was stated in section 3.6.4 (page 119), the words selected for metaphor source forms were those produced by learners that were infrequent, potentially influenced by the L1, included in classroom input, or unconventional in some way. In the control condition, 157 sources of metaphor production were provided by the learners from 111 forms, while the experimental group provided 197 sources from 142 forms. The distribution of metaphor sources in each condition can be seen in Table 7.21. Between the two conditions, the rate at which learners drew on the various sources remained broadly similar with one clear exception. This was the increase in learners' use of words that had come from course input. A log-likelihood test found this increase to be significant, with a large effect size. However, it should be remembered that presence in course input was one of the criteria for selecting words for the feedback forms. There is also the possibility that learners in the experimental condition believed that this was the most desirable option to select, since it would indicate that they could recall prior course content. At the same time, it is likely that the experimental treatment made metaphors in input more salient for learners. In only 14.89% of cases (seven of 47) in which the language had been selected by the researcher because it had appeared in input did control condition learners actually indicate they that had used words from class. The corresponding figure for experimental group learners was 50.57% (44 of 87 cases). This raises the

possibility that learners may not always be aware of the influence of classroom input unless it is made salient in some fashion.

Table 7.21. Feedback form results for sources of learner metaphor in the two conditions

Source	Condition	Uses	Rate of use (% of forms)	Log-likelihood [a]	Bayes factor effect size
L2 dictionary	Control	15	13.51%	2.00	-3.54
	Experimental	11	7.75%		
L1 dictionary	Control	39	35.14%	0.45	-5.08
	Experimental	43	30.28%		
Translated without dictionary	Control	43	38.74%	1.08	-4.45
	Experimental	44	30.99%		
Course input	Control	8	7.21%	-20.18****	14.65+++
	Experimental	45	31.69%		
Knew this language before class	Control	44	39.64%	1.33	-4.20
	Experimental	44	30.99%		
Other	Control	8	7.21%	0.00	-5.53
	Experimental	10	7.04%		

Note: Rate of use and log-likelihood values calculated for proportions of metaphor source forms collected (Control = 111, Experimental = 142), rather than the total number of sources indicated on the forms. This was because learners could select more than one source on each form.

a = Negative values indicate a lower rate of use in the control condition.

**** = Significant at the 0.0001 level

+++ = Very strong evidence against the null hypothesis

The influence of the first language can be seen in learners' frequent use of L1 dictionary entries and translation when writing. In combination, these strategies were selected on 73.87% of all forms in the control group, and 61.27% in the experimental condition. Clearly, learners were drawing on their L1 to a considerable degree when writing, especially when compared to the rates for L2 dictionary usage.

Another point of interest is the reasoning that learners gave for selecting the 'others' option on the feedback form. Learning that had occurred outside the classroom environment was one reason; in four cases, learners cited teachers of

other courses, personal friends, or songs as the inspiration behind a particular word choice. The most common reason given, in nine cases, was that it was a personal choice that was made at the time of writing. These learners were to some degree experimenting with language, saying a word “sounded cool” or that it seemed to fit a meaning they wished to express. In four other cases, learners combined a personal choice with one of the other sources, such as selecting from a range of options provided by a dictionary. In the final case, the learner simply said that the word had been chosen in error.

A second question to consider was how successfully learners were able to produce language drawn from the various sources. Table 7.22 shows the number of times each source was used in the two conditions and the rate of error as identified with the NICT JLE error tag set in each case. It is noteworthy that the error rate is higher for experimental group learners in every category, although none of the differences were found to be significant with log-likelihood tests. It is also somewhat surprising that learners who translated without using a dictionary were apparently no less successful than those who did, either in the L1 or L2. The other notable feature of metaphors that were drawn from course input is that the error rate was much lower than that of any other source, whether L1 or L2. To some extent, this was because learners could simply base their own language on the course example, but this does at least provide them with a starting point for incorporating this new chunk of language into their productive range.

Table 7.22. Log-likelihood tests for the relative frequency of errors from each source between conditions

Source	Condition	Uses	Errors	Error rate (% of uses)	Log-likelihood [a]	Bayes factor effect size
L2 dictionary	Control	15	7	46.67%	-0.34	-2.92
	Experimental	11	7	63.64%		
L1 dictionary	Control	39	22	56.41%	-0.40	-4.00
	Experimental	43	29	67.44%		
Translated without dictionary	Control	43	20	46.51%	-0.15	-4.32
	Experimental	44	23	52.27%		
Used words from class	Control	8	2	25.00%	-0.24	-3.73
	Experimental	45	16	35.56%		
Knew this language before class	Control	44	25	56.82%	-0.08	-4.40
	Experimental	44	27	61.36%		
Other	Control	8	4	50.00%	-0.62	-2.27
	Experimental	10	8	80.00%		

Note: Log-likelihood values calculated for proportions of uses that contained errors. Minimum number of errors = 5 in either corpus.
a = Negative values indicate a lower rate of error in the control condition.

Table 7.23. Log-likelihood tests for relative frequency of error types between conditions

Error type	Errors (% of forms)		Log-likelihood [a]	Bayes factor effect size
	Control	Experimental		
No error	56 (50.45%)	67 (47.18%)	0.14	-5.40
Grammatical	24 (21.62%)	44 (30.99%)	-2.07	-3.46
Lexical	18 (16.22%)	18 (12.68%)	0.54	-4.99
Phraseological	11 (9.91%)	12 (8.45%)	0.15	-5.39
Other	2 (1.80%)	1 (0.70%)	/	/

Note: Error rate and log-likelihood values calculated for proportions of metaphor source forms collected (Control = 111, Experimental = 142). Minimum number of errors = 5 in either corpus.
a = Negative values indicate a lower rate of this error type in the control condition.

Finally, the categories of error produced by learners in each condition were compared (Table 7.23). It appears that the proportion of forms containing a

grammatical error was higher in the experimental condition, but in fact log-likelihood tests found no significant differences in the proportion of error types produced between the conditions. This would be an interesting area to investigate in an expanded study, as the possibility that an increased emphasis on lexical information may lead to a relative inattention to syntax echoes the Resource Depletion for Output hypothesis (Barcroft, 2004, p. 325 - 327, 2006, p. 494 - 496)

The main limitation of the data provided here is that the metaphors were not selected randomly for feedback surveys, and so they may not be truly representative of the entire range of metaphors produced in the study. The sample size also limits the power of the statistical analyses performed here. Nevertheless, the surveys were able to demonstrate that learners draw on a variety of sources for their productive output, and that the influence of the L1 is a prominent feature of learners' lexical selections that exhibits neither a wholly positive nor negative influence on learner writing.

7.6.2. L1 metaphors and word frequency as a predictor of L2 output

Another way that learners' output may be affected is through the influence of the L1 on their conceptions of the world around them and their intuitions about language. It has been claimed that if two language groups share a mental representation of how the world is ordered, then it might be expected that idioms realising that shared perception would be more easily learned, and perhaps produced, by language learners than expressions that draw on an unshared metaphor (Boers & Demecheleer, 2001; Türker, 2016; Yeganehjoo, Yap, Abdullah & Tan, 2012). The question to be considered here was whether the frequency of individual words

might also be a factor, with forms that appear more often in the L1 being easier to draw on in L2 use than expressions that are relatively infrequent in the L1.

To obtain a sense of the linguistic patterning for metaphorical verbs used to describe cultures in Japanese, the Sketch Engine corpus analysis software was used to create a list of common collocates of the word *bunka* (*culture*) in the jpTenTen11 corpus. Japanese translations of all metaphorical verbs that appeared in collocations with *culture* in learner writing were searched for in jpTenTen11, and those found to collocate significantly under either the MI or t-score measure were listed. In cases where different translated forms were found to collocate with *bunka*, the most frequently-occurring collocate was selected. The complete list of common verb collocates is provided in Appendix S (page 511), and Table 7.24 lists the Japanese verb collocates of *bunka* whose English translations were classed as target metaphors in Appendix G and Appendix H (page 443, 450). The forms *bring up* and *grow up*, whose translations are strong collocates of *bunka* but which are rare as collocates of *culture* in English and may appear marked in learner writing (see examples 136 - 137), were included on the list in Table 7.24 since they were a potential source of L1 interference. The information in the table seems to show that at least some metaphor themes in Japanese bear a reasonably close similarity to the English metaphors covered in class. Further evidence for this came from the metaphor surveys that were described in the previous section. Learners who had translated from their L1 were asked to provide the Japanese expression they had worked from, and these translations sometimes included verbs from the table above, as in (135). However, there were also translations provided that contained other

metaphorical verbs not in this table, so it was clear that learners were drawing on a wider range of expressions than those captured in the list of collocates.

(135) Japanese culture absorb cultures from other countries. (Nihon wa hoka no kuni kara bunka wo kyuushuu shita) [CON3]

Table 7.24. Japanese metaphorical verb collocates of *bunka* (culture) and their English translations

Metaphor theme in English	Japanese verb collocate of *bunka* (culture)	English target word translation	Taught or untaught target word?
MOVING FORWARD IS IMPROVEMENT	進歩 (advance)	advance	Taught
CULTURES ARE LIVING THINGS	進化 (shinka)	evolve	Taught
	育つ (sodatsu)	grow	Taught
	育む (hagukumu)	bring up; grow up	Untaught
CULTURES ARE SPACES	入り込む (hairikomu)	come into	Taught
	入る (hairu)	enter	Taught
	輸出 (yushutsu)	export	Taught
	輸入 (yunyuu)	import	Taught
	取り入れる (toriireru)	take in	Taught
CULTURES ARE SUBSTANCES	吸収 (kyuushuu)	absorb	Taught
	流れ込む (nagarekomu)	flow	Taught
	混在 (konzai)	mix	Taught
	形作る (katachidzukuru)	shape	Taught
	広める (hiromeru)	spread	Taught
	結び付ける (musubitsukeru)	combine	Untaught
	形成 (keisei)	form	Untaught
CULTURES ARE CONSTRUCTIONS	合う (au)	fit	Taught
CULTURES ARE POSSESSIONS	失う (ushinau)	lose	Taught
	共有 (kyouyuu)	share	Taught
	受け入れる (ukeireru)	accept; receive	Untaught
	取り上げる (toriageru)	adopt	Untaught
	持つ (have)	have	Untaught
ABSTRACT ENTITIES ARE PHYSICAL STRUCTURES	繋がり (tsunagari)	connect	Untaught

The process of drawing up the list of Japanese verbs to express culture led to some possible insights into learners' word choices. One pattern in learners' writing, also commented on in section 7.5.2, was the use of the phrasal verbs *grow up* and *bring up* as collocates of *culture*. The corpus searches revealed that verbs with basic senses referring to the raising or growth of children (e.g., *hagukumu*, *sodateru*) collocate more strongly with *bunka* than verbs that refer to plant growth. This suggests that although a similar broad metaphor theme exists in English and Japanese, specific realisations of that metaphor may differ between the two languages. Had learners selected *foster* or *nurture* instead of *grow up* or *bring up*, the sentences in (136) and (137) would appear less marked.

(136) Japan's society changed enormously of course we must save Japanese unique culture, but culture should grow up more and more. [EXP20]

(137) [Japanese people] try to mix foreign culture to bring up their culture. [EXP1]

Facilitative effects were also in evidence. The verb *spread* was frequently used to refer to culture by learners in both conditions, and *extend* and *expand* also appeared occasionally. *Spread* is a frequent collocate of *culture* in English according to COCA, and its translations *hiromeru, hirogaru* and *fukyuu* are comparably frequent as collocates of *bunka* in Japanese. Therefore, it is probable that for learners, the word *spread* would be an obvious choice when describing cultural flow, and the L1 can be seen as providing a facilitative effect. However, in one case, a learner who reported in the feedback form that they had used an L1 dictionary reference translated *hirogaru* as *widen*, leading to example (138). Although this usage was not acting as a collocate of culture and so was not included in the analysis

for this section, it serves as an illustration of the difficulties of selecting natural expressions that learners face even when the same metaphor exists across two languages.

(138) It is good for Japanese people to using new technology to create new music because it can widen to the world. [EXP21]

A further hint that L1 frequency might be influencing learners' choice of expressions came from the verb *shape*. In COCA, this was the second most frequent collocate of *culture* amongst the verbs in Table 7.24 - only *have* collocated more often. However, in Japanese, the strongest translation of *shape* (*katachidzukuru*) appeared in a collocation with *bunka* 20 times less frequently. In learner writing too, the word *shape* appeared only twice with direct reference to cultures. It is possible that the lower likelihood of collocation in the learners' L1 meant that they were not primed to see *shape* as a candidate for collocation in English. An alternative hypothesis is that the relatively greater frequency of *shape*, and its translations, *katachi* or *kakkou*, as nouns have the effect of blocking access to verbal usage (Ellis, 2006b, p. 176 - 179).

(139) [The] concept of hybridity is a way of better understanding the transnational forces that shape of the identities of nations and cultures. [CON6]

(140) Moreover Japanese society are shaped by new popular culture in Japan. [EXP4]

To estimate the strength of the relationship between learners' lexical choices and frequency effects in the L1 and English, frequencies were obtained for the

individual Japanese collocate verbs as well as the *bunka* collocations for each verb in Appendix S (page 511). The equivalent frequencies of the English verbs and collocations were also taken from COCA to estimate the effects from English in general and the two input corpora to estimate more localised L2 influences. Spearman's Rho correlations were then determined between word frequencies in the learner output corpora and in COCA, jpTenTen11, and the input corpora (Table 7.25). These calculations were also performed for the frequencies of collocations in these corpora (Table 7.26).

The results in Table 7.25 show that the strengths of the relationships in word frequency between the L1 and L2 (as determined by COCA and jpTenTen11) and learner output were very similar. In all cases, moderate correlations were found. The correlation between the word frequencies in COCA and those in jpTenTen11 was also investigated and found to be strong (r_s = .585, p = .0004), which may explain the similarity in the correlations between the larger corpora and learner output. In the case of correlations with the input corpora, the experimental group's output correlated strongly with the input they received. This suggests that learner output was somewhat responsive to the input they received, at least regarding the use of metaphorical verbs. A Fisher r to z transformation was then used to determine the significance of the differences between the correlations in the two conditions, but none were found to be significant.

Regarding collocations, the correlations were weaker in almost all cases than those for individual word frequencies. The values for correlations with Japanese collocation frequencies were marginally higher than the frequencies determined in COCA, but the difference is too small to make any strong claims. In general, the

Table 7.25. Spearman correlations between the normalised frequencies of metaphorical verbs in the learner output corpora and those in COCA, jpTenTen11, and the input corpora

Condition	COCA word freq.		jpTenTen11 word freq.		Input corpus word freq.	
	n	***r_s* (*p*)**	***n***	***r_s* (*p*)**	***n***	***r_s* (*p*)**
Control output word freq.	29	.444 (.016*)	29	.406 (.029*)	29	.354 (.059)
Experimental output word freq.	31	.463 (.009**)	31	.448 (.011*)	31	.641 (.0001***)
Fisher r to z transformations between r_s values						
Fisher *r* to *z* transformation (p)		-0.09 (.928)		-0.19 (.849)		-1.43 (.153)

Note: Verbs listed in Appendix S with a frequency of zero were not included in the correlations.
* = Significant at the 0.05 level; ** = 0.01 level; *** = 0.001 level

Table 7.26. Spearman correlations between the normalised frequencies of metaphorical verb + *culture* collocations in the learner output corpora and those in COCA, jpTenTen11, and the input corpora

Condition	COCA collocation freq.		jpTenTen11 collocation freq.		Input corpus word freq.	
	n	***r_s* (*p*)**	***n***	***r_s* (*p*)**	***n***	***r_s* (*p*)**
Control output collocation freq.	21	.209 (.364)	21	.347 (.124)	10	.615 (.058)
Experimental output collocation freq.	29	-.115 (.554)	29	.289 (.128)	17	.283 (.272)
Fisher r to z transformations between r_s values						
Fisher *r* to *z* transformation (*p*)		1.07 (.285)		0.21 (.834)		0.92 (.358)

Note: Collocations listed in Appendix S with a frequency of zero were not included in the correlations.

results do not support the notion that learner output of metaphorical verb + *culture* collocations is strongly related to the frequency of those collocations in wider use either in the L1 or L2. The one case where the *r* value is high is that for the correlation between output and input in the control condition, but here the *n* size has become rather low, which may be affecting the results. The correlation between frequencies of collocations in COCA and jpTenTen11 was found to be r_s = .347 (*p* = .051), which is noticeably weaker than that for individual words.

7.7. Concluding comments

This chapter has examined some of the features of language itself that can affect learner production of metaphorical vocabulary. Considerations such as word frequency, part of speech, phraseology, and the first language can influence learners' initial selection of language elements as well as their likelihood of successfully incorporating their choices into conventional lexico-grammatical patterns. The findings have highlighted the non-linearity of language development, with signs of underused linguistic features as well as forms that learners had perhaps become overly-dependent on being repeatedly uncovered in the data. This is fully in-line with studies that have investigated learner output from the perspective of dynamic systems theory (de Bot, Lowie, & Verspoor, 2007; Larsen-Freeman, 2006) or cognitive linguistics (Ellis, 2006a, 2006b), as well as those that have applied corpus-driven techniques to learner writing (Granger & Rayson, 1998; Groom, 2009; Shaw & Ting-Kun Liu, 1998).

Frequency certainly impacts on learners' lexical choices. A general overusing or dependence on high frequency forms was seen when learner output was compared with that of native speakers writing on similar topics. Limited vocabulary sizes may account for this; the comparative lack of more precise terms forces them to recycle high frequency words in their output, such as *have*, *thing*, and *big*. However, the analysis of individual words which were significantly favoured by one group over another (Appendix P, page 493) revealed that many high frequency words were also relatively underused as metaphors by learners. These included the words *see*, *give*, and *show*, which were commonly used by BAWE writers to offer comments or

summaries of content. On a broad scale, fewer differences were found in the use of metaphorical vocabulary from the mid-frequency bands between the learner groups and the native speakers, although again the analysis at the level of individual words revealed a trend for language learners to underuse metaphorical nouns. Less frequent words were shown to be much more frequent in native speaker writing. Nouns and verbs that are especially frequent in academic discourse seemed to comprise the better part of this difference, while learners drew on adverbial polywords as discourse organisers much more than native speakers. These results seem to imply that vocabulary analyses based on frequency should be carried out on both macro- and micro-levels. On the one hand, relatively low level learners appear to be dependent on highly frequent words and make relatively lesser use of infrequent terms. But on the other hand, we see that even within the frequency bands shown to be significantly favoured by one group over another, several individual words are exceptions to the general trend. These exceptions may be of particular value for bringing to learners' attention, either to raise awareness of unexploited potential in the case of high frequency terms, or to stress the value of alternatives to words and phrases that learners already make extensive use of.

Patterns of over- and underuse were not solely determined by frequency, however. There were noticeable differences between learners' use of various parts of speech to produce metaphors. In particular, nouns were underused as metaphors in both conditions, and instances of relative overuse were often driven by a reliance on certain word choices. For example, although adverbs were used very frequently by learners, this was almost entirely due to the used of fixed expressions such as *on the other hand* and *in short*, rather than the more frequent, yet perhaps less salient,

metaphorical adverbs *far* and *here*. Similarly, learner production of metaphorical adjective/noun collocations was largely based on exploiting basic adjectives such as *big* and *high*, rather than the class of adjectives as a whole. This is probably a typical feature of learner development. The pattern of use, overuse, and presumably later diversification would allow learners to become comfortable with the latest additions to their productive repertoires before moving on to new expressions.

As examples of learner error in this chapter demonstrate, acquisition is certainly not a single-step process. The learning of new lexical items requires developing mastery of the whole range of semantic and syntactic patterns in which they are used, as well as building awareness of the conceptual referents they subsume (Langacker, 2008b, p. 225). Variations in usage of the different phraseological patterns seen in this chapter also suggest that educators should not expect all learners to respond in the same manner to activities designed to raise awareness of metaphor, as for some this may entail learning entirely new expressions, while others merely require linguistic fine-tuning.

The influence of the L1 was investigated in combination with the effects of frequency. It was seen that the L1 frequency of translations of metaphorical verbs had as strong a relationship with learners' English output as did the frequencies of those words in English. However, the fact that the frequencies obtained from large-scale corpora of Japanese and English also correlated strongly may suggest that what we are seeing here is a similarity in the frequency of particular concepts across languages, rather than the influence of L1 and L2 lexical frequency on learners' word choices. It was interesting that the correlations between collocation frequencies in the L1 and L2 and learner production were lower than those of individual words. This

is a similar finding to that of Macis and Schmitt (2017, p. 331 - 335), who found no evidence of a frequency effect in learners' knowledge of collocations. One explanation the authors put forward for this is that the relative infrequency of collocations compared with individual words means that the effects of learning through exposure are correspondingly weaker than those believed to facilitate the acquisition of other language patterns.

Finally, although it may not be possible in all contexts, there are apparent benefits to giving consideration to learners' first language when analysing productive use of metaphorical vocabulary. The feedback form data suggest that the L1, either in the form of dictionaries or translation, represents a substantial influence in learners' lexical selection processes. As the analysis of collocates of *bunka* in Japanese showed, there are often similarities in how different languages encode abstract concepts, and the L1 can therefore have a facilitative effect on lexical selection. Nevertheless, the examples with *widen*, *bring up*, and *grow up* in this chapter also demonstrate that incorrect or marked choices can still be made even when the broader metaphorical theme aligns between the L1 and L2. If it is accepted that L1 influences are unavoidably present in the second language learning environment, then the question becomes one of how to maximise the benefits of this situation. External resources, such as collocation dictionaries or concordancers, may be of some use here, if they are detailed enough to provide information on specialised usage and yet still accessible by learners of a range of proficiencies. Within the classroom, the answer may well be to raise awareness of language features such as collocation and the unfortunate reality that shared metaphor themes do not necessarily entail shared actualisations of those themes. In doing so,

the goal would be to train learners to notice phraseological patterns and to learn to second guess their own instincts regarding linguistic combinations.

CHAPTER 8. CONCLUSIONS, IMPLICATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

Obtaining the metaphorical competence required to master academic writing is a formidable challenge for language learners. As well as becoming familiar with the conventions for expressing abstract concepts, they must learn to indicate stance, emphasise and moderate opinions, and to combine all of this into coherent, flowing text that accords with the norms of their discipline. Yet metaphor, and indeed linguistic competence itself, should also be considered as more than just the knowledge base required to join a community; it is also the development of a personal style, a finding of an individual voice that will allow learners to take ownership of language as a tool to exploit in their own ways.

This study has explored the effects of a sustained classroom intervention designed to raise awareness of content-specific metaphorical patterns. It has considered the impact of the experimental treatment from multiple perspectives so as to capture as much detail as possible. The main focus of the study has been on learners' written output in response to course content. Learner development has been examined by comparing the metaphorical output from a class who received explicit instruction on metaphor themes with that of another group whose instruction focussed on high frequency and academic vocabulary, with metaphor exposure largely coming through incidental input. It has also considered the metaphor production of learners at different levels of proficiency, learners' use of metaphorical target language, and their harnessing of metaphor for rhetorical

purposes. Taking a broader perspective, the study has looked at how different approaches to instruction led to changes in the quality and quantity of input, and on how this might relate to learners' written output. It has also examined the longitudinal effects of sustained metaphor instruction, and finally, with regard to the influence of language itself, the study has considered the impact of word frequency and part of speech on learner writing, as well as how both phraseology and learners' first language play a part in their written production.

The study has revealed a mixed picture of outcomes from raising learners' awareness of metaphor. At the broadest levels of analysis, production of metaphorical vocabulary was found to be only marginally higher in the experimental condition, and these differences were non-significant. At a narrower level of detail, however, some differences did emerge. Evidence was found for increases in production of explicitly taught target language both in terms of quantity and the variety of types used, and there were signs of change in the use of metaphor to achieve different rhetorical functions. There were also interesting outcomes within the conditions, as the general pattern for increased use of metaphor with higher proficiency in the control condition was much less evident in the experimental group. This suggested that lower-level learners (those with TOEIC scores ≤ 420) appeared to show the greatest difference in performance between the conditions.

The data have also revealed an enormous complexity within the learning environment. It is clear that metaphorical and lexical development can and should be considered in various ways, with changes in the amount, variety, and accuracy metaphorical output all evident in the data. Nor should we only view development solely in positive terms; avoidance and error are part of the learning process, as are

declining rates of usage, since they can reflect learners moving away from overused forms.

This final chapter will review the outcomes of the study and discuss the main findings. It will consider the effectiveness of the metaphor identification procedure used and the benefits and limitations of classroom research designs. It will review the place of metaphor in vocabulary instruction and offer suggestions on pedagogical practices for bringing metaphor awareness into EAP settings. Finally, it will consider the limitations of the study and how they might be addressed in future investigations that would enhance understanding of learner metaphor production.

8.1. The effectiveness of the metaphor identification procedure

Instances of metaphor production were identified with a variation of MIPVU. The first variation (in chapter four and sections 5.3 - 5.5 of chapter five) allowed for a small number of target forms to be coded as metaphors despite not appearing in separate numbered senses from a more basic sense in their dictionary entry (section 3.4.2.2, page 68). This was to allow for the inclusion of metaphorically derived forms that were part of important content themes in the data set. The second variation (in chapter five, section 5.6 and chapters 6 - 7) adopted a more inclusive approach to polyword identification by basing this on dictionary entries instead of the BNC polyword list and by coding these forms as metaphorical if any component word would be coded as a metaphor when analysed individually. This allowed for a more consistent treatment of multi-word units, which was important for the analysis of rhetorical function use, particularly in the case of text-structuring metaphors. These variant procedures were applied to over 100,000 words of learner writing in the

output corpora. The input corpora, at over 250,000 words each, were too large to be fully examined in this way. Instead, they were partially analysed for metaphorical content using a combination of the MIPVU and the semantic tagging feature of the Wmatrix software tool.

In general, the MIPVU proved to be an effective method of extracting metaphors from learner writing. In most cases, the detailed steps laid out in Steen et al. (2010) could be applied without difficulty to the data, and in cases where the idiosyncrasies of learner writing did cause issues, the Nacey (2013, p. 117 - 120) and Littlemore et al. (2014, p. 130 - 141) studies of metaphors in learner language were precedents upon which to base decisions. The inter-rater reliability levels that were reported in sections 3.4.4 (page 88) and 7.2.2.1 (page 295) indicate that the data coding was consistent with other researchers' interpretations of the procedure.

However, while reliability is an important feature of measurement validity, other aspects of the MIPVU should also be evaluated. One concern is that while the procedure has been shown to perform reliably, there must be some question over the consistency of the sources upon which it is based. The list of polywords used in the MIPVU is an example of this. The actual list is the one used for tagging the British National Corpus and it contains 541 polywords, including Latin phrases, cohesive devices, slang expressions and names of foods. What may be problematic is that there is no explanation of the principles that guided the creation of this list, and in tagging data with the MIPVU, one frequently encounters phrases that would appear to be viable candidates for the list but are not included. For example, *in comparison with* and *in conjunction with* appear on the list, but not *in contrast with*, and other frequent, relatively fixed phrases such as *on the other hand*, *in order to* and *in fact*

are also absent. Since MIPVU does not analyse parts of polywords for metaphor coding, and many of the potential polywords that do not appear on the list contain metaphorical prepositions, this has the potential to skew results. The list of polywords in Appendix D (page 433) shows that between 0.2% and 0.4% of all lexical units in the output corpora and BAWE sample analysed in this study were potential polywords that did not appear on the BNC list. The effect of this on overall metaphor counts is harder to determine, but in more narrowly-focussed analyses, it could be significant; 15.21% of the structuring metaphors in the control condition and 10.69% of those in the experimental condition were not on the BNC list.

Other issues noted during the coding process related to the dictionaries themselves. Although dictionaries are considered to provide "complete and culturally sanctioned representation[s]" of lexical knowledge (Steen et al., 2010, p. 7) it can be surprising when carrying out MIPVU to discover how differently words can be defined in separate dictionaries. The verb *combine* appears in the Macmillan dictionary in a single numbered sense that offers a general definition covering both physical and abstract combinations: 'if you combine things, or if they combine, you use, do, or put them together.' Two (non-numbered) sub-senses offer definitions relating to the combining of substances and organizations. The Longman dictionary, meanwhile, uses five numbered senses to define *combine*, including the most general sense, 'if you combine two or more different things, or of they combine, they begin to exist or work together,' and two that match the sub-senses in Macmillan. Of the remaining senses, the first ('to do two different activities at the same time') appears to be related to Macmillan sense 1, while the other ('to have two or more different features or qualities at the same time') does not seem to match a

Macmillan definition. MIPVU requires that senses appear as separate numbered entries in order to show sufficient distinctness of meaning, but as we can see, the distinction between senses can be drawn in various ways. Hanks (2013, p. 70) makes a similar point, noting that particular dictionaries, or even individual lexicographers, can be more or less prone to 'lumping' or 'splitting' word senses. What is being encountered here is a tension between the need for metaphor identification to move beyond subjectivity into clearly defined categories, for dictionaries to present information as concisely and as clearly as possible, and the natural tendency for language to defy hard and fast categorisation.

Despite these concerns, MIPVU should be considered a step forward in the development of metaphor research tools, and one that can work well in conjunction with corpus-based approaches to metaphor identification. While it is undoubtedly a time-consuming process that cannot be applied to large datasets, MIPVU is a useful step towards the analysis of larger corpora. It worked well in tandem with Wmatrix, providing a thorough inspection of the smaller output corpora that yielded a list of search terms which could then in turn be expanded upon with the Wmatrix semantic tagger.

Regarding the analysis of developments in learner language, the MIPVU should probably be seen as a first step only. It is a maximally inclusive method of identification that produces estimates of the quantity of metaphor in a text, but as this study has shown, raw quantities can mask more subtle developments in language learning. Nevertheless, the MIPVU output is valuable grist for deeper investigations into learners' interlanguage.

8.2. Classroom studies on metaphor development

One major concern with studies that have applied cognitive linguistic principles to metaphor instruction has been the relative lack of evidence of long-term gains for learners. It could be the case that many learners are unfamiliar with the notion of motivated language, which leads them to perceive it as difficult and thus increases the chance of avoidance (Philip, 2007, p. 7). Consequently, it has been suggested that awareness-raising activities should be a recurring feature in the classroom so that learners can gradually gain confidence and be allowed to learn from errors and supportive feedback (Boers, 2004, p. 216 - 217). This points to a need for classroom-based studies (as opposed to more tightly controlled experimental designs) that can provide such repetition and reinforcement in an authentic environment. Evidence for the need to provide repeated encounters and explicit learning opportunities with metaphor can be seen in Figure 6.9 (page 269), which suggests that the experimental group overcame an initial deficit in production but rapidly gained in either confidence or willingness to produce metaphorical language.

Another benefit of classroom-based studies is that they allow for the rich specificity of context to support findings. Rather than relying solely on fixed data sources such as pre- and post-tests to describe the outcomes of the study, teacher-researchers are able to bring in contextual evidence to support the narrative of the study. This can most clearly be seen in instances where learner output could be related to input features, such as the particular prompts used for writing topics and in the responses to the metaphor source feedback forms used in chapter seven. It was also possible to allow for the data to guide the investigation itself to some

degree, reflecting the spirit of corpus-driven research. This can be seen in the analyses of phraseological patterns and errors produced by learners.

The drawback of this kind of data collection procedure is that it is incredibly time-consuming and cannot easily be downsized for smaller-scale studies. The degree of insight into the learning process gained is a direct result of intensive data collection. Nor do some of the findings of the study offer much in the way of 'quick-fix' solutions that can easily be taken into the classroom. To some extent, teachers must accept as faith that a metaphor awareness-raising approach will pay off in the long term, even if learners may initially produce less metaphor.

Another limitation to this study is that learners were writing in aided, rather than examination conditions. As such, claims cannot easily be made for learners' actual knowledge; they are instead providing evidence more of their attempts to add to their productive repertoire. An obvious follow up to this study would, therefore, be to obtain samples of learner writing without the aid of dictionaries or other reference materials. Not only would this allow for claims to be made for acquisition, but it would also provide a useful comparison with learner writing both with and without language support.

Finally, although it may well be a simplification to view language acquisition as progressing directly from receptive to productive knowledge, adding a measure of receptive uptake of metaphorical forms would undoubtedly enhance the findings of a follow-up study. As was pointed out earlier, lack of productive use cannot be taken as evidence of learning not having taken place, and it is likely that learners acquired an understanding of metaphorical words that was not evinced in their writing. However, such a measure would need to be carefully designed so as to avoid a

testing effect in which the measurement instrument itself provoked the learning, rather than the course instruction.

8.3. The relationship of metaphor to lexical knowledge

The basic paradox of vocabulary teaching is that while it is increasingly recognised that explicit instruction is a vital part of foreign language vocabulary learning (Laufer, 2005; Nation, 2013, Chapter 3), the sheer enormity of the lexical challenge means that much of a learner's knowledge will be gained from implicit learning opportunities. Therefore, it seems logical that, alongside the direct instruction of high frequency words, awareness of important general features of vocabulary should also be raised so as to increase the likelihood that learners will notice such features and improve the efficiency of implicit learning (Schmidt, 1990, p. 149).

Two of the most important features of vocabulary that would have to be covered in this approach would be phraseology and polysemy. Both are ubiquitous in language, particularly in academic discourse, but they are also highly complex, diverse phenomena that resist precise definition (Moon, 1997; Taylor, 2003; 2012, Chapter 10). Because of this, they may be ideal candidates for approaches that promote noticing and awareness, rather than there being an expectation of mastery. Metaphor clearly has a part to play in this, as it drives the creation of polysemous word forms and in so doing helps to create the phraseological patterns that distinguish word senses. It has been argued in this thesis that metaphor awareness raising is ideally suited to CLIL or content-based instruction settings firstly because metaphor is a common feature in academic discourse, and secondly because tertiary level learners are more likely to have acquired basic vocabularies and would in

theory require less study to acquire additional senses of known words. Thus, there is a need for more investigations that explore metaphor awareness raising across other disciplines.

Studies into learner metaphor use would also help to fill gaps in our understanding of how learners' vocabularies grow. Currently, this is an under-explored area of research; as was mentioned in chapter two, studies of vocabulary growth are somewhat rare, and those that measure aspects of lexical depth rarer still. The findings of this study, along with those of the Littlemore et al. (2014) and Turner (2014) studies described earlier, have suggested similar trends in productive use of figurative vocabulary, with a gradual increase interspersed with short plateaus or even slight drops in production. With regard to the experimental treatment in particular, the finding in this study that lower proficiency learners responded positively to metaphor awareness raising implies that there is scope for improving the lexical variety of learners at the A2 level or with TOEIC scores of less than 500. However, these studies have all relied on comparing the production of learners at different levels of ability to make these claims about changes in metaphorical output, and it would also be very useful to obtain data from individuals over longer periods of time to observe change in single learners.

Although metaphor can be considered present in the lexicon, it would be far too simplistic to view it as a mere element of vocabulary as there are clearly other aspects to its nature. One other way that metaphor could be used to support the development of vocabulary would be from a functional perspective. While language programs and some textbooks may take a functional approach to instruction, vocabulary is more often addressed in topical themes or in frequency bands.

Therefore, drawing on metaphor to explain the motivation behind language used to achieve rhetorical goals may enable learners to use their lexical knowledge more purposefully.

8.4. Pedagogical implications raised by the study

This study has mainly focussed on the effect of metaphor instruction on language learners, but several of the findings obtained in the study can be seen as having import for classroom practice as well.

The first is that when there is an expectation that learners will use target metaphorical language as part of their class production, it may be beneficial to limit the number of metaphorical themes that are introduced. In this study, 19 metaphorical themes were explicitly taught in the experimental condition, and although the rate of production increased in many cases, for eight of the themes, target forms were produced on average less than once per learner in the experimental condition (i.e., less than 23 uses) (Table 4.4, page 140). This indicates that there were many learners who did not use some of the expressions or some themes, which may have been due to them being overloaded with target language. A similar impression can be gained from Tables 6.13 - 6.18 (pages 259-263), which shows that only three of the taught target metaphorical words used to describe culture were produced by more than half of the class members. Certainly, some topics appeared only fleetingly in the course, so there were fewer opportunities to produce metaphorical language related to them, but the theme of cultures was more or less omnipresent, and yet still many words did not appear in learner output. Had the study focussed only on metaphors of culture, it is possible that learners

could have concentrated their efforts on producing a narrower range of metaphors, which would have led to more feedback, greater confidence and potentially more accurate or appropriate usage.

In a similar vein, the data presented in Figure 6.9 imply that experimental group learners may have initially avoided using metaphors of culture. This may have been due to the perceived difficulty of using such language; some early forays into metaphor use led to errors that were picked up for class feedback, and this could have had an inhibiting effect. Without further investigation, no firm claims can be made for the reason why learners initially made less use of target forms than their control group peers. The literature on vocabulary development, however, does stress the value of periodic review and the incremental nature of lexical acquisition (Nation, 2013, Chapter 14; Schmitt, 2010, p. 19 - 22), and none of the findings of this study would contradict this. Therefore, it may well be prudent to build repetition and supportive feedback into a course when attempting to teach metaphor to students who are mostly at the A2 level (earlier than studies of examination writing have found much evidence of open-class metaphor use).

The data from this study have also shown tendencies in learner behaviour that may have been caused by the experimental treatment itself or some external influence. One example of this is that learners in this study appeared less willing or able to use the metaphorical senses of some parts of speech than others. It is possible that salience and perceived difficulty lie behind the apparent disinclination among learners to use nouns in their metaphorical senses. Textbooks typically prioritise verb phrases in grammar instruction, while treatment of nouns is often limited to the countable and uncountable distinction. As such, learners may simply

feel more comfortable constructing their sentences based around verbs, leading to the noted feature of learner writing bearing more in common with spoken rather than written discourse. It is possible that lower-level learners may be benefit from the provision of phraseological patterns that metaphorical nouns commonly appear in, so that they have a template around which to construct a sentence using the new language.

It is also important to consider the effect of explicit instruction on the lexical selections that learners make. Sections 4.5 and 5.5 dealt with learner production of target language, and in both cases, the findings suggest that the experimental treatment had the effect of steering learners towards metaphors that were explicitly taught in class and away from other metaphors that belonged to target metaphor themes but had not been explicitly taught. Whether this unintended homogenising effect on learner writing is a negative outcome or not is debatable. It could be claimed that instructors should be aware of this likely outcome and make efforts to encourage learners to continue experimenting with word forms in their writing and not to be dependent on teachers for new language. Equally, this finding could be seen as a stage in learner development during which target forms are given greater attention, presumably so that learners can consider them more carefully, and as a result, other forms are given less emphasis. Further data collection of learner writing after the period of explicit instruction has ended may reveal whether learners' use of particular word forms is permanently altered by explicit instruction or whether this is just a temporary foregrounding effect. This also raises the question of whether learners in the study were responding to the teaching on metaphor themes or to the

actual lexical items that exemplified those themes, a point that will be returned to in the final section.

8.5. Developments in metaphor production

It has been claimed in this study that learners in academic settings can benefit from an extended focus on content-specific metaphor in their courses. Comparing the data from the two conditions provides indications of how learners' written output may have been affected and offers insights into how raising awareness of metaphor can influence lexical development.

In the control condition, when exposure to metaphor came largely through incidental exposure (Figure 6.5, page 253), production of open-class metaphor was shown to increase with overall proficiency. This is an important finding, as other studies that have suggested this same trend have not been able to entirely remove the effect of writing topic on learners' output. However, in this study, learners of all abilities and in both conditions wrote on the same topics, so it can be claimed more conclusively that one marker of developing lexical proficiency is that learners will increasingly draw on the extended, figurative senses of open-class parts of speech in their written output. The caveat that must be added is that they were doing so with the aid of dictionaries and course materials, so the data are less reflective of their own internal lexical acquisition than the richness of their productive output in an aided environment.

In the experimental condition, learners received sustained instruction intended to raise awareness of content-specific metaphor. There is evidence that this had an impact on written output, though the effects varied with proficiency. Lower level

learners (those in the A2 CEFR band, or with TOEIC scores <500) tended to show an increase in the quantity of open-class metaphors that they produced relative to the control group (Figures 5.5 - 5.7, pages 177-178). This is a predictable outcome, as these learners are least likely to have developed core vocabularies or the ability to use words in extended senses. Since the experimental treatment was targeting an identifiable gap in these learners' capabilities, it is only to be expected that they showed the clearest response to the treatment. In the case of higher ability learners, the effects of the experimental treatment were less evident in quantity of production than in the variety of metaphors produced. In fact, all ability groups produced a greater variety of taught target metaphors in the experimental condition (Table 5.20 and Figure 5.13, pages 188-189), but the high level learners did so despite producing fewer taught target metaphors overall than the high level control group learners. That the higher ability learners should exhibit greater changes in variety of use than quantity seems plausible, given that the control group results indicated that these learners were already capable of producing metaphorical vocabulary without explicit instruction. However, a major limitation of the study was that the power of many statistical analyses was limited by the relatively small sample of participants. Moderate to large effect sizes were not infrequent in the investigations of learner ability, but relatively few results reached statistical significance. As such, the results are more suggestive than conclusive regarding developments in metaphor production, but they do point to the value of an expanded study that might overcome the limitations of sample size.

Taken together, these findings support a model of metaphor development that broadly mirrors the findings of Ellis and Ferreira-Junior (2009), who suggest that

production of language features is initially based around a prototypical construction, with later diversification following once learners have consolidated knowledge of the original form. In this case, the prototypical form of individual words would most likely be the basic sense; this would be expected to appear earliest in learner output, with metaphorical senses appearing later. Of course, exceptions to this would exist in cases where the metaphorical sense has become dominant in the language or perhaps due to L1 influence.

The results plotted on the vertical axes of Figures 5.25, 5.27, 5.29 and 5.30 (pages 213, 217, 220, and 222, and tabulated in Appendix K, page 473) may provide some indication of how learners of different proficiencies begin to draw on the metaphorical senses of lexical items in their writing. The few lexical items significantly favoured by low-level learners are fixed phrases, target forms, or the noun *world*. The words favoured by high-level learners include some that perform evaluative or emphatic functions (*dark*, *strongly*, *dramatically*) and the class of nouns in general (*field*, *role*, *fashion*, *look*). Yet it is also notable that several of the metaphors favoured by high ability learners are extremely high frequency words that would almost certainly have been known in at least their basic senses by all learners (*look*, *way*, *see*, *say*, *bring*). Low ability learners, however, were clearly making less use of these words as metaphors in their writing, which suggests that one important facet of learner development is 'returning' to already known words and adding to their range of use.

It was also suggested that metaphor production increased in a non-linear fashion in both conditions. The lowest and highest ends of the ability scale were marked by signs of increasing production of open-class metaphors (Figure 5.6, page

178), but for learners in the TOEIC range of 500 - 600, there was an apparent slowdown in both conditions. A developmental period in which greater experimentation takes precedence over increased production has been suggested by other studies (Littlemore et al., 2014, p. 128), although this appeared around CEFR level B2 (TOEIC: 785) in examination writing. In this study, it is possible that the greater level of linguistic support (in the form of dictionaries and course materials being available and there being a specific instruction to attempt to use new language) allowed for this experimentation period to become apparent at a noticeably earlier stage.

What might be happening during this experimental stage? Littlemore et al. (2014, p. 138 - 141) suggested that learners begin to use metaphor more creatively and consequently produce more erroneous forms. Although creativity was not a focus of this study, there was a finding that error rates can change with ability grouping depending on the construction under investigation; verbal constructions sometimes had higher error rates among low-level learners, while for nouns there was a suggestion that fluctuating error rates may have been due to avoidance on the part of low-level learners. It must be remembered, however, that the statistical support for this is limited due to the small sample sizes. The analyses in sections 5.6.2.3 (page 218) and 7.3 - 7.4 (pages 305, 314) lend support to the conjecture that metaphorical noun use is a marker of proficiency, and it is likely that the early stages of such use will be marked by a greater amount of errors. Another factor to consider would be lexical variety. It was found in the study that the higher the learners' ability, the greater the variety of target metaphors they used in their output. Lexical richness is a quality long considered indicative of more advanced composition

(Laufer & Nation, 1995, p. 307 - 308), and Figure 5.13 (page 189) implies that variety does increase broadly in line with ability, albeit with considerable variation at an individual level. Viewed in a different way, this finding could also imply that lower ability learners are overly dependent on certain metaphors in their writing, leading them to repeat the same language and increase quantity but not variety of use.

Productive lexical variety can therefore be considered in two ways; learners can gradually acquire various senses of individual words as well as broadening their range of production of words themselves. Thus it is important that lexical and metaphorical development not be seen only in terms of increased use of a particular feature, as limiting dependence on a particular expression in favour of other terms can be just as much a sign of improvement. Again, this reflects Ellis and Ferreira-Junior's (2009) findings on the progression of construction learning, with overused phrases being put aside in favour of other, more recently-learnt expressions.

Dependence on certain metaphors took several forms in the data. Most strikingly, learners in the control condition, and low proficiency individuals especially, made extensive use of metaphorical coherence-forming devices such as *in addition*, *on the other hand*, *and so on* and *in summary* (Appendix K, page 473; Appendix P, page 493). Why this pattern appeared in the data is unclear, as structuring metaphors were not an explicit feature of either condition's teaching. One speculative conclusion might be that control condition learners and those of lower proficiency were more likely to use such expressions as 'filler' in their writing, in lieu of other language features to prioritise.

Overuse could also be seen in learners' production of extremely high frequency forms (Figures 7.1 and 7.2, pages 307, 311). The words *big*, *thing*, and *have* were

significantly overused as metaphors by learners in both conditions relative to BAWE data. All of these are highly prototypical referents for the concepts they represent, which indicates that learners in this study generally lacked the diverse productive vocabularies of native speakers. This should come as no surprise, but it does serve as a reminder that learners would benefit from revisiting previously studied language features with the purpose of adding variety to their output.

Of course, it should also be remembered that the overuse of high frequency forms was far from uniform across all elements of language. In the case of adverbs, for example, learners greatly overused relatively low frequency expressions such as *and so on* and *on the other hand* while generally avoiding more frequent adverbs such as *far* and *here*. This suggests that other factors, including the salience of a given expression, will affect uptake.

Also with regard to overuse, the findings of section 7.6 raise the question of how much learner dependence on L2 forms might be influenced by the norms of their L1. The metaphor feedback surveys suggested that learners were making at least some reference to their first language in 60 - 75% of cases. This figure may be skewed somewhat, as the feedback surveys were only used for cases of interesting, innovative or unusual language, but they do imply that the L1 is very much present in learners' minds when they construct meaning. In the case of metaphorical verbs being used to describe culture(s), the analyses in section 7.6.2 (page 351) suggest that there is considerable overlap between Japanese and English in the metaphorical themes that conceptualise cultural change. At the level of individual collocates of *culture* or *bunka*, some examples were found where differences in frequency of use between the L1 and L2 might have influenced learners' lexical choices, such as the

infrequent use of *shape* in collocation with *culture*. Other cases of potential cross-linguistic interference were those where the actual lexical realisations of a shared metaphorical theme differed, as with learners' use of *grow up* and *bring up* as collocates of culture, although this might also be simply ascribed to incorrect or even unlucky dictionary use since the synonymous *foster* or *nurture* do collocate with *culture* in English. However, facilitative effects from the L1 were just as evident, as could be seen in the case of forms such as *spread* and the significant correlations between output word frequency and L1 and L2 frequency as determined by COCA and jpTenTen11. The finding that the correlations between output and L1/L2 frequency for collocations with metaphorical verbs and culture or *bunka* were weaker than those for individual words may actually support the frequency model. Since collocations are less common than their individual component words, learners may be less responsive to their presence in input.

Finally, phraseology is an important aspect of language that is closely related to metaphor. The majority of learners in this study had not yet reached the stage where we might expect them to be combining several chunks of language to structure discourse, as was the case with the subject of Bell's (2009, p. 125 - 126) study, and indeed the analysis found a preponderance of fixed phrases and simple constructions, but signs of development were also in evidence. Metaphorical adjective/noun collocations, as might be expected, were frequent in both conditions, and in the majority of cases, learners were able to produce attested uses of high frequency patterns. The main effect of the experimental treatment appears to have been an increase in the output of particular adjectives, especially *big* and *high*, rather than any great change in variety. Since the awareness-raising activities in the

metaphor workbook focussed only on high frequency adjectives, this may not be unexpected. It is also possible that A2 level learners have simply not yet reached the point at which they can exploit less frequent adjectives in their writing.

Collocation was also investigated for metaphorical verbs with the noun *culture*. Here, learners made a number of selections for verbs that were not strong or frequent collocate when judged against COCA, but which could be explained when consideration was given to the context of the classroom by using input corpora frequencies instead of COCA. This may reveal a limitation of using large scale corpora for baseline data - some of the target expressions were ideally suited to an anthropology course dealing with culture, but did not actually appear frequently in this way in general usage.

Colligation patterns were examined for the six most frequent metaphorical verbs (*evolve*, *take in*, *absorb*, *mix*, *spread* and *lose*) and [part] nouns, and the data for both patterns provided evidence of developmental stages that learners must work through. For metaphorical verbs, learner output usually adopted the most frequent construction as determined by COCA, and error rates were significantly higher when learners produced one of the less frequent constructions. Lower proficiency learners in the experimental condition were seen to be more likely to make errors than their more proficient classmates, with both grammatical and phraseological issues relatively common. The [part] nouns construction may be one that appears at a later stage of development - it was more frequently found in mid and high level learners' writing in both conditions, and the error rates showed no clear trend in line with ability. This reflects the claim earlier in this chapter that

metaphorical noun use may become more frequent in higher ability ranges, and that it will be more likely to contain error in the early stages of this development.

Learners are thought to gradually acquire an awareness of the 'chunked' nature of language (Ellis, 2003) through a process of trial, error, feedback and retrial, with patterns typically emerging around a core construction which is the most frequent of its type (Ellis & Ferreira-Junior, 2009). Usage of constructions will likely then ebb and flow as learners initially rely on trusted phrases and later acquire new ones to partially replace them. Studies into learner production of collocations have suggested that there is a tendency for high frequency forms to become overused relative to native speaker norms (Durrant & Schmitt, 2009, p. 174), which may be simply because learners are reluctant to let go of a safe island in their range of productive expression merely to bring variety to their output.

In summary, these findings suggest a general developmental process in which constructions begin to appear in learner output at certain levels of proficiency. Once a construction has been picked up for use, output will typically centre on the most prototypical form, and errors may be more frequent. Through sustained practice, however, accuracy should improve. At some point, learners may begin to experiment with the pattern, perhaps opting for other lexical items to use in the main slot or adapting the construction into a less prototypical form. The effect of instruction on this process is to foreground target constructions, and learners' response will depend on the state of their interlanguage at that time. It could be that the response is simple avoidance if the construction is perceived as difficult in some way, or learners might attempt to use the construction, but struggle to produce it accurately. Alternatively, if the construction is one that has already been acquired to some

extent, learners might be prompted to try out some variation on its form. Throughout this process, it is likely that feedback will play an important role in bringing inaccuracies or errors to learners' attention, although whether such feedback is attended to will depend on various factors, including the salience and specificity of the feedback, the medium in which it is presented, and other competing sources of input (Bitchener & Knoch, 2009; Bitchener, Young, & Cameron, 2005; Ruegg, 2015).

8.6. Implications of the study and suggestions for further research

One of the main objectives of this study has been to combine the rigour of an experimental research design with the authenticity of a classroom setting. It has been argued from both empirical and theoretical perspectives that cognitive linguistic principles reflect the true nature of language and language learning (Boers & Lindstromberg, 2008, p. 17 - 37; Langacker, 2008a), and that learners will benefit from increased awareness of the motivated aspects of language. Highly controlled studies carried out over usually short periods of time have shown positive results in this regard, but post-tests and investigations into learners' independent use of CL knowledge have been less convincing (Boers, 2004, p. 215 - 217; Condon, 2008, p. 150 - 151; Skoufaki, 2008, p. 118). The 'gold-standard' for CL research must be to demonstrate that its principles are beneficial when applied over the long term in courses whose goals are tied into institutional curricula and therefore cover more than just the focus of the CL treatment itself. In other words, we need to know whether CL approaches can be successfully integrated with established programs, and if so, what issues instructors might face and how they can be overcome.

This study has attempted to break new ground in this area by taking a CL-inspired approach to metaphor teaching into a CLIL course and capturing as much data as possible from the learning environment. In order to evaluate the effectiveness of metaphor awareness raising, a comprehensive picture of classroom occurrences was required. It is pleasing to note that several outcomes of the study were very positive - despite possible early hesitation to use metaphors, learners were increasingly willing to try out target expressions in their written responses to course content. This, it is argued, gave learners greater expressive freedom and allowed them to better engage with the themes of the course. Learners also showed signs of reducing their overuse of certain safe choices, particularly fixed phrase adverbs, and of shifting towards a greater variety of use. Other outcomes, rather than showing signs of development in learners, revealed aspects of development that may inspire further studies. These included the different patterns of use for taught and untaught target metaphors, the general trend for learners to make less use of metaphorical nouns than verbs, and the need to provide instruction for phraseological constructions in order to help encourage learners to adopt new language successfully. It is hoped that this study can serve as a starting point for other longitudinal investigations into learner development.

Regarding the practical concerns instructors will likely face when bringing metaphor into the classroom, several issues and limitations became evident during the study. The first was the problem of learner avoidance, a common issue when new language features are introduced. The approach taken here was to provide guided instruction in the form of controlled production activities and personalisation of target language as well as regular review of metaphors which were not appearing

in learner output. Coupled with this, it was repeatedly emphasised that the notebook writing would be graded based on learners' attempts to engage with course themes and to experiment with new language rather than on accuracy. This may have been an unusual assignment for learners, which would explain their taking some time to fully commit to experimentation. On reflection, a better design for the metaphor workbook activities may have been to ensure that the exercises for every metaphor theme required learners to practice writing their own sentences using the new forms so as to build confidence and to allow for feedback to be provided before learners attempted to use the language in free production with their reflective writing. From the point of view of research, if the ability to use a linguistic feature accurately in free production is seen as indicative of mastery, then it may be prudent to design studies that allow learners to demonstrate knowledge of metaphorical language in tasks that are less demanding. The fact that data was only collecting from a task with a high level of demand in this study may have masked developments in receptive knowledge or controlled productive capabilities.

In the same vein, when presenting target metaphorical forms to learners, instructors should take care to also raise awareness of appropriate collocations or other phraseological considerations. With hindsight, the materials developed for the metaphor workbook used in this study probably did not provide enough assistance in this area, which may have led to the errors reported in examples (54 - 58) of section 7.1 and was a possible cause of some learner avoidance of particular metaphors. Language learners, particularly those of lower proficiency, cannot be expected to process language much beyond the level of individual words, and the subtle differences of meaning created by collocations are likely to elude them. Of

course, there needs to be moderation in the amount of detail course materials provide, but without some sense of phraseological appropriacy, learners will not be able to access the intended meanings of metaphorical expressions.

A broader issue was that in general, learners appeared to be steered towards only language that had been explicitly taught in class, rather than learning to apply language from metaphorical themes independently. It must be said that the findings of chapter six run counter to this somewhat (Figure 6.9 and Table 6.22, pages 269-270), as there was evidence that untaught target metaphor use was increasing among metaphor themes relating directly to culture. However, the general trend was less positive. Certainly, fostering the ability to apply learning independently is a desirable outcome of language studies, and this should be considered in line with research into learner autonomy. This raises a question regarding the degree to which metaphor itself was influencing learners' output; were lexical selections based on a conscious awareness of a particular metaphor theme, or were did learners merely use the vocabulary items that they had been taught, irrespective of their metaphorical status? This question motivated the distinction between taught and untaught target metaphors (section 3.6.2.1, page 112), but as was stated above, the results were mixed. Experimental group learners were producing more untaught target metaphors related to culture as the semester progressed, but at no point did they produce significantly more of these metaphors than the control group. An expanded study with a greater number of participants focusing on fewer target metaphors may help to clarify this situation. Furthermore, the examples of *grow up* and *bring up* as collocates of *culture* cited earlier demonstrate that even if knowledge of a metaphor theme, or conceptual metaphor, is assumed, this does not

entail an understanding of linguistic norms. This is what Holme (2004, p. 97) suggests when he states, "a conceptual metaphor schema is not a facet of the rules governing the language that we produce but a feature of how we conceptualise and interpret the meanings that language will utilise." Philip (2005) picks up this theme, pointing out that receptive knowledge of metaphor is vastly different from an awareness of how to produce it, as conventional metaphor usage cannot easily be separated from phraseology, grammar, and collocation. And unfortunately, just as receptive knowledge of a metaphor theme does not explain productive language use, examples of language production cannot offer any definitive proof of receptive knowledge of metaphor.

Future studies to overcome these limitations would need to incorporate other methods of data collection that might be able to access learners' awareness of metaphor itself. One possibility would be to conduct interviews with learners about the motivations behind their lexical choices, but in in the case of conventional metaphor, learners may not necessarily be aware of an underlying metaphor theme. An alternative would be to develop tasks that assess metaphoric awareness across a continuum of knowledge. For example, in vocabulary assessment, one model of word knowledge posits a scale that encompasses both receptive and productive (or passive and active) knowledge as well as a distinction between recognition (the ability to correctly identify a word or meaning when it is present) and recall (the ability to produce a word or meaning by oneself) knowledge (Laufer & Goldstein, 2004). If assessment tools covering a similar range of metaphor knowledge could be developed, and assuming the study were designed in such a way to prevent a testing

effect influencing later performance, stronger claims could be made for the degree to which metaphorical awareness develops in response to explicit instruction.

Clearly, consideration should also be given to providing a balanced set of vocabulary development activities in individual courses and across broader programs. It has been argued in this study that a metaphor awareness-raising approach can benefit learners by encouraging them to draw on their lexical resources in more diverse, and perhaps more contextually-appropriate, ways, but this is not to overlook the merits of the frequency-based approach to developing breadth of vocabulary knowledge (Nation, 2013, p. 11 - 35). Another limitation of the analyses presented in this study is that comparisons have only been made between the two groups' use of metaphorical vocabulary; it is of course possible that the control group learners benefitted in ways that the experimental group did not through the greater amount of instruction they received on high frequency or academic vocabulary. This is something that should definitely be examined in future studies and in other contexts. To return to the point raised in chapters one and two (pages 1, 8), however, it is not that learners simply need to know a lot of words (although that is a formidable enough challenge), but that they also need to know about the range of meaning those words can encompass, so an argument can be made for fostering an interest and an analytical approach to vocabulary learning (Nation, 2008, p. 172 - 173).

Another implication raised by the study was simply that in teaching discipline-specific metaphorical themes, a less-is-more approach may be of more benefit than attempting to provide comprehensive coverage of all themes. Given that metaphoric awareness involves an understanding of phraseological patterning as well as the

learning of new senses of word meaning, it could be argued that learners require extended practice with the most prominent metaphors so as to allow for more feedback and revision and to boost the salience of target language. A similar study to this one that focussed only on a small group of metaphors and provided qualitative insights into learners' thought processes as they attempt to incorporate these new expressions into their output would be hugely informative.

This study has also taken an ambitious approach to data collection by creating corpora that reflect to the greatest possible degree classroom language exposure and learner output over time. This was carried out with the goal of furthering understanding of how learners go about adding to their productive lexica. Ellis's notion of contingency learning (2006a, 2006b) was adopted as a theory of acquisition that took into account the complexities of exposure and noticing of language forms. This allowed for concepts such as frequency, dispersion and salience to be investigated with regard to the production of metaphorical vocabulary. These are particularly difficult concepts to measure outside of controlled experimental designs, but the input corpora allowed for them to be operationalized and for their effects to be estimated. This was a valuable opportunity, as other studies into learners' longitudinal development have had to rely solely on output corpora to investigate learner development (e.g., Bell, 2009; Ellis & Ferreira-Junior, 2009). One way in which this study might be improved would be to narrow the focus to just a single learner or a small group of learners, and to attempt to record their language production during each class and their exposure to language from other learners as well as the course instructors. This would provide a comprehensive record of classroom input and output, which, if accompanied by detailed feedback or a range

of metaphor assessment instruments, might enable a better understanding of how input and learners' own thought processes lead to output.

Lastly, with regard to vocabulary learning, this study has demonstrated the various ways that lexical development is manifested in learners' written production. In doing so, it has argued for a principled approach to addressing knowledge of vocabulary depth through the teaching of metaphor in a CLIL environment. While such an approach is not intended to replace instruction for new lexical items, the results from the experimental treatment suggest that there is merit in addressing the range of meanings lexical items can cover, and that learners are able to add variety to their output when provided with focussed instruction into prominent metaphorical themes. In general terms, the study has highlighted the value of returning to previously studied language (in this case, vocabulary) and then expanding and improving on this knowledge. There can be a tendency for textbooks, curricula and research to prioritise new elements of language and quantity of production over review and revision, but the findings of this study suggest that there is also clear merit in considering what has come before and then adding variety and subtlety to learners' range of expression.

LIST OF REFERENCES

Aizawa, K. (2006). Rethinking frequency markers for English-Japanese dictionaries. In M. Murata, K. Minamide, Y. Tono, & S. Ishikawa (Eds.), *English lexicography in Japan* (pp. 108 - 119). Tokyo: Taishukan Publishing Company.

Allwright, D. (2005). Developing principles for practitioner research: The case of exploratory practice. *The Modern Language Journal*, *89*(3), 353 - 366.

Alsop, S. & Nesi, H. (2009). Issues in the development of the British Academic Written English (BAWE) corpus. *Corpora, 4*(1), 71 - 83.

Bachman, L. F. (1990). *Fundamental considerations in language testing.* Oxford: Oxford University Press.

Baddeley, A. (1997). *Human memory: Theory and practice (Revised edition)*. Hove: Psychology Press Ltd.

Bamgbose, A. (1998). Torn between the norms: Innovations in world Englishes. *World Englishes 17*(1), 1 - 14.

Barcelona, A. (2007). The role of metonymy in meaning construction at discourse level: A case study. In G. Radden, K. Köpcke, T. Berg, & P. Siemund (Eds.) *Aspects of meaning construction* (pp. 51 - 75). Amsterdam: John Benjamins Publishing Company.

Barcroft, J. (2004). Effects of sentence writing in second language lexical acquisition. *Second Language Research, 20*(4), 303 - 334.

Barcroft, J. (2006). Can writing a new word detract from learning it? More negative effects of forced output during vocabulary learning. *Second Language Research, 22*(4), 487 - 497.

Batty, A. (2012). Identifying Dimensions of Vocabulary Knowledge in the Word Associates Test. *Vocabulary Learning and Instruction, 1*(1), 70 - 77.

BAWE. (n.d.) British Academic Written English Corpus. Retrieved from: http://www.coventry.ac.uk/research/research-directories/current-projects/2015/british-academic-written-english-corpus-bawe/

Beglar, D. (2009). A Rasch-based validation of the Vocabulary Size Test. *Language Testing, 26*(4), 1 - 22.

Bell, H. (2009). The messy little details: A longitudinal case study of the emerging lexicon. In T. Fitzpatrick, & A. Barfield (Eds.), *Lexical processing in second language learners* (pp. 111 - 127). Bristol: Multilingual Matters.

Beréndi, M., Csábi, S., & Kövecses, Z. (2008). Using conceptual metaphors and metonymies in vocabulary teaching. In F. Boers, & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 65 - 99). Berlin: Mouton de Gruyter.

Biber, D. & Conrad, S. (2009). *Register, genre, and style*. Cambridge: Cambridge University Press.

Bitchener, J., & Knoch, U. (2009). The value of a focused approach to written corrective feedback. *ELT Journal 63*(3), 204 - 211.

Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing 14*(3), 191 - 205.

Boers, F. (2000a). Enhancing metaphoric awareness in specialised reading. *English for Specific Purposes, 19*(2), 137 - 147.

Boers, F. (2000b). Metaphor awareness and vocabulary retention. *Applied Linguistics, 21*(4), 553 - 571.

Boers, F. (2003). Applied linguistic perspectives on cross-cultural variation in conceptual metaphor. *Metaphor and Symbol, 18*(4), 231 - 238.

Boers, F. (2004). Expanding learners' vocabulary through metaphor awareness: What expansion, what learners, what vocabulary? In M. Achard, & S. Niemeier (Eds.), *Cognitive linguistics, second language acquisition, and foreign language teaching* (pp. 211 - 232). Berlin: Mouton de Gruyter.

Boers, F., & Demecheleer, M. (2001). Measuring the impact of cross-cultural differences on learners' comprehension of imageable idioms. *ELT Journal, 55*(3), 255 - 262.

Boers, F., Demecheleer, M., & Eyckmans, J. (2004). Cross-cultural variation as a variable in comprehending and remembering figurative idioms. *European Journal of English Studies, 8*(3), 375 - 388.

Boers, F., Eyckmans, J., & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research, 11*(1), 43 - 62.

Boers, F., & Lindstromberg, S. (2008). How cognitive linguistics can foster effective vocabulary teaching. In F. Boers, & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 1 - 61). Berlin: Mouton de Gruyter.

Boers, F., Lindstromberg, S., Littlemore, J., Stengers, H., & Eyckmans, J. (2008). Variables in the mnemonic effectiveness of pictorial elucidation. In F. Boers, & S.

Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 189 - 216). Berlin: Mouton de Gruyter.

Boers, F., & Littlemore, J. (2000). Cognitive style variables in participants' explanations of conceptual metaphors. *Metaphor and Symbol, 15*(3), 177 - 187.

Boers, F., Piquer Píriz, A. M., Stengers, H., & Eyckmans, J. (2009). Does pictorial elucidation foster recollection of idioms. *Language Teaching Research, 13*(4), 367 - 382.

Boers, F., & Stengers, H. (2008). A quantitative comparison of the English and Spanish repertoires of figurative idioms. In F. Boers, & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 355 - 374). Berlin: Mouton de Gruyter.

Bovair S. & Kieras D. E. (1981). A guide to propositional analysis for research on technical prose. Technical report No. 8. University of Arizona. Department of Psychology.

Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review, 112*(1), 193 - 216.

Brown, D. (2012). The frequency model of vocabulary learning and Japanese learners. *Vocabulary Learning and Instruction, 1*(1), 20 - 28.

Brown, H., & Iyobe, B. (2014). The growth of English medium instruction in Japan. In N. Sonda, & A. Krause (Eds.), *JALT2013 Conference Proceedings.* Tokyo: JALT.

Brown, J. D. (2001). *Using surveys in language programs*. Cambridge University Press.

Burton, R. F. (2001). Quantifying the effects of chance in multiple choice and true/false tests: Question selection and guessing of answers. *Assessment & Evaluation in Higher Education*, *26*(1), 41 - 50.

Cambridge English. (2017) Cambridge English Language Assessment. Retrieved May 1, 2017, from http://www.cambridgeenglish.org/exams/general-and-business-english.

Cambridge University Press. (2015). Cambridge English Corpus. Retrieved April 5, 2017, from http://www.englishprofile.org/cambridge-english-corpus

Cameron, L. (2003). *Metaphor in educational discourse.* London: Continuum.

Cameron, L., & Deignan, A. (2006). The emergence of metaphor in discourse. *Applied Linguistics, 27*(4), 671 - 690.

Caspi, T., & Lowie, W. (2010). A dynamic systems perspective on L2 lexical development in academic English. In R. Chacón-Beltrán, C. Abello-Contesse, & M. Torreblanca-López (Eds.), *Insights into non-native vocabulary teaching and learning* (pp. 41 - 58). Bristol: Multilingual Matters.

Chapetón, C. M. (2010). Metaphor identification in EFL argumentative writing: A corpus-driven study. *Folios, 32*, 125 - 140.

Chapman, M. (2006, April). *An over-reliance on discrete item testing in the Japanese business context*. Paper presented at the International Conference in English Instruction and Assessment, National Chung Cheng University, Taiwan.

Charteris-Black, J. (2000). Metaphor and vocabulary teaching in ESP economics. *English for Specific Purposes, 19*(2), 149 - 165.

Charteris-Black, J. (2004). *Corpus approaches to critical metaphor analysis.* Basingstoke: Palgrave Macmillan.

Chen, Y., & Lai, H. (2013). Teaching English idioms as metaphors through cognitive-oriented methods: A case in an EFL writing class. *English Language Teaching, 6*(6), 13 - 20.

Childs, M. (1995). Good and bad uses of TOEIC by Japanese companies. In J. D. Brown & S. O. Yamashita (Eds.), *Language Testing in Japan* (pp. 66 - 75). Tokyo, Japan: The Japan Association for Language Testing.

Chujo, K. & Nishigaki, C. (2003). Bridging the vocabulary gap: From EGP to EAP. *JACET Daigaku eigo kyōiku gakkai kiyō*, (37), 73 - 84.

Chujo, K. & Oghigian, K. (2009). How many words do you need to know to understand TOEIC, TOEFL & EIKEN? An examination of text coverage and high frequency vocabulary. *The Journal of Asia TEFL*, *6*(2), 121 - 148.

Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System 25*(3), 301 - 315.

Cobb, T., & Horst, M. E. (1999). Vocabulary sizes of some City University students. *Journal of the Division of Language Studies of City University of Hong Kong, 1*(1), 59 - 68.

Condon, N. (2008). How cognitive linguistic motivations influence the learning of phrasal verbs. In F. Boers, & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 133 - 158). Berlin: Mouton de Gruyter.

Conklin, K. & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics, 32*, 45 - 61.

Cooper, T. C. (1999). Processing of idioms by L2 learners of English. *TESOL Quarterly*, *33*(2), 233 - 262.

Council of Europe. (2001). *Common European framework of reference for languages: learning, teaching assessment*. Cambridge: Cambridge University Press.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213 - 238.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671 - 684.

Crookes, G. (1991). Action research for second language teachers: It's not just teacher research. *University of Hawai'i Working Papers in ESL*, *10*(2), 73 - 90.

Crossley, S., Salsbury, T., & McNamara, D. (2010). The development of polysemy and frequency use in English second language speakers. *Language Learning, 60*(3), 573 - 605.

Crossley, S. A., Subtirelu, N., & Salsbury, T. (2013). Frequency effects or context effects in second language word learning: What predicts early lexical production? *Studies in Second Language Acquisition 35*(4), 727 - 755.

Danesi, M. (1994). Recent research on metaphor and the teaching of Italian. *Italica 71*(4), 453 - 464.

Danesi, M. (1995). Learning and teaching languages: The role of “conceptual fluency”. *International Journal of Applied Linguistics 5*(1), 3 - 20.

Daulton, F. E. (2008). *Japan's built-in lexicon of English-based loanwords.* Clevedon: Multilingual Matters.

de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism, 10*(1), 7 - 21.

de Groot, A. M. B. & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning 50*(1), 1 - 56.

Dearden, J. (2014). *English as a medium of instruction - a growing global phenomenon.* British Council. Oxford: University of Oxford.

Deignan, A. (2003). Metaphorical expressions and culture: An indirect link. *Metaphor and Symbol, 18*(4), 255 - 271.

Deignan, A. (2005). *Metaphor and corpus linguistics.* Amsterdam: John Benjamins Publishing Company.

Deignan, A. (2006). The grammar of linguistic metaphors. In A. Stefanowitsch, & S. T. Gries (Eds.), *Corpus-based approaches to metaphor and metonymy* (pp. 106 - 122). Berlin: Mouton de Gruyter.

Deignan, A., Gabryś, D., & Solska, A. (1997). Teaching English metaphors using cross-linguistic awareness-raising activities. *ELT Journal, 51*(4), 352 - 360.

Deignan, A., Littlemore, J., & Semino, E. (2013). *Figurative language, genre and register.* Cambridge: Cambridge University Press.

Deignan, A., & Potter, L. (2004). A corpus study of metaphors and metonyms in English and Italian. *Journal of Pragmatics, 36*(7), 1231 - 1252.

Dekeyser, R. M. (1997). Beyond explicit learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition 19*(2), 195 - 221.

Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. New York: Macmillan. Retrieved November 22, 2016, from https://archive.org/details/democracyeducati1916dewe

Díaz-Negrillo, A. & Fernández-Domínguez, J. (2006). Error tagging systems for learner corpora. *Revista española de lingüística aplicada 19*, 83 - 102.

Dóczi, B. & Kormos, J. (2016). *Longitudinal developments in vocabulary knowledge and lexical organization*. New York: Oxford University Press.

Dörnyei, Z. (1994). Motivation and motivating in the foreign language classroom. *The Modern Language Journal*, *78*(3), 273 - 284.

Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. Lawrence Erlbaum Associates.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition.* New York: Lawrence Erlbaum Associates.

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching, 47*(2), 157 - 177.

Dziemianko, A. (2010). Paper or electronic? The role of dictionary form in language reception, production and the retention of meaning and collocations. *International Journal of Lexicography 23*(3), 257 - 273.

Edge, J. (2001). Attitude and access: Building a new teaching/learning community in TESOL. In J. Edge (Ed.) *Action research: Case studies in TESOL practice series* (pp. 1 - 11). Arlington: Kirby Lithographic Company, Inc.

Elliot, J. (2004). The struggle to redefine the relationship between 'knowledge' and 'action' in the academy: Some reflections on action research. *Educar* (30), 11 - 26.

Ellis, N. C. (2002). Reflections on frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*(2), 143 - 188.

Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. J. Doughty, & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 33 - 68). Oxford: Blackwell.

Ellis, N. C. (2006a). Language acquisition as rational contingency learning. *Applied Linguistics, 27*(1), 1 - 24.

Ellis, N. C. (2006b). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*(2), 164 - 194.

Ellis, N. C. & Beaton, A. (1993). Psycholinguistic determinants of foreign language learning. *Language Learning*, *43*(4), 559 - 617.

Ellis, N. C., & Ferreira-Junior, F. (2009). Constructions and their acquisition: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics, 7*(1), 187 - 220.

ETS (2014). *2013 Report on test takers worldwide: The TOEIC® listening and reading test*. Retrieved December 27, 2016, from http://f1.hjfile.cn/file/201408/2013%20Report%20on%20Test%20Takers%20Worldwide.pdf

ETS (2015). *2014 Report on test takers worldwide: The TOEIC® listening and reading test*. Retrieved December 27, 2016, from http://pdfstori.com/7937+84543_cover-june-16-2015-156-pm-2014-report-on-test-takers-worldwide-the-toeic-listening

ETS (2016). *The TOEIC® Tests - the Global Standard for Assessing English Proficiency for Business*. Retrieved May 22, 2017, from: https://www.ets.org/toeic/succeed

ETS (2017). *Compare TOEFL® scores*. Retrieved May 22, 2017, from: https://www.ets.org/toefl/institutions/scores/compare/

Fauconnier, G., & Turner, M. (2003). *The way we think: Conceptual blending and the mind's hidden complexities.* New York: Basic Books.

Firth, J. R. (1957). A synopsis of linguistic theory, 1930 - 1955. In C. E. Bazell (Ed.), *Studies in linguistic analysis* (pp. 1 - 32). Oxford: Basil Blackwell.

Fitzpatrick, T. (2009). Word association profiles in a first and second language: Puzzles and problems. In T. Fitzpatrick, & A. Barfield (Eds.), *Lexical processing in second language learners* (pp. 38 - 52). Bristol: Multilingual Matters.

Gardner, D. & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics 35*(3), 305 - 327.

Gibbs, R. (2013). Metaphoric cognition as social activity. *Metaphor and the Social World, 3*(1), 54 - 76.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics, 8*(3), 183 - 206.

Grady, J. E. (1997). Theories are buildings revisited. *Cognitive Linguistics, 8*(4), 267 - 290.

Granger, S. (1998). The computer learner corpus: A versatile new source of data for SLA research. In S. Granger (Ed.), *Learner English on computer* (pp. 3 - 18). London: Routledge.

Granger, S. (2002). A bird's eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.) *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 3 - 33). Amsterdam & Philadelphia: Benjamins.

Granger, S. (2003). The international corpus of learner English: A new resource for foreign language learning and teaching and second language acquisition research. *TESOL Quarterly, 37*(3), 538 - 546.

Granger, S., & Rayson, P. (1998). Automatic profiling of learner texts. In S. Granger (Ed.), *Learner English on computer* (pp. 119 - 131). London: Routledge.

Groom, N. (2009). Effects of second language immersion on second language collocational development. In A. Barfield, & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 21 - 33). Basingstoke: Palgrave Macmillan.

Hanks, P. (2013). *Lexical analysis: Norms and exploitations.* Cambridge, Massachusetts: The MIT Press.

Hasselgren, A. (1994). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied of Applied Linguistics, 4*(2), 237 - 260.

Hemchua, S., & Schmitt, N. (2006). An analysis of lexical errors in the English compositions of Thai learners. *Prospect, 21*(3), 3 - 25.

Higginbotham, G., Munby, I., & Racine, J. P. (2015). A Japanese word association database of English [Online article]. *Vocabulary Learning and Instruction*. Retrieved November 5, 2015, from http://vli-journal.org/wp/online-first/

Hirai, M. (2002). Correlations between active skill and passive skill test scores. *Shiken: JALT Testing & Evaluation SIG Newsletter*, *6*(3), 2 - 8.

Hoàng, T. Đ. (2015). *Metaphorical language in second language learners' essays: Products and processes* (Doctoral dissertation, Victoria University of Wellington). Retrieved April 24, 2015, from http://researcharchive.vuw.ac.nz/xmlui/bitstream/handle/10063/4195/thesis.pdf?sequence=2.

Hoey, M. (2005). *Lexical Priming.* Abingdon: Routledge.

Holme, R. (2004). *Mind, metaphor and language teaching.* Basingstoke: Palgrave Macmillan.

Hulstijn, J. H. (2001). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 258 - 286). Cambridge: Cambridge University Press.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*(3), 539 - 558.

Hunston, S., & Francis, G. (1999). *Pattern grammar: A corpus-driven approach to the lexical grammar of English.* Amsterdam: John Benjamins Publishing Company.

Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics, 32*, 150 - 169.

Izumi, E., Uchimoto, K., & Isahara, H. (2005, October). *Error annotation for corpus of Japanese learner English*. Retrieved April 14, 2017, from http://acl-arc.comp.nus.edu.sg/archives/acl-arc-090501d3/data/pdf/anthology-PDF/I/I05/I05-6009.pdf

Jakubíček, M., Kilgariff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013). *The TenTen corpus family.* Retrieved February 24, 2014, from http://www.sketchengine.co.uk/documentation/ wiki/Corpora/TenTen

Johnson, J., & Rosano, T. (1993). Relation of cognitive style to metaphor interpretation and second language proficiency. *Applied Psycholinguistics, 14*(2), 159 - 175.

Johnson, K. E. (2006). The sociocultural turn and its challenges for second language teacher education. *TESOL Quarterly*, *40*(1), 235 - 257.

Kachru, B. B. (1991). World Englishes and applied linguistics. In M. L. Tickoo (Ed.) *Languages and standards: Issues, attitudes, case studies* (pp. 178 - 205). SEAMEO Regional Language Centre: Singapore.

Kachru, B. B. (1992). World Englishes: Approaches, issues and resources. *Language Learning 25*(1), 1 - 14.

Kachru, B. B. & Nelson, C. L. (2001). World Englishes. In A. Burns & C. Coffin (Eds.) *Analysing English in a global context: A reader* (pp. 9 - 25). Routledge: London.

Lado, R. (1957). *Linguistics Across Cultures: Applied Linguistics for Language Teachers.* Ann Arbor: The University of Michigan Press.

Lakoff, G. (1990). The invariance hypothesis: Is abstract reason based on image-schemas? *Cognitive Linguistics, 1*(1), 39 - 74.

Lakoff, G., Espenson, J., & Schwartz, A. (1991). *Master metaphor list (2nd edition).* Retrieved February 4, 2013, from http://araw.mede.uic.edu/~alansz/metaphor/METAPHORLIST.pdf

Lakoff, G., & Johnson, M. (2003). *Metaphors we live by (Second edition).* Chicago: University of Chicago Press.

Langacker, R. W. (1987). *Foundations of cognitive grammar (Vol. 1: Theoretical prerequisites).* Stanford: Stanford University Press.

Langacker, R. W. (1991). *Foundations of cognitive grammar (Vol. 2: Descriptive application).* Stanford: Stanford University Press.

Langacker, R. W. (2008a). Cognitive grammar as a basis for language instruction. In P. Robinson, & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 66 - 88). New York: Routledge.

Langacker, R. W. (2008b). *Cognitive grammar: A basic introduction.* New York: Oxford University Press.

Langacker, R. W. (2009). *Cognitive Linguistics Research* (42): *Investigations in cognitive grammar.* Berlin: Mouton de Gruyter.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics, 27*(4), 590 - 619.

Laufer, B. (1997). What's in a word that makes it hard or easy: Some intralexical factors that affect the learning of words. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 140 - 155). Cambridge: Cambridge University Press.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics, 19*(2), 255 - 271.

Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook, 5*, 223 - 250.

Laufer, B. (2006). Comparing focus on form and focus on forms in second-language vocabulary learning. *The Canadian Modern Language Review 63*(1), 149 - 166.

Laufer, B., & Eliasson, S. (1993). What causes avoidance in L2 learning: L1-L2 difference, L1-L2 similarity, or L2 complexity. *Studies in Second Language Acquisition, 15*(1), 35 - 48.

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics, 29*(4), 694 - 716.

Laufer, B. & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning 54*(3), 399 - 436.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*(3), 307 - 322.

Laufer, B., & Rozovski-Roitblat, B. (2015). Retention of new words: Quantity of encounters, quality of task, and degree of knowledge. *Language Teaching Research*, *19*(6), 687 - 711.

Lazar, G. (2003). *Meanings and metaphors: Activities to practise figurative language.* Cambridge: Cambridge University Press.

Lindstromberg, S. (2010). *English prepositions explained*. Amsterdam: John Benjamins Publishing Company.

Littlemore, J. (2001). Metaphoric competence: A language learning strength of students with a holistic cognitive style? *TESOL Quarterly, 35*(3), 459 - 491.

Littlemore, J. (2003). The effect of cultural background on metaphor interpretation. *Metaphor and Symbol, 18*(4), 273 - 288.

Littlemore, J. (2004). *Applying cognitive linguistics to second language learning and teaching.* Basingstoke: Palgrave Macmillan.

Littlemore, J. (2015). *Metonymy: Hidden shortcuts in language, thought and communication*. Cambridge: Cambridge University Press.

Littlemore, J., Krennmayr, T., Turner, J., & Turner, S. (2014). An investigation into metaphor use at different levels of second language writing. *Applied Linguistics, 35*(2), 117 - 144.

Littlemore, J., & Low, G. (2006a). *Figurative thinking and foreign language learning.* Basingstoke: Palgrave Macmillan.

Littlemore, J., & Low, G. (2006b). Metaphoric competence, second language learning, and communicative language ability. *Applied Linguistics, 27*(2), 268 - 294.

Littlemore, J., Trautman Chen, P., Koester, A., & Barnden, J. (2011). Difficulties in metaphor comprehension faced by international students whose first language is not English. *Applied Linguistics, 32*(4), 408 - 429.

Low, G. (1988). On teaching metaphor. *Applied Linguistics, 9*(2), 125 - 147.

Low, G., Littlemore, J., & Koester, A. (2008). Metaphor use in three UK university lectures. *Applied Linguistics, 29*(3), 428 - 455.

MacArthur, F. (2010). Metaphorical competence in EFL: Where are we and where should we be going? A view from the language classroom. *AILA Review, 23*, 155 - 173.

MacArthur, F., & Littlemore, J. (2008). A discovery approach to figurative language learning with the use of corpora. In F. Boers, & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 159 - 188). Berlin: Mouton de Gruyter.

Macis, M. & Schmitt, N. (2017). Not just 'small potatoes': Knowledge of the idiomatic meanings of collocations. *Language Teaching Research 21*(3), 321 - 340.

Mahon, J. E. (1999). Getting your sources right: What Aristotle didn't say. In L. Cameron, & G. Low (Eds.), *Researching and Applying Metaphor* (pp. 69 - 80). Cambridge: Cambridge University Press.

Martinez, R. & Murphy, V. A. (2011). Effect of frequency and idiomaticity on second language reading comprehension. *TESOL Quarterly*, *45*(2), 267 - 290.

McCrostie, J. (2010). The TOEIC® in Japan: A scandal made in heaven. *Shiken: JALT Testing & Evaluation SIG Newsletter*, *14*(1), 2 - 10.

McDonough, J. & McDonough, S. (1997). *Research methods for English language teachers*. London: Arnold.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London: Routledge.

McTaggart, R. (1994). Participatory action research: Issues in theory and practice. *Educational Action Research*, *2*(3), 313 - 337.

Meara, P. (2009). *Connected words: Word associations and second language vocabulary acquisition.* Amsterdam: John Benjamins Publishing Company.

Meara, P. (2010). *EFL vocabulary tests (Second edition)*. Retrieved September 6, 2011, from http://www.lognostics.co.uk/vlibrary/meara1992z.pdf

Meara, P., & Wolter, B. (2004). V_Links: Beyond vocabulary depth. *Angles on the English-speaking world, 4*, 85 - 96.

Mele, A. R. (1997). Real self-deception. *Behavioral and Brain Sciences*, *20*(1), 91 - 136.

Messick, S. (1988). *Validity*. Educational Testing Service Research Report. Retrieved November 30, 2016, from http://onlinelibrary.wiley.com/doi/10.1002/j.2330-8516.1987.tb00244.x/epdf

MEXT. (n.d.). *About Global 30*. Retrieved October 13, 2015, from http://www.uni.international.mext.go.jp/global30/

MEXT. (2014). *Selection for the FY 2014 top global university project.* Retrieved October 13, 2015, from http://www.mext.go.jp/b_menu/houdou/26/09/__icsFiles/afieldfile/2014/10/07/1352218_02.pdf

Miller, G. A. (1993). Images and models, similes and metaphors. In A. Ortony (Ed.), *Metaphor and thought (Second edition)* (pp. 357 - 400). New York: Cambridge University Press.

Milton, J. (2009). *Measuring second language vocabulary acquisition.* Bristol: Multilingual Matters.

Milton, J. (2010). The development of vocabulary breadth across the CEFR levels: A common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, and textbooks across Europe. *Eurosla Monographs Series I, 1*(1), 211 - 232.

Moon, R. (1997). Vocabulary connections: Multi-word items in English. In N. Schmitt, & M. McCarthy (Eds.) *Vocabulary: Description, acquisition and pedagogy* (pp. 40 - 63). Cambridge: Cambridge University Press

Murai, S. (2016, January 25). Changes in store for TOEIC, but test still not total gauge of fluency: Experts. The Japan Times. Retrieved December 27, 2016, from: http://www.japantimes.co.jp/news/2016/01/25/reference/changes-store-toeic-test-still-not-total-gauge-fluency-experts/#.WKMMIxhh2Wa

Na, L., & Nation, I. S. (1985). Factors affecting guessing vocabulary in context. *RELC Journal, 16*(1), 33 - 42.

Nacey, S. (2013). *Metaphors in learner English.* Amsterdam: John Benjamins Publishing Company.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review 63*(1), 59 - 81.

Nation, I. S. P. (2008). Lexical awareness in second language learning. In J. Cenoz & N. H. Hornberger (Eds.), *Encyclopedia of language and education. (Second edition)*

Volume 6: Knowledge about language (pp. 167 - 177). New York: Springer Science.

Nation, I. S. P. (2013). *Learning vocabulary in another language. (Second edition)*. Cambridge: Cambridge University Press.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*(2), 223 - 242.

Nunan, D. (1991). *Language teaching methodology: A textbook for teachers*. Hemel Hempstead: Prentice Hall International.

Nurweni, A., & Read, J. (1999). The English vocabulary knowledge of Indonesian university students. *English for Specific Purposes, 18*(2), 161 - 175.

Ozturk, M. (2015). Vocabulary growth of the advanced EFL learner. *The Language Learning Journal, 43*(1), 94 - 109.

Ozturk, M. (2016). Second language vocabulary growth at advanced level. *The Language Learning Journal*, 44(1), 6 - 16.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76*(3), 241 - 263.

Palmer, H. E. (1917). *The Scientific Study and Teaching of Languages.* Yonkers-on-Hudson: World Book Company.

Panther, K. & Thornburg, L. L. (2004). The role of conceptual metonymy in meaning construction. *Metaphorik.de*, (6), 91 - 116.

Pearson. (n.d.). *Longman Corpus Network*. Retrieved April 5, 2017, from http://www.pearsonlongman.com/dictionaries/corpus/learners.html

Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: An eye tracking study. *Studies in Second Language Acquisition 38*(1), 97 - 130.

Perfetti, C. & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189 - 213). Amsterdam: John Benjamins Publishing Company.

Philip, G. (2005). From concept to wording and back again: Features of learners' production of figurative language. Retrieved March 15, 2014, from https://www.academia.edu/1090320/_From_Concept_to_Wording_and_Back_Again_Features_of_learners_production_of_figurative_language

Philip, G. (2006). Metaphor, the dictionary, and the advanced learner. Retrieved March 15, 2014, from http://amsacta.unibo.it/2062/

Philip, G. (2007). Decomposition and delexicalisation in learners' collocational (mis)behaviour. Retrieved March 15, 2014, from http://ucrel.lancs.ac.uk/publications/cl2007/paper/170_Paper.pdf

Philip, G. (2010). "Drugs, traffic, and many other dirty interests": Metaphor and the language learner. Retrieved August 2, 2013, from http://amsacta.unibo.it/2125/1/RaAM06.pdf

Piersman, Y. & Geeraerts, D. (2006). Metonymy as a prototypical category. *Cognitive Linguistics 17*(3), 269 - 316.

Prabhu, N. S. (1995). Concept and conduct in language pedagogy. In G. Cook & B. Seidlhofer (Eds.) *Principle & Practice in applied linguistics: Studies in honour of H. G. Widdowson*. (pp. 57 - 71). Oxford: Oxford University Press.

Pragglejaz Group (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, *22*(1), 1 - 39.

Prawat, R. S. (1991). Conversations with self and settings: A Framework for thinking about teacher empowerment. *American Educational Research Journal*, *28*(4), 737 - 757.

Radden, G. (2002). How metonymic are metaphors? In R. Dirven & R. Pörings (Eds.) *Metaphor and metonymy in comparison and contrast* (pp. 407 - 434). Berlin: Mouton de Gruyter.

Radden, G., & Dirven, R. (2007). *Cognitive English grammar.* Amsterdam: John Benjamins Publishing Company.

Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics, 13*(4), 519 - 549.

Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards, & B. Laufer (Eds.), *Vocabulary in a second language* (pp. 209 - 227). Amsterdam: John Benjamins Publishing.

Riding, R., & Cheema, I. (1991). Cognitive styles: An overview and integration. *Educational Psychology, 11*(3 - 4), 193 - 215.

Rodgers, T. S. (1969). On measuring vocabulary difficulty: An analysis of item variables in learning Russian-English vocabulary pairs. *IRAL 7*(4), 327 - 343.

Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Acquisition, 21*(4), 589 - 619.

Ruegg, R. (2015). Differences in the uptake of peer and teacher feedback. *RELC Journal 46*(2), 131 - 145.

Ryan, R. M. & Deci, E. L. (2000). Self-determination theory and facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68 - 78.

Salsbury, T., Crossley, S. A., & McNamara, D. S. (2011). Psycholinguistic word information in second language oral discourse. *Second Language Research 27*(3), 343 - 360.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129 - 158.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual.* Basingstoke: Palgrave Macmillan.

Schmitt, N. & Zimmerman, C. B. (2002). Derivative word forms: What do learners know? *TESOL Quarterly*, *36*(2), 145 - 171.

Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal, 59*(4), 339 - 341.

Selinker, L. (1972). Interlanguage. *IRAL: International Review of Applied Linguistics 10*(3), 209 - 231.

Semino, E. (2006). A corpus-based study of metaphors for speech activity in British English. In A. Stefanowitsch, & S. T. Gries (Eds.), *Corpus-based approaches to metaphor and metonymy* (pp. 36 - 62). Berlin: Mouton de Gruyter.

Semino, E., Deignan, A. & Littlemore, J. (2013). Metaphor, genre, and recontextualization. *Metaphor and Symbol 28*(1), 41 - 59.

Shaw, P., & Ting-Kun Liu, E. (1998). What develops in the development of second-language writing? *Applied Linguistics, 19*(2), 225 - 254.

Shibatani, M. (1990). *The languages of Japan*. New York: Cambridge University Press.

Sinclair, J. (1991). *Corpus, concordance, collocation.* Oxford: Oxford University Press.

Siyanova-Chanturia, A., Conklin, K. & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research 27*(2), 251 - 272.

Skoufaki, S. (2008). Conceptual metaphoric meaning clues in two idiom presentation methods. In F. Boers, & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 101 - 132). Berlin: Mouton de Gruyter.

Srdanović, I., Suchomel, V., Ogiso, T., & Kilgarriff, A. (2013). Japanese language lexical and grammatical profiling using the web corpus JpTenTen. Retrieved March 15, 2017, from http://www.ninjal.ac.jp/event/specialists/project-meeting/files/JCLWorkshop_no3_papers/JCLWorkshop_No3_29.pdf

Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal, 36*(2), 139 - 152.

Steen, G. (1999). From linguistic to conceptual metaphor in five steps. In R. W. Gibbs & G. J. Steen (Eds.), *Metaphor in Cognitive Linguistics: Selected Papers from the Fifth International Cognitive Linguistics Conference, Amsterdam, July, 1997* (pp. 57 - 78). Amsterdam: John Benjamins Publishing Company.

Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A. A., Krennmayr, T., & Pasma, T. (2010). *A method for linguistic metaphor identification: From MIP to MIPVU.* Amsterdam: John Benjamins Publishing Company.

Stefanowitsch, A. (2006a). Corpus-based approaches to metaphor and metonymy. In A. Stefanowitsch, & S. T. Gries (Eds.), *Corpus-based approaches to metaphor and metonymy* (pp. 1 - 16). Berlin: Mouton de Gruyter.

Stefanowitsch, A. (2006b). Words and their metaphors: A corpus-based approach. In A. Stefanowitsch, & S. T. Gries (Eds.), *Corpus-based approaches to metaphor and metonymy* (pp. 63 - 105). Berlin: Mouton de Gruyter.

Stefanowitsch, A., & Gries, S. T. (2006). *Corpus-based approaches to metaphor and metonymy.* Berlin: Mouton de Gruyter.

Stewart, J., Batty, A. O., & Bovee, N. (2012). Comparing multidimensional and continuum models of vocabulary acquisition: An empirical examination of the vocabulary knowledge scale. *TESOL Quarterly, 46*(4), 695 - 721.

Tannenbaum, R. J., & Wylie, E. C. (2006). *Mapping the TOEIC and TOEIC Bridge tests on the Common European Framework of Reference for Languages.* Retrieved February 24, 2014, from http://www.ets.org/toeic/research /mapping_toeic

Tannenbaum, R. J. & Wylie, E. C. (2008). Linking English-language test scores onto the Common European Framework of Reference: An application of standard-setting methodology. TOEFL iBT Research Report. Educational Testing Service. Retrieved December 31, 2016, from https://www.ets.org/Media/Research/pdf/RR-08-34.pdf

Taylor, J. R. (2003). Polysemy's paradoxes. *Language Sciences, 25*(6), 637 - 655.

Taylor, J. R. (2012). *The mental corpus: How language is represented in the mind*. Oxford: Oxford University Press.

TOEIC (2016). *TOEIC® Listening & Reading Test nyūgaku shiken tan'i nintei ni okeru katsuyō jōkyō* [TOEIC® Listening & Reading Test Entrance examination / Credit utilization] Retrieved December 23, 2016, from http://www.toeic.or.jp/toeic/about/data/search.html

Türker, E. (2016). The role of L1 conceptual and linguistic knowledge and frequency in the acquisition of L2 metaphorical expressions. *Second Language Research 32*(1), 25 - 48.

Turner, S. (2014). *The development of metaphoric competence in French and Japanese learners of English* (Unpublished doctoral dissertation). The University of Birmingham, Birmingham.

Université catholique de Louvain. (2011). *ICLE*. Retrieved from https://uclouvain.be/en/research-institutes/ilc/cecl/iclev2.html

Université catholique de Louvain. (2013). *LOCNESS*. Retrieved from https://uclouvain.be/en/research-institutes/ilc/cecl/locness.html

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics, 22*(2), 217 - 234.

Verspoor, M., & Lowie, W. (2003). Making sense of polysemous words. *Language Learning, 53*(3), 547 - 586.

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning, 61*(1), 219 - 258.

Waring, R. & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language 15*(2), 130 - 163.

Webb, S. (2007). The effect of repetition on vocabulary knowledge. *Applied Linguistics, 28*(1), 46 - 65.

Webb, S., & Chang, A. C. (2012). Second language vocabulary growth. *RELC Journal, 43*(1), 113 - 126.

West, M. (1953). *A general service list of English words.* London: Longman.

Widdowson, H. G. (1997). EIL, ESL, EFL: Global issues and local interests. *World Englishes, 16*(1), 135 - 146.

Wilson, D. (2011). Parallels and differences in the treatment of metaphor in relevance theory and cognitive linguistics. *Studia Linguistica, 128*, 195 - 213.

Woodford, P. E. (1982). An introduction to TOEIC: The initial validity study. TOEIC Research Summaries. Educational Testing Service. Retrieved December 23, 2016, from https://www.ets.org/Media/Research/pdf/TOEIC-RS-00.pdf

Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC Journal*, *37*(3), 308 - 328.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, *86*(1), 54 - 66.

Yazaki, G. (1975). *Nihon no gairaigo*. Japan: Iwanami Shoten.

Yeganehjoo, M., Yap, N. T., Abdullah, M. H., Tan, B. H. (2012). The influence of cross-linguistic similarities on L2 idiom production. *The Southeast Asian Journal of English Language Studies 18*(4), 91 - 107.

LIST OF ELECTRONIC RESOURCES

Anthony, L. (2015). AntConc (Version 3.3.5) [Computer Software]. Tokyo. Available online at http://www.laurenceanthony.net/software.html

Cambridge Dictionary. (n.d.). Available online at http://dictionary.cambridge.org

Cambridge University Press. (2012). Available online at http://vocabulary.englishprofile.org/staticfiles/terms.html

Cobb, T. (n.d.). Vocabprofile. Available online at http://www.lextutor.ca/vp

Davies, M. (2004-). *BYU-BNC*. (Based on the British National Corpus from Oxford University Press). Available online at http://corpus.byu.edu/bnc/

Davies, M. (2008-) *The Corpus of Contemporary American English (COCA): 520 million words, 1990-present*. Available online at http://corpus.byu.edu/coca/

Jisho (n.d.) Available online at http://jisho.org

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. & Suchomel, V. (2013). jpTenTen11 corpus (v3.0). Available online at https://the.sketchengine.co.uk/auth/corpora/

Longman Dictionary of Contemporary English. (n.d.) Available online at http://www.ldoceonline.com

Macmillan Dictionary. (n.d.). Available online at http://www.macmillandictionary.com

Mason, O. (n.d.). Qtag part of speech tagger.

Merriam-Webster dictionary (n.d.) Available online at https://www.merriam-webster.com

Oxford English Dictionary. (n.d.). Available online at http://www.oed.com

Oxford dictionaries. (n.d.) Available online at https://en.oxforddictionaries.com

Rayson, P. (n.d.). Wmatrix3. Available online at

http://ucrel.lancs.ac.uk/wmatrix3.html

Sketch Engine. (n.d.) Available online at https://the.sketchengine.co.uk/login/

LIST OF COURSE MATERIALS USED

British, U.S. music no longer dominates world. (2013, June 7). *The Japan Times*. Retrieved from http://www.japantimes.co.jp/culture/2013/06/07/music/british-u-s-music-no-longer-dominates-world/#.VBBPEUuVsVg

Campbell, E. (Reporter) (2005, October 25). Japan – Generation Z [Television series episode]. *Foreign Correspondent*. Sydney: ABC. Transcript retrieved from http://www.abc.net.au/foreign/content/2005/s1491707.htm

Hagihara, B., Channa, V. (Producers) & Tsutomu, T. (Director). (2002). *Sumo Story* [Documentary]. Discovery Networks Asia.

Harrison, E. (2010, May 26). *Idle idol* [Pecha Kucha presentation]. Retrieved from http://www.pechakucha.org/presentations/idle-idol

Hayashi, K. & Tanizawa, T. (Directors). (2009). *Vocaloid Hatsune Miku, the world's virtual diva*. [Internet video]. News Special: The Asahi Shinbun.

Hilditch, N., Nishiyama, M., Hart, L., (Researchers) & Boyle, M. (Director). (2000). Music in Japan [Television series episode]. In K. Parsons (Producer), *Japan TV*. London: BBC Choice.

Kolker, A. & Alvarez, L. (Producers & directors). (1991). *The Japanese Version* [Documentary]. The Center for New American Media.

Nishiyama, M., Hart, L. (Researchers) & Boyle, M. (Director). (2000). Fashion in Japan [Television series episode]. In K. Parsons (Producer), *Japan TV*. London: BBC Choice.

Parker, J. (2008, May 27). Home Alone. *The Japan Times*. Retrieved from http://www.japantimes.co.jp/life/2008/05/27/lifestyle/home-alone/#.VBBJ30uVsVg

Plath, D. W. (Writer) & Ikeda, K. (Producer). (1999). *Makiko's New World* [Documentary]. Media Production Group.

Rusak, D. (2009). Karate, baseball and politics: Hybridity and the martial arts in modern Japan. *Undergraduate Journal of Anthropology, 1*, 63 - 71.

Appendix A. Participant data

Control group data

Code	Gender	Age*	Nationality	Maximum TOEIC	Estimated CEFR level	TOEIC group
CON1	F	20	Japanese	730	B1	High
CON2	F	19	Japanese	365	A2	Low
CON3	F	19	Japanese	245	A2	Low
CON4	M	19	Japanese (Korean fluency)	625	B1	High
CON5	M	19	Japanese	485	A2	Mid
CON6	M	20	Japanese	510	A2	High
CON7	F	19	Japanese	525	A2	High
CON8	M	19	Japanese	840	B2	High
CON9	M	19	Japanese	410	A2	Mid
CON10	F	19	Japanese	360	A2	Low
CON11	F	19	Japanese	405	A2	Low
CON12	F	19	Japanese	350	A2	Low
CON 13	M	21	Japanese	475	A2	Mid
CON14	M	19	Japanese	620	B1	High
CON15	F	19	Japanese	310	A2	Low
CON16	M	19	Japanese	520	A2	High
CON17	M	19	Japanese	575	B1	High
CON18	M	19	Japanese	435	A2	Mid
CON19	M	19	Japanese	320	A2	Low
CON20	M	19	Japanese	485	A2	Mid
CON21	M	19	Japanese	470	A2	Mid
CON22	F	19	Japanese	420	A2	Mid
CON23	F	19	Japanese	445	A2	Mid

Experimental group data

Code	Gender	Age*	Nationality	Maximum TOEIC	Estimated CEFR level	TOEIC group
EXP1	F	19	Japanese	645	B1	High
EXP2	F	19	Japanese	470	A2	Mid
EXP3	M	19	Japanese	285	A2	Low
EXP4	F	19	Japanese	435	A2	Mid
EXP5	F	19	Japanese	420	A2	Low
EXP6	F	19	Japanese	335	A2	Low
EXP7	M	19	Japanese	570	B1	High
EXP8	F	19	Japanese	860	B2	High
EXP9	M	19	Japanese	515	A2	High
EXP10	F	19	Japanese	510	A2	Mid
EXP11	M	19	Japanese	860	B2	High
EXP12	F	19	Japanese	420	A2	Low
EXP13	F	19	Japanese	440	A2	Mid
EXP14	M	20	Japanese	350	A2	Low
EXP15	M	19	Japanese	515	A2	High
EXP16	F	19	Japanese	455	A2	Mid
EXP17	M	19	Japanese	500	A2	Mid
EXP18	M	20	Japanese	435	A2	Mid
EXP19	M	19	Japanese	565	B1	High
EXP20	F	19	Japanese	415	A2	Low
EXP21	F	19	Japanese	385	A2	Low
EXP22	M	19	Japanese	670	B1	High
EXP23	M	19	Japanese	410	A2	Low

* = Age at beginning of course

Appendix B. Course activities

Lesson	(Date) Activities		Writing assignment
	Control group	**Experimental group**	
1	(05.04.13) Introduction lesson Discussing popular culture Map of Japan - marking cultural entities Japanese-ness Foreign images of Japanese culture	(04.04.14) Introduction lesson Discussing popular culture ***Metaphor work*** Introduction to metaphor PRECIOUS MATERIALS ARE VALUABLE, WONDERFUL OR BEAUTIFUL ENTITIES CULTURES ARE LIVING THINGS	What is your personal image of Japanese popular culture? What does this phrase mean to you?
2	(09.04.13) Music in Japan Japanese music overseas Visual bands Manufactured bands	(08.04.14) Music in Japan Japanese music overseas Visual bands ***Metaphor work*** Review	
3	(12.04.13) Music in Japan Hiphop in Japan Discussion Language work Word roots / affixes	(11.04.14) Music in Japan Manufactured bands Hiphop in Japan Discussion ***Metaphor work*** MORE OF AN ABSTRACT THING IS AN INCREASE IN SIZE OR HEIGHT	Responding to discussion questions about music in Japan. Do Japanese musicians copy western styles? Is it important for Japanese musicians to succeed overseas? Should Japanese musicians have to sing in English to become popular overseas?
4	(16.04.13) Hatsune Miku / Vocaloid Music in Japan review Hatsune Miku Vocaloid discussion Fashion in Japan Kogyaru	(15.04.14) ***Metaphor work*** Sources of metaphor: plants Hatsune Miku / Vocaloid Hatsune Miku Vocaloid discussion	
5	(19.04.13) Fashion in Japan Kogyaru Youth fashion Ginza girls The Japanese and fashion	(18.04.14) Language work Fixing common errors (inc. metaphor) Fashion in Japan Kogyaru ***Metaphor work*** MOVING FORWARD IS IMPROVEMENT Fashion in Japan Youth fashion	Hatsune Miku / Vocaloid What is your aesthetic reaction to this music? How do you feel about people using new technology to create their own music? How does this music compare to non-vocaloid music? Who is influencing whom here?

(continued)

Lesson	(Date) Activities		Writing assignment
	Control group	**Experimental group**	
6	(23.04.13) Language work Vocabulary review Fashion in Japan Japanese fashion designers Foreign models in Japan	(22.04.14) Fashion in Japan Ginza Girls The Japanese and fashion Japanese fashion designers Foreign models in Japan	
7	(26.04.13) Japanese tribes Gothic Lolitas Otaku Freeters NEETs Language work Word parts Hybridity and hybridism Reading activity	(25.04.14) Japanese tribes Gothic Lolitas Otaku Freeters NEETs ***Metaphor work*** MOVEMENT IS FREEDOM STRONG EMOTIONS ARE SICKNESS/INSANITY BEING INTERESTED IS ENTERING SOMETHING STRONG EMOTION IS PHYSICAL FORCE Language work Word parts	Give your opinions about one or more of the Japanese tribes. Can you identify with any of these groups or can you at least understand their opinions?
8	(30.04.13) Hybridity and hybridism Reading activity Word parts: -ity vs. -ism Karate and baseball	(02.05.14) Hybridity and hybridism Japanese-ness Discussing hybrid cultural entities Reading activity	
9	(10.05.13) Hybridity and hybridism Examining hybridity and hybridism Planning research paper Reading model essay Language work Cohesive writing	(09.05.14) Hybridity and hybridism Examining hybridity and hybridism ***Metaphor work*** CULTURES ARE SPACES Hybridity and hybridism Karate and baseball	What examples of cultural hybridity and hybridism do you see around you in Japan? How do you feel about them?
10	(14.05.13) Makiko's New World Imagining Meiji era Japan Makiko's diary Urban lifestyles in Meiji Japan	(13.05.14) Makiko's New World Imagining Meiji era Japan Makiko's diary Urban lifestyles in Meiji Japan Hybridity and hybridism Planning research paper	

(continued)

Lesson	(Date) Activities		Writing assignment
	Control group	**Experimental group**	
11	(17.05.13) Makiko's New World Diaries and women's lives Hybridity and hybridism Collocation Research paper drafting	(16.05.14) Makiko's New World Diaries and women's lives Hybridity and hybridism Reading model essay Text analysis Language work Cohesive writing Collocation	What are some similarities and differences between your life and Makiko Nakano's life?
12	(21.05.13) Language work Vocabulary review Makiko's New World Houses in the Meiji Meiji influences on modern Japan	(20.05.14) Language work Vocabulary building Makiko's New World Houses in the Meiji Meiji influences on modern Japan	
13	(24.05.13) Makiko's New World Review activity Family relationships Japanese wedding ceremonies Hybridity and hybridism Editing research papers	(23.05.14) Makiko's New World Family relationships Japanese wedding ceremonies Hybridity and hybridism Editing research papers ***Metaphor work*** Review	How do you feel about Japanese wedding ceremonies? Which elements show hybridity? Can you find examples of hybridism?
14	(28.05.13) Hybridity and hybridism Editing research papers Makiko's New World Western food Makiko's character	(27.05.14) Hybridity and hybridism Editing research papers Makiko's New World Western food	
15	(31.05.13) Hybridity and hybridism Discussing ideas for research papers Makiko's New World Makiko's character Conclusion	(30.05.14) Makiko's New World Makiko's character Conclusion ***Metaphor work*** Review of metaphor in Makiko's New World	What did you learn from Makiko's New World? What surprised you? What did you feel was good or positive about their lives? What did you think was not so good about their lives?
16	(04.06.13) Hybridity and hybridism Editing research papers Language work Vocabulary review test Makiko's New World Social class Language work Describing trends	(03.06.14) Hybridity and hybridism Editing research papers Language / metaphor work Vocabulary review test Makiko's New World Social class ***Metaphor work*** MORE IS UP AND LESS IS DOWN	

(continued)

Lesson	(Date) Activities		Writing assignment
	Control group	**Experimental group**	
17	(07.06.13) Language work Describing trends Describing pie chart data Home Alone Reading on trends in modern living arrangements	(06.06.14) ***Metaphor work*** Describing trends with metaphor Language work Describing trends Home Alone Reading on trends in modern living arrangements	Discuss the popularity of Japanese versus foreign music in Japan with reference to newspaper article.
18	(11.06.13) Language work Review vocabulary test answers Home Alone Reading activity Language work Describing bar chart data	(10.06.14) Home Alone Review text Language work Describing pie chart data Describing bar chart data ***Metaphor work*** Metaphorical adjectives	Choose an element of Japanese culture that is a hybrid. Describe how it shows hybridity and whether it shows hybridism.
19	(14.06.13) Sumo Story What is sumo? Senior yobidashi Sumo and religion Yobidashi Sumo rankings	(13.06.14) Language work Describing charts Sumo Story What is sumo? Senior yobidashi Sumo and religion ***Metaphor work*** ABSTRACT ENTITIES ARE PHYSICAL STRUCTURES	Compare sumo with other sports played in Japan. Makes comparisons in terms of traditions, history, popularity etc.
20	(18.06.13) Language work Vocabulary review Sumo Story Sumo rankings Winning in sumo History of sumo Sumo in the Meiji era Modernization of sumo Challenges for sumo	(17.06.14) Sumo Story Yobidashi Sumo rankings Winning in sumo History of sumo Sumo in the Meiji era Modernization of sumo	
21	(21.06.13) Language work Vocabulary review Sumo Story Akebono The Tokyo tournament Conclusion Discussion Language work Collocations	(20.06.14) Sumo Story Challenges for sumo Akebono The Tokyo tournament Conclusion Language work Collocations	Sumo: Past, present and future How has sumo changed? How has it resisted change? What are some of the problems it faces?

(continued)

Lesson	(Date) Activities Control group	Experimental group	Writing assignment
22	(25.06.13) Language work Appositives The Japanese Version High vs. low culture Introduction Love hotels	(24.06.14) The Japanese Version High vs. low culture Introduction Love Hotels ***Metaphor work*** CULTURES ARE A MASS CULTURES ARE CONSTRUCTIONS CULTURES HAVE MANY PARTS	
23	(28.06.13) Language work Fixing common errors in writing The Japanese Version Japanese weddings Final project Pecha Kucha video Explanation	(27.06.14) The Japanese Version Japanese weddings ***Metaphor work*** CULTURES ARE POSSESSIONS Final project Pecha Kucha video Explanation	Japanese weddings in the future – how do you think they will continue to evolve? What trends will emerge? What cultural influences will appear?
24	(02.07.13) The Japanese Version Japanese Cowboys English in Japan Western style and appearance Final project Preparation time	(01.07.14) Language / metaphor work Vocabulary review The Japanese Version Cherry blossom revellers Japanese cowboys ***Metaphor work*** Sources of metaphor: the body	
25	(05.07.13) Language work Vocabulary review The Japanese Version Japanese culture Final project Preparation time	(04.07.14) ***Language / metaphor work*** Vocabulary review The Japanese Version Japanese culture Final project Preparation time	Are the Japanese losing their culture? Is there too much American influence? Are the Japanese westernized or modernized?
26	(09.07.13) The Japanese Version Cherry blossom revellers Final project Preparation time	(08.07.14) ***Language / metaphor work*** Vocabulary review The Japanese Version English in Japan Western style and appearance Final project Preparation time	
27	(12.07.13) The Japanese Version Foreign TV stars Foreigners in advertising Final project Preparation time	(11.07.14) The Japanese Version Foreign TV stars Foreigners in advertising Final project Preparation time	

(continued)

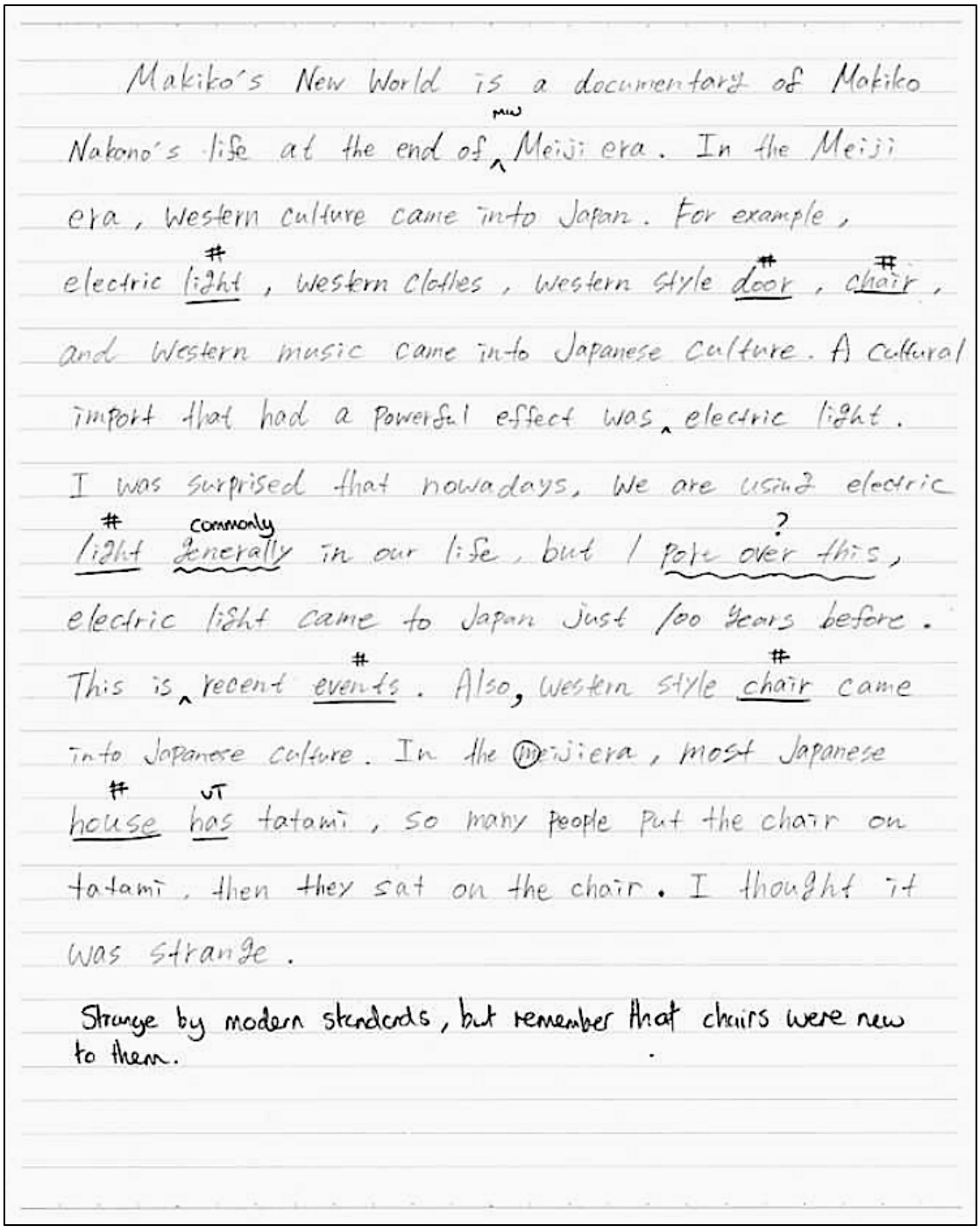
Lesson	(Date) Activities		Writing assignment
	Control group	**Experimental group**	
28	(16.07.13) Final project Presentations Exam Vocabulary test	(15.07.14) ***Language / metaphor work*** Vocabulary review Final project Preparation time	
29	*(no lesson)*	(18.07.14) Final project Presentations	
30	*(no lesson)*	(22.07.14) Final project Presentations	
Exam	(19.07.13) Final project Presentations Exam Written exam	(25.07.14) Exam Vocabulary test Written exam	Exam topic 1: What element of Japanese Popular Culture would you choose to teach? Exam topic 2: How do you think Japanese society has changed since Makiko's era, and how is that reflected in popular culture? Exam topic 3: Define 'hybridity' and 'hybridism' and give your opinion on why Japan might use hybridism to construct its identity in global society.

Appendix C. Error correction code and written feedback example

Error correction code used in the study

Symbol	Meaning
SP	spelling
#	singular/plural
P	punctuation
VT	verb tense
A/P	active/passive voice
SVA	subject/verb agreement
WO	word order
WW	wrong word
NA	not appropriate
WF	word form
^	missing word
═══	unnecessary word
NC	not a complete sentence
?	I don't understand.

Sample of learner writing with feedback and error correction symbols



Makiko's New World is a documentary of Makiko Nakano's life at the end of ^Meiji era. In the Meiji era, Western culture came into Japan. For example, electric light, Western clothes, Western style door, chair, and Western music came into Japanese culture. A cultural import that had a powerful effect was ^electric light. I was surprised that nowadays, We are using electric light generally in our life, but I pore over this, electric light came to Japan just 100 years before. This is ^recent events. Also, Western style chair came into Japanese culture. In the Meiji era, most Japanese house has tatami, so many people put the chair on tatami, then they sat on the chair. I thought it was strange.

Strange by modern standards, but remember that chairs were new to them.

Appendix D. Exceptions to MIPVU coding

1. Words listed in separate senses in dictionaries

As was discussed in section 3.4.2.2, the following words were coded as MRWs despite appearing in their own separate listing in the Macmillan dictionary.

Metaphor theme	Item	Freq. in output?		
		Con.	Exp.	BAWE
Sources of metaphor: plants	branch out (ph v.)	0	1	0
	deep-rooted (adj)	1	0	0
	grassroots (adj)	0	0	0
	leaf through (ph. v)	0	0	0
	stem from (ph. v)	0	1	2
MOVING FORWARD IS IMPROVEMENT	forefront (n)	0	0	0
	progressive (adj)[45]	0	1	0
MOVEMENT IS FREEDOM	free-spirited (adj)[46]	0	1	0
MORE OF AN ABSTRACT THING IS AN INCREASE IN SIZE OR HEIGHT	mushroom (v)	0	1	0
	skyrocket (v)	0	1	0
	snowball (v)	0	3	0
MORE IS UP, LESS IS DOWN	bottom out (ph. v)	0	0	0
	peak (v)	0	0	0
	plateau (v)	0	0	0
	rock bottom (n)	0	0	0
STRONG EMOTION IS PHYSICAL FORCE	carried away (adj)[46]	0	0	0
BEING INTERESTED IS ENTERING SOMETHING	absorbed in (adj)	3	2	0
	immersed in (adj)	0	2	0
	TOTALS	4	13	2
	% of lexical units (MIPVU)	0.01%	0.02%	0.01%

[45] The Macmillan entry for *progressive* actually contains four senses, but none related to physical forward movement, which is how the word was taught.

[46] Macmillan only contains definitions for the noun form *free spirit* and the verb phrase *get carried away*. Since these have only one listed sense each, both *free-spirited* and *carried away* were analysed based on these meanings. This is consistent with Steen et al. (2010, p. 36).

2. Words with conflated senses in dictionaries

The following words were selected for teaching in the metaphor workbook as members of the metaphor themes listed in the table, although their dictionary entries conflate physical and abstract senses.

Metaphor theme	Item	Freq. in output?		
		Con.	Exp.	BAWE
MOVING FORWARD IS IMPROVEMENT	static (adj)	0	1	1
CULTURES ARE CONSTRUCTIONS	undamaged (adj)	0	0	0
MORE IS AN INCREASE	enormous (adj)	0	0	0
	immense (adj)	0	0	0
	TOTALS	0	1	1
	% of lexical units (MIPVU)	0.00%	0.00%	0.00%

3. Metonymy and metaphor

The table below lists the frequencies with which figurative uses of body part words appeared in the learner output corpora and the BAWE corpus. Below are some examples to distinguish the three cases.

Metaphor theme	Item	Metonymy only Freq. in output?			Metaphor and metonymy Freq. in output?			Metaphor only Freq. in output?		
		Con.	Exp.	BAWE	Con.	Exp.	BAWE	Con.	Exp.	BAWE
Sources of metaphor: the body	head (n)	0	1	0	0	0	2	0	0	0
	eye (n)	0	1	0	0	1	4	0	0	0
	eye (v)	0	0	0	0	0	0	0	0	0
	ear (n)	1	0	0	0	0	0	0	0	0
	mouth (n)	0	0	0	0	0	0	0	0	0
	mouth off (ph. v)	0	0	0	0	0	0	0	0	0
	shoulder (n)	0	0	0	0	0	0	0	0	0
	shoulder (v)	0	0	0	0	0	0	0	0	0
	heart (n)	0	0	0	4	5	0	0	0	1
	hand (n)	0	0	1	0	0	2	0	1	1
	hand (v)	0	0	0	0	0	1	0	0	0
	TOTALS	1	2	1	4	6	9	0	1	2
	% of lexical units (MIPVU)	0.00%	0.00%	0.00%	0.01%	0.01%	0.04%	0.00%	0.00%	0.01%

Metonymy only:

In these cases, the reference of the highlighted words remains within the human body, although a PART FOR WHOLE relationship of meaning exists. These uses were not coded as MRWs in the study.

- [A singer's] voice stuck in their head. [EXP12]
- Otaku should turn their eyes to not only the virtual world but also real world. [EXP4]

- ...the novice hunter learns by accompanying more experienced hands in the woods... [BAWE3016c]

Metaphor and metonymy:

In these cases, there is a relationship of meaning both within and across domains. The highlighted words are still referring to individual human beings, but there is also a crossing of domains from the physical body to actions, emotions, or opinions. These cases were coded as MRWs in the study.

- They are putting their heart and soul into various things that they like or love. [CON23]
- ... I can feel a tender heart of Makiko from [a diary]. [EXP16]
- To write an ethnography with no consideration of gender could in many scholars' eyes deem it as incomplete. [BAWE3014b]

Metaphor only:

The cases below are further removed from the physical domain of a human body. I the first case, there is perhaps an argument for metonymy still being present, but the broader context describes does not indicate any sense of physical action on the part of leaders. The second case is more of a 'pure' metaphor, in which the highlighted word is being used in its 'central part' sense (Macmillan sense 4), and there is no connection to the human body. Both cases were coded as MRWs.

- ...[leaders] hold the fate of the society in their hands... [BAWE3135c]
- Conversely the rationality of the optimal forager is installed at the every heart of nature... [BAWE3016c]

4. Polywords not recorded on the BNC list

Phrases that were not included on the BNC polyword list were treated as follows:

- In chapter fours and five (sections 5.3 - 5.5), they were recorded as separate lexical units and analysed individually, as is detailed in MIPVU.
- In chapter five (section 5.6) and chapters 6 - 7, they were recorded as single lexical units but tagged as MRWs if any of the component words would have been coded as MRWs under MIPVU.

Polyword	Listed in Macmillan dictionary as...	Freq. in output?		
		Con.	Exp.	BAWE
a great deal of	Phrase	0	0	0
above all	Phrase	2	2	0
after all	Phrase	0	1	0
as it is	Phrase	0	0	1
at first glance	Phrase	0	0	1
at the same time	Phrase	2	0	1
bring to light	Phrase	0	0	1
by the way	Phrase	6	2	0
come to a head	Phrase	1	0	0
come true	Phrase	1	2	0
get rid of	Verb	1	0	0
in collaboration with	Preposition	0	0	2
in comparison	Phrase	0	1	0
in comparison to	Phrase	0	0	2
in conclusion	Phrase	17	7	7
in contrast	Phrase	5	6	0
in contrast to	Phrase	0	0	3
in detail	Phrase	0	2	0
in fact	Phrase	12	13	4
in order that	Phrase	0	0	1
in order to	Phrase	12	1	11
in other words	Phrase	4	0	3

(*continued*)

Polyword	Listed in Macmillan dictionary as...	Freq. in output?		
		Con.	Exp.	BAWE
in place	Phrase	0	0	1
in summary	Phrase	8	7	0
in turn	Phrase	0	0	1
make sense	Phrase	0	0	1
on sale	Phrase	0	1	0
on the contrary	Phrase	0	0	1
on the edge of	Phrase	1	0	0
on the other hand	Phrase	62	42	3
on the spot	Phrase	0	1	0
once in a while	Phrase	2	0	0
out of line	Phrase	0	0	1
out of order	Phrase	1	0	0
out of place	Phrase	0	0	1
out of the ordinary	Phrase	1	0	0
outside of	Preposition	0	0	1
pay attention	Phrase	3	4	0
rack their brain	Phrase	1	0	0
so far	Phrase	4	2	0
take care	Phrase	1	0	0
take hold	Phrase	0	1	1
take part	Phrase	0	0	2
take place	Phrase	5	10	5
take root	Phrase	1	0	0
these days	Phrase	13	22	1
	TOTALS	166	127	56
	% of lexical units (MIPVU)	0.38%	0.22%	0.23%

Appendix E. Annotations used in input corpora

Annotation type	Examples
Time stamps	<01:30:15>
Periods of lessons	<ACTIVITY BEGINS>
	<ACTIVITY EXPLANATION>
	<ACTIVITY MODELLING>
	<ACTIVITY REVIEW>
	<ANSWER SHARING>
	<CLASS EXPLANATION>
	<DVD BEGINS> <TITLE/SECTION/PLAY>
	<HANDING OUT MATERIALS>
	<HOMEWORK EXPLANATION BEGINS>
	<LESSON BREAK BEGINS>
	<OTHER EXPLANATION>
	<STUDENT PRESENTATION BEGINS>
	<STUDENTS MOVING>
	<TEACHERS COLLECT HOMEWORK>
	<TEACHERS TALK TO INDIVIDUALS>
	<TEXT>
Student involvement	<STUDENT> <COMMENTS>
	<STUDENT> <GESTURES>
	<STUDENT> <QUESTIONS>
	<STUDENT> <RESPONDS>
	<STUDENT> <SNEEZES>
Instructor actions	<CLAPS>
	<COUGHS>
	<FOLDS PAPER>
	<LAUGHS>
	<LAUGHTER>
	<MOVES WHITEBOARD>
	<SNEEZES>
	<TEACHERS DISCUSS>
	<WHISTLES>
	<WIPES WHITEBOARD>
	<WRITES ON BLACKBOARD>
	<WRITES ON WHITEBOARD>
Instructor speaking	<TEACHER1>
	<TEACHER2>
Other	<INAUDIBLE>

Appendix F. Input corpora data

Control condition

Lesson	Teacher 1[a]	Teacher 2[a]	Audio materials	Written materials[b]	Total
1	4,246	3,395	0	95	**7,736**
2	5,054	2,614	1,898	236	**9,802**
3	4,738	2,572	974	907	**9,191**
4	3,324	2,165	1,975	81	**7,545**
5	4,769	2,440	1,463	372	**9,044**
6	4,310	4,206	1,401	234	**10,151**
7	1,828	635	0	550	**3,013**
8	4,007	4,644	0	422	**9,073**
9	4,375	4,012	0	355	**8,742**
10	4,389	3,663	2,051	296	**10,399**
11	4,533	4,292	1,573	406	**10,804**
12	4,707	3,670	2,255	352	**10,984**
13	3,877	3,573	1,674	124	**9,248**
14	4,728	4,027	2,379	328	**11,462**
15	3,444	2,290	1,867	325	**7,926**
16	3,900	3,828	0	607	**8,335**
17	4,124	2,406	0	1,293	**7,823**
18	3,373	2,203	0	1,901	**7,477**
19	3,783	2,466	2,824	533	**9,606**
20	5,349	1,956	3,233	999	**11,537**
21	4,566	2,456	1,012	1,106	**9,140**
22	5,756	2,822	550	700	**9,828**
23	5,679	4,790	2,297	394	**13,160**
24	5,202	3,539	1,874	633	**11,248**
25	3,876	1,832	1,382	630	**7,720**
26	1,690	1,491	859	267	**4,307**
27	3,142	1,444	1,995	477	**7,058**
28	1,041	633	0	349	**2,023**
Exam	571	275	0	135	**981**
Total	**114,381** (46.62% of total)	**80,339** (32.74% of total)	**35,536** (14.48% of total)	**15,107** (6.16% of total)	**245,363**

Experimental condition

Lesson	Teacher 1[a]	Teacher 2[a]	Audio materials	Written materials[b]	Total
1	7,501	4,742	0	843	**13,086**
2	6,791	3,464	1,180	648	**12,083**
3	4,430	3,601	1,412	610	**10,053**
4	4,272	2,190	2,550	303	**9,315**
5	5,985	2,714	1,474	1,013	**11,186**
6	6,132	2,724	1,916	636	**11,408**
7	4,753	1,906	0	709	**7,368**
8	5,560	4,813	0	740	**11,113**
9	3,781	6,433	0	564	**10,778**
10	5,167	4,005	1,663	428	**11,263**
11	7,834	0	1,804	1,587	**11,225**
12	4,310	4,804	2,926	542	**12,582**
13	3,836	4,579	2,707	850	**11,972**
14	3,809	2,776	1,704	324	**8,613**
15	3,317	2,933	1,944	522	**8,716**
16	4,133	2,926	0	931	**7,990**
17	5,027	2,275	0	2,956	**10,258**
18	6,066	3,686	0	1,331	**11,083**
19	4,086	3,240	2,253	1,404	**10,983**
20	4,659	3,017	3,748	1,624	**13,048**
21	4,346	3,111	1,529	1,503	**10,489**
22	5,578	4,731	1,138	641	**12,088**
23	5,139	4,784	2,120	368	**12,411**
24	4,653	3,994	1,452	1,116	**11,215**
25	8,629	0	1,225	1002	**10,856**
26	4,725	3,265	1,290	687	**9,967**
27	3,401	2,991	900	459	**7,751**
28	1,513	1,668	0	283	**2,464**
29	1,355	583	0	0	**1,938**
30	1,022	477	0	129	**1,628**
Exam	585	0	0	992	**1,577**
Total	**142,395** (47.86% of total)	**92,432** (31.07% of total)	**36,935** (12.41% of total)	**25,745** (8.65% of total)	**297,507**

Note: Teacher 1 = author (language instructor), Teacher 2 = teaching partner (anthropology instructor)

a = Teacher input counts include teacher talk to the whole class. Teacher talk to individuals or small groups, student talk to the whole class, and student-to-student talk was not recorded.

b = Includes all printed material distributed to students. Does not include text written on whiteboards.

Adjusting the data for teacher absences

To investigate whether teacher absences were a possible cause of some of the significant differences in input between the conditions, the data were recalculated with different values for teacher input for the periods teacher 2 was absent. In the data for lessons 11 and 25, the values for teachers 1 and 2 were replaced with values reflecting the ratio of words produced by each teacher in the other regular lessons as follows:

- Total teacher 1 words in lessons 1 - 10, 12 - 24, and 26 - 30: 125,347
- Total teacher 2 words in lessons 1 - 10, 12 - 24, and 26 - 30: 92,432
- Ratio teacher 2 words to teacher 1 words: 92,432/125,347 = 0.74
- Mean teacher 1 words in lessons 1 - 10, 12 - 24, and 26 - 30: 125,347/28 = 4,477
- Approximated teacher 2 input for lessons with absences: 4477 x 0.74 = 3,301

In lessons 11 and 25, the input word count values for teacher 1 and teacher 2 were replaced with 4,477 and 3,301 respectively. Since the examination period contained fewer words of teacher input than regular lessons, a different calculation was used. Here, the approximated value for teacher 2 was simply 74% of the words produced by teacher 1 during the examination period (585 x 0.74 = 431), while the value for teacher 1 remained unchanged.

Using these values, the new total size for the experimental input corpus became 297,031 tokens, of which teacher 1 provided 134,886 tokens and teacher 2 99,465 tokens. These values were only used for the calculations in Table 3.8 on page 111.

Appendix G. Taught target Metaphors

Examples are based on samples from course input and output, but have been modified for clarity and brevity.

Metaphor theme	Metaphor-related word	POS	Example
PRECIOUS MATERIALS ARE VALUABLE, WONDERFUL, OR BEAUTIFUL ENTITIES	gem	noun	That idea is a gem.
	jewel	noun	Kyoto is a jewel of a city.
	treasure	noun	Sumo is a treasure of Japan.
	treasure	verb	I will treasure this memory.
	treasured	adjective	One of my treasured memories.
Sources of metaphor: Plants	blossom	verb	She started a small restaurant ten years ago, and now it has blossomed into a very successful local business.
	branch	noun	The business has several branches throughout Kyoto.
	branch out	phrasal verb	Rakuten began as an online shopping service, but has branched out into banking and travel.
	budding	adjective	Hatsune Miku's budding creators met on the Internet.
	deep-rooted	adjective	His fashion style are deep-rooted in Japanese culture.
	flourish	verb	How did the government help Kyoto to flourish again?
	fruit	noun	This agreement is the fruit of two years of discussion between our companies.
	grassroots	adjective	Because of Vocaloid technology, more music is being created at a grassroots level.
	leaf through	phrasal verb	I was leafing through a magazine.
	roots	noun	Tempura has its roots in Portuguese food.
	seeds	noun	These people are planting the seeds of a successful future for the Tohoku region.
	stem from	phrasal verb	The popularity of anime stems from its interesting storylines and beautiful artwork.
	uproot	verb	The Nakano family were uprooted because of the air raids on Kyoto.
	wilt	verb	He has been training for the marathon every day, but now that the weather is getting colder, his enthusiasm is wilting.
	wither	verb	Kyoto's economy withered after Tokyo became the capital of Japan.
MOVING FORWARD IS IMPROVEMENT	advance	noun	The advance of technology has changed our lives.
(*continued*)	advance	verb	As technology advances, using synthesized voices to create music will probably become more common.

Metaphor theme	Metaphor-related word	POS	Example
MOVING FORWARD IS IMPROVEMENT (*cont.*)	barrier	noun	Japan has always been a mysterious country, and I think people need to break down those barriers.
	driving force	noun	Teenagers are the driving force of Japanese fashion.
	forefront	noun	Japan is regarded as being at the forefront of the animated movie business.
	hold back	phrasal verb	The sumo association is often misunderstood as being held back by old traditions.
	keep up	phrasal verb	I used to try to keep up with the latest fashion trends, but now I just wear what I want.
	lead	verb	Women now have a more independent way of dressing, and street fashion leads the way.
	move	verb	After its sudden decline, Kyoto's mayors did all they could to get the city moving again.
	movement	noun	Japanese designers are leading some of the most important fashion movements.
	obstacle	noun	Being successful in sumo requires having the spirit to overcome any obstacle in your way.
	progress	noun	Banzuke are signs of a wrestler's progress in the sumo world.
	progress	verb	She enjoyed playing the piano so much that she progressed very quickly.
	progressive	adjective	Makiko's father was a progressive man, willing to give his young wife plenty of freedom to be away from the house.
	slow down	phrasal verb	Business slowed down across Japan after the 2011 Tohoku disasters.
	static	adjective	Even in the seemingly static world of sumo, there is change.
	verge	noun	Some Japanese artists are on the verge of international success.
MOVEMENT IS FREEDOM	confine	verb	The stereotype of the old Kyoto bride is that she was squeezed and confined into a tiny space.
	escape	verb	Hakone is a nice place to escape from the pressure of living in Tokyo.
	flexible	adjective	This company offers flexible working hours.
	free-spirited	adjective	I wonder if there is anybody as free-spirited as I am, flitting around outside the house every day.
	release	noun	Reading is a great release from the problems of the real world.
(*continued*)	rigid	adjective	Work environments in Japan can be very rigid, so some young people prefer temporary work.

Metaphor theme	Metaphor-related word	POS	Example
MOVEMENT IS FREEDOM (*cont.*)	squeeze	verb	The stereotype of the old Kyoto bride is that she was squeezed and confined into a tiny space.
Sources of metaphor: The body	ear	noun	She has a real ear for music.
	eye	noun	He has a good eye for detail.
	eye	verb	She eyed the stranger suspiciously.
	foot	noun	The town lies at the foot of the mountain.
	hand	noun	Can you give me a hand with this package?
	hand	verb	She handed him the package.
	head	noun	Makiko became the head of the family business after her husband died.
	heart	noun	The university stands in the heart of the city.
	mouth	noun	Miyazaki sits at the mouth of the Oyodo River.
	mouth off	phrasal verb	He is always mouthing off about other people.
	shoulder	noun	After her father's death, a lot more responsibility fell on Makiko's shoulders.
	shoulder	verb	The hikers shouldered their rucksacks and began walking again.
	spine	noun	The author's name is written on the spine of the book.
	spineless	adjective	He's too spineless to disagree with the president.
Metaphorical adjectives	big	adjective	I'm a big fan of Hayao Miyazaki.
	bright	adjective	She has a bright future in this company.
	broad	adjective	The music market is much broader now.
	dark	adjective	He has a dark side to his character.
	deep	adjective	We had a deep conversation for over an hour.
	high	adjective	Pale skin used to be a sign of high status.
	low	adjective	Nearly half of single 20- to 34-year-olds still live with a parent, though the proportion in Tokyo is much lower.
	narrow	adjective	He can be very narrow-minded.
	shallow	adjective	The characters in this movie are so shallow.
	small	adjective	I have a small question to ask.
	wide	adjective	They have a wide selection of cheeses.
CULTURES ARE LIVING THINGS	alive	adjective	They are trying to keep the sport alive.
	evolution	noun	Technology has helped the evolution of music.
	evolve	verb	Their music has evolved.
	grow	verb	The popularity of hip-hop is growing rapidly in Japan.
	hybrid	adjective	California Roll is a hybrid food.
	hybrid	noun	Many popular Japanese dishes are cultural hybrids.

Metaphor theme	Metaphor-related word	POS	Example
CULTURES ARE SPACES	come into	phrasal verb	This was a time when a lot of Western styles came into Japanese culture.
	enter	verb	Baseball entered Japan in the 19th century.
	export	noun	Techno has been Japan's most successful musical export.
	export	verb	Japan has exported its anime styles into other cultures.
	import	noun	Valentine's Day is a cultural import.
	import	verb	Japan has imported a lot of fashion styles from the West.
	inside	preposition	Inside every culture, there are both original and borrowed entities.
	outside	adverb	At some point in the future, you are going to step outside your own culture.
	outside	noun	Then you will see your own culture from the outside.
	outside	preposition	This came from outside of your culture.
	outsider	noun	How do you think your culture appears to outsiders?
	surround	verb	I grew up with Japanese culture surrounding me.
	take in	phrasal verb	Japan took in many Western products and ideas in the Meiji era.
	transcend	verb	Anime is something that has transcended Japanese culture.
CULTURES ARE SUBSTANCES	absorb	verb	Japanese culture has absorbed many things from overseas.
	flow	noun	Japan has been influenced by many other countries, and that flow of influence is still going on.
	flow	verb	Today, foreign culture is flowing into Japan.
	fusion	noun	Modern Japanese music contains a multitude of different sounds in a fusion of east and west.
	mix	noun	Their music is an eclectic mix of influences.
	mix	verb	Omuraisu mixes the cultures of Japan and France.
	mixture	noun	This culture is a mixture of what's come in from the outside.
	shape	noun	Modern Japanese weddings come in every shape and style.
	shape	verb	We are using hybridity as an example of how cultures are shaped.
	spread	verb	Eventually, sumo spread across the whole country.
CULTURES ARE CONSTRUCTIONS	construct	verb	The Japanese have constructed their identity in response to Western pressure and influences.

(*continued*)

Metaphor theme	Metaphor-related word	POS	Example
CULTURES ARE CONSTRUCTIONS (*cont*.)	fit	verb	Japanese sushi was adapted to fit American people's tastes.
	fit into	phrasal verb	The 'borrowing' country adapts its imports to fit into its own culture.
	made up	phrasal verb	Cultures are made up of various influences.
	manufacture	verb	Many modern bands are manufactured by record companies.
	(un)damaged	adjective	The Meiji leaders tried to present Japan as undamaged by the influence of Western culture.
CULTURES HAVE MANY PARTS	aspect	noun	The documentary focuses on aspects of Japanese culture that are seldom seen outside the country.
	component	noun	Hierarchy is an essential component of sumo.
	divide	verb	In Japan, food is divided into Western and Japanese styles.
	element	noun	Sumo demonstrates clear links to other elements of Japanese culture.
	ingredient	noun	History, adaptation, and bushido are essential ingredients of the soul of Japan.
	part	noun	Japan claims tempura as part of its cuisine.
CULTURES ARE POSSESSIONS	borrow	verb	All countries borrow things from other countries and change them.
	claim	verb	Japan claims tempura as part of its cuisine.
	lose	verb	We are not losing our culture.
	shared	adjective	It's a shared cultural entity.
	share	verb	We are the same nationality, so we share the same culture and background.
	steal	verb	We feel as though they are stealing our culture.
	take	verb	America is happy to go around the world taking cultural entities and bringing them home.
MORE OF AN ABSTRACT THING IS AN INCREASE IN SIZE OR HEIGHT	boom	noun	The disaster created a boom in the construction industry.
	boom	verb	The Japan music scene is booming.
	enormous	adjective	Japanese culture is made up of an enormous amount of imported ideas, products, and even festivals.
	explode	verb	Hip-hop music and fashion arrived in Japan in the 1980s, and it has exploded since then.
	explosion	noun	There has been an explosion in fuel prices.
	huge	adjective	Making it to the top of the rankings is a huge endeavour.
(*continued*)	immense	adjective	He has immense popularity among young people.

Metaphor theme	Metaphor-related word	POS	Example
MORE OF AN ABSTRACT THING IS AN INCREASE IN SIZE OR HEIGHT (*cont*.)	massive	adjective	This group has developed a massive following.
	mushroom	verb	The number of aesthetic salons for men has mushroomed in recent years.
	rocket	verb	After a TV program reported on its health benefits, the demand for natto rocketed.
	skyrocket	verb	His popularity skyrocketed.
	snowball	verb	Lolita fashion originated in the 1970s in Japan, but its popularity snowballed in the 1990s when some musicians began dressing this way.
	soar	verb	Since its first release in 2004, the popularity of Vocaloid music has soared.
	spiral	verb	Prices have spiralled.
MORE IS UP AND LESS IS DOWN	bottom out	phrasal verb	The birth rate fell for a long time, and finally bottomed out in 2005.
	crash	verb	The value of the company's stock crashed after the announcement.
	dive	noun	His popularity took a dive after he was arrested.
	drift along	phrasal verb	Unemployment has been drifting along at around 8% for a long time.
	drop	noun	There was a slight drop in the birth rate.
	escalate	verb	Three days later, her anger has escalated.
	escalation	noun	There has been an escalation in crime.
	fall	verb	The Japanese population rose throughout the 20th century, then began to fall in the mid 2000s.
	go down	phrasal verb	The birth rate went down for a while.
	go up	phrasal verb	The price of oil often goes up during a war.
	hit rock bottom	verb phrase	The government's ratings hit rock bottom after the scandals.
	leap	noun	There was a leap in the exam scores.
	mount	verb	Now we have mounting prices.
	peak	noun	The economy peaked, and now we're in the post-peak period.
	peak	verb	The economy peaked, and now we're in the post-peak period.
	plateau	verb	The values rise and then plateau towards the end.
	plummet	verb	After the announcement, the company's stock price plummeted.
	plunge	verb	As winter began, the temperature plunged.
	rise	verb	The Japanese population rose throughout the 20th century, then began to fall in the mid 2000s.
	shoot up	phrasal verb	After study abroad, my TOEIC score shot up.

(*continued*)

Metaphor theme	Metaphor-related word	POS	Example
MORE IS UP AND LESS IS DOWN (*cont*.)	surge	noun	We have had a surge in petty crime.
	surge	verb	The value of the yen surged last year.
	tumble	verb	The number of people who eat rice is tumbling.
STRONG EMOTIONS ARE SICKNESS/INSANITY	crazy	adjective	I know people who are crazy about Gothic Lolita.
	fanatical	adjective	Japanese fans are truly fanatical in their devotion to their heroes.
	fever	noun	All of Japan caught the Hatsune Miku fever.
BEING INTERESTED IS ENTERING SOMETHING	absorbed in	adjective	I was so absorbed in the novel that I didn't hear the phone ring.
	get into	phrasal verb	When you get into this sport, you realize it's very complex.
	immersed in	adjective	It isn't healthy to be too immersed in your job.
STRONG EMOTION IS PHYSICAL FORCE	blow	noun	Failing at university was a severe blow.
	blow away	phrasal verb	He lives in Tokyo and is blown away by the Japanese music scene.
	carried away	phrase	He was so carried away that he broke into tears.
	move	verb	Their music moves us.
	strike	verb	He was struck by sudden shame.
	take by storm	verb phrase	They took the fashion world by storm.
ABSTRACT ENTITIES ARE PHYSICAL STRUCTURES	base	noun	Sumo's strong base ensures its survival.
	base	verb	What did you base your decision to come to this university on?
	cement	verb	The aim of the president's visit was to cement relations between the two countries.
	erode	verb	Sumo's popularity has been eroded by several scandals.
	foundation	noun	The foundation of a culture doesn't change even if it takes things in from overseas.
	link	noun	Sumo demonstrates clear links to other elements of Japanese culture.
	link	verb	Sumo has always been inextricably linked to this ritual.
	support	noun	The government provided support to local businesses.
	support	verb	They supported the weaving trades.
	undermine	verb	This has undermined our efforts to protect the environment.

Appendix H. Untaught Target Metaphors

Examples are based on samples from course input and output, but have been modified for clarity and brevity. Those examples with * are based on marked usage produced by learners.

Metaphor theme	Metaphor-related word	POS	Example
MOVING FORWARD IS IMPROVEMENT	leading	adjective	Two leading economists reported on changing patterns of musical influence.
	momentum	noun	Western music has always been popular, but other countries' music is gaining momentum.
	overtake	verb	Sports coaches expect that their pupils will overtake them in terms of ability.
	resistance	noun	Sumo had to show resistance to government pressure to modernize.
	road	noun	This was the only road that they could follow in their life.
	step	noun	An important step in sumo's history was when Akebono became the first foreign yokozuna.
	withdraw	verb	Otaku cannot deal with the real world and tend to withdraw from it.
MOVEMENT IS FREEDOM	bind	verb	Freeters are not bound by contracts with a company.
	stick	verb	Modern weddings show us that Japanese people are not sticking to traditional styles.
Sources of metaphor: The body	face	noun	There is a bad face to hybridism.
	face	verb	Sumo still faces challenges as it has in the past.
Metaphorical adjectives	clear	adjective	Sumo demonstrates clear links to other elements of Japanese culture.
	close	adjective	They depend on their parents or people who are close to them.
	comfortable	adjective	Compared to the Meiji era, our lives are very convenient and comfortable.
	coming	adjective	We need to preserve Japanese culture for the coming ages.
	following	adjective	This is explained in the following section.
	free	adjective	Japanese musicians value free thinking.
	fresh	adjective	Every day in Makiko's life was fresh and important.
	full	adjective	Japanese wedding ceremonies are full of ambiguities.
	great	adjective	This was a great achievement for Japan.
	hard	adjective	People in the Meiji era had a hard life.
	kind	adjective	Life has been kind to me.
	large	adjective	They used a large amount of money.

(*continued*)

Metaphor theme	Metaphor-related word	POS	Example
Metaphorical adjectives (*cont*.)	little	adjective	The documentary had a little impact on me.
	live	adjective	Many fans go to her live concerts.
	lost	adjective	Sumo has to regain its lost audience.
	middle	adjective	They are part of the middle class.
	natural	adjective	This wedding style will become natural.
	near	adjective	This style will spread throughout Japan in the near future.
	poor	adjective	He put up a poor fight.
	powerful	adjective	Animated musicians cannot give powerful performances.
	related	adjective	Lolita fashion and Otaku are related to Japanese anime culture.
	rich	adjective	The ceremony is carried out to pray for a rich harvest.
	strong	adjective	Many Japanese have a strong desire to be like Westerners.
	uncomfortable	adjective	I don't feel uncomfortable when I hear Japanese singers singing in English.
	young	adjective	This is a young fashion trend.
CULTURES ARE LIVING THINGS	ancestor	noun	Jeu de paume was a direct ancestor of real tennis.
	birth	noun	The Meiji era saw the birth of many new cultural practices in Japan.
	birthplace	noun	The birthplace of karate is in fact Okinawa.
	born	verb	Taco rice was born in Okinawa in the 1980s.
	bring up	phrasal verb	This is a culture that has been brought up to see itself as unique.
	grow up	phrasal verb	Japanese culture has grown up in this way.
	growth	noun	That gave a big boost to Kyoto's post-war economic growth.
	habitat	noun	This is the habitat of the otaku.
	recover	verb	Can our culture recover from this experience?
	revitalize	verb	Japanese people are revitalizing Western cultural entities.
	survival	noun	Like the sport of sumo itself, the dohyou's strength ensures its survival.
	survive	verb	Sumo has had to be nimble to survive.
CULTURES ARE SPACES	access	verb	People can now easily access Western music through various forms of media.
	borderline	noun	This style is on the borderline between Otaku and non-Otaku culture.
	bring into	phrasal verb	Karate developed in Okinawa, and then it was brought into Japan.
	contain	verb	All cultures contain hybrid elements.
	deport	verb	After breaking the rules, he was deported* from sumo society.
(*continued*)	exporter	noun	The US is a major exporter of culture.

Metaphor theme	Metaphor-related word	POS	Example
CULTURES ARE SPACES (*cont*.)	go into	phrasal verb	Entities are often hybridized when they go into other cultures.
	intake	verb	Japan intakes* the US's good points, and produces new things.
	leave	verb	Things can come into that culture, or things might leave that culture.
	outside	adjective	This is an outside view of your culture.
	penetrate	verb	Foreign music no longer penetrates global charts as it used to.
	take in	phrasal verb	These martial arts were taken into Japan.
	within	preposition	We're talking about high culture within Japanese culture.
CULTURES ARE SUBSTANCES	combination	noun	It's an interesting combination of fashion styles.
	combine	verb	Their style of hip-hop combines humour, music, and poetry.
	derive	verb	Spaghetti Neapolitan is derived from several countries' cultures.
	flood	verb	Modern Japan is being flooded with cultural influences.
	form	noun	Westerners are often surprised to see how cultural forms are adapted in Japan.
	form	verb	Individual cultural entities are formed by hybridization.
	permeate	verb	The American lifestyle has permeated our lifestyle.
	refine	verb	Rules were developed and sumo was refined until it became a popular sport.
	refined	adjective	Kyoto is known for its refined culture.
	spread out	phrasal verb	Western culture spread out* in the Meiji era.
CULTURES ARE CONSTRUCTIONS	break	verb	Sumo came under pressure to change, which could have broken* its style.
	destroy	verb	Japanese don't want their society to be destroyed.
	destruction	noun	Hybridization causes the destruction of culture.
	reconstruct	verb	America helped to reconstruct Japanese political systems and society after the war.
	restoration	noun	This will lead to the restoration of Japanese wedding culture.
CULTURES HAVE MANY PARTS	add	verb	I feel like we're adding things to our culture.
	piece	noun	Traditional weddings might disappear as one piece of Japanese culture.
CULTURES ARE POSSESSIONS	accept	verb	It is difficult to accept a different culture.
(*continued*)	adopt	verb	The documentary focuses on how Western culture is adopted and hybridized in Japan.

Metaphor theme	Metaphor-related word	POS	Example
CULTURES ARE POSSESSIONS (*cont.*)	adoption	noun	Donald Ritchie talks about Japanese adoption of foreign cultures.
	exchange	noun	Cultural exchanges are helping to break down barriers.
	exchange	verb	Japanese people realize that exchanging cultural entities helps countries to develop,
	give	verb	Japan is being given a new culture by other countries.
	hand down	phrasal verb	We should hand down our culture.
	have	verb	Japan has a lot of popular culture.
	inherit	verb	We have inherited Meiji era styles.
	keep	verb	We can keep our culture.
	pass	verb	Japanese should pass* their traditions to the future.
	possess	verb	These countries possess similar cultures.
	receive	verb	Japan has received* the influence of music from the West.
	retain	verb	Sumo has always tried to retain the essence of its traditions.
	save	verb	Japanese society has changed enormously, and we must try to save our unique culture.
MORE OF AN ABSTRACT THING IS AN INCREASE IN SIZE OR HEIGHT	broaden	verb	Even less well known countries can become famous if we broaden our tastes.
	deepen	verb	The interchange between Japan and other countries deepened.
	expand	verb	Sumo expanded in Japan as the sport of samurai.
	extend	verb	It is very good to extend the culture of Japan.
	shrink	verb	Rather than shrinking the globe, pop music has encouraged the growth of indigenous music.
	slump	noun	Both countries have had a slump in their musical output.
MORE IS UP AND LESS IS DOWN	above	preposition	Britain's share of global music sales were 52% above its share of world gross GDP.
	fall down	phrasal verb	The popularity of Mt. Fuji has fallen down*.
	upward	adjective	We have seen an upward trend in the value of the yen.
BEING INTERESTED IS ENTERING SOMETHING	into	preposition	Some people are really into one subject and know a lot about it.
STRONG EMOTION IS PHYSICAL FORCE	impact	noun	The history of Japan has had a strong impact on people's lives.

Metaphor theme	Metaphor-related word	POS	Example
ABSTRACT ENTITIES ARE PHYSICAL STRUCTURES	attach	verb	With Japanese kimonos, greater importance is attached to design than to usefulness.
	build	verb	Tension builds as the sumo tournament approaches.
	concrete	adjective	This word does not have a concrete meaning.
	connect	verb	Quite extreme elements of society are connected through hip-hop music.
	connection	verb	This is an example of karate's connection to China.
	structure	noun	This may change the structure of the music industry.

Appendix I. Output corpora data

The tables below provide the total number of lexical units and mean number of lexical units per week produced by each participant. Because the data were coded under two systems (section 3.4.2.2, page 68), counts for both systems are provided.

1. Data coded following the MIPVU procedure for identifying lexical units (chapter four and chapter five, sections 5.3 - 5.5)

Control group			Experimental group		
Code	Total lexical units	Per week	Code	Total lexical units	Per week
CON1	2715	181.00	EXP1	2732	182.13
CON2	1860	124.00	EXP2	2527	168.47
CON3	942	62.80	EXP3	1927	128.47
CON4	2090	139.33	EXP4	2681	178.73
CON5	2480	165.33	EXP5	3097	206.47
CON6	2272	151.47	EXP6	2648	176.53
CON7	1421	94.73	EXP7	2322	154.80
CON8	3098	206.53	EXP8	2657	177.13
CON9	1124	74.93	EXP9	2804	186.93
CON10	1970	131.33	EXP10	3700	246.67
CON11	1989	132.60	EXP11	2749	183.27
CON12	1507	100.47	EXP12	2073	138.20
CON13	1372	91.47	EXP13	1597	106.47
CON14	2245	149.67	EXP14	2192	146.13
CON15	1348	89.87	EXP15	2072	138.13
CON16	1282	85.47	EXP16	2690	179.33
CON17	2525	168.33	EXP17	2387	159.13
CON18	1720	114.67	EXP18	2285	152.33
CON19	1002	66.80	EXP19	2700	180.00
CON20	2017	134.47	EXP20	2435	162.33
CON21	2474	164.93	EXP21	2096	139.73
CON22	2048	136.53	EXP22	2718	181.20
CON23	2666	177.73	EXP23	2111	140.73

2. Data coded with alternative procedure, in which polywords appearing in the Macmillan dictionary were recorded as lexical units (chapter five, section 5.6 and chapters 6 - 7)

Control group			Experimental group		
Code	**Total lexical units**	**Per week**	**Code**	**Total lexical units**	**Per week**
CON1	2693	179.53	EXP1	2726	181.73
CON2	1854	123.60	EXP2	2524	168.27
CON3	942	62.80	EXP3	1899	126.60
CON4	2082	138.80	EXP4	2678	178.53
CON5	2466	164.40	EXP5	3079	205.27
CON6	2254	150.27	EXP6	2646	176.40
CON7	1410	94.00	EXP7	2307	153.80
CON8	3082	205.47	EXP8	2655	177.00
CON9	1118	74.53	EXP9	2797	186.47
CON10	1958	130.53	EXP10	3673	244.87
CON11	1968	131.20	EXP11	2737	182.47
CON12	1504	100.27	EXP12	2056	137.07
CON13	1362	90.80	EXP13	1592	106.13
CON14	2231	148.73	EXP14	2179	145.27
CON15	1342	89.47	EXP15	2055	137.00
CON16	1269	84.60	EXP16	2686	179.07
CON17	2482	165.47	EXP17	2379	158.60
CON18	1709	113.93	EXP18	2271	151.40
CON19	986	65.73	EXP19	2694	179.60
CON20	1996	133.07	EXP20	2428	161.87
CON21	2459	163.93	EXP21	2089	139.27
CON22	2032	135.47	EXP22	2716	181.07
CON23	2631	175.40	EXP23	2106	140.40

The tables below detail the mean number of lexical units per writing topic. Values under both counting systems are provided.

1. Data coded following the MIPVU procedure for identifying lexical units (chapter four and chapter five, sections 5.3 - 5.5)

Topic	Control group	Experimental group
1	87.52	148.05
2	84.09	150.65
3	115.04	148.18
4	107.04	164.39
5	123.59	152.39
6	123.13	153.74
7	109.50	142.57
8	116.67	161.22
9	118.10	155.00
10	259.00	275.91
11	116.87	150.70
12	161.39	161.96
13	126.48	149.13
14	107.61	145.96
15	194.83	240.00

2. Data coded with alternative procedure, in which polywords appearing in the Macmillan dictionary were recorded as lexical units (chapter five, section 5.6 and chapters 6 - 7)

Topic	Control group	Experimental group
1	87.09	141.13
2	83.30	150.39
3	114.22	141.04
4	106.52	163.78
5	117.70	151.65
6	121.87	152.83
7	103.70	141.52
8	105.65	160.91
9	107.22	154.70
10	257.17	275.26
11	115.78	149.09
12	159.78	161.52
13	125.35	148.74
14	107.35	145.43
15	192.96	239.04

Appendix J. MIPVU coding data

1. The following tables display the raw counts of parts of speech produced by each participant under the data coding system used in chapter four and sections 5.3 - 5.5 of chapter five.

Code	Total nouns	Total verbs	Total phrasal verbs	Total adjectives	Total adverbs	Total conjunctions	Total determiners	Total prepositions	Total other	Total lexical units
CON1	724	424	9	191	146	202	234	317	468	2715
CON2	557	291	6	191	111	142	123	173	266	1860
CON3	324	155	2	95	37	51	49	89	140	942
CON4	547	364	0	189	123	189	114	188	376	2090
CON5	758	302	2	185	159	160	241	330	343	2480
CON6	645	327	2	198	119	174	218	248	341	2272
CON7	380	246	3	122	82	102	113	123	250	1421
CON8	766	468	10	231	191	232	356	350	494	3098
CON9	319	195	7	112	76	62	73	95	185	1124
CON10	550	284	8	183	120	153	174	171	327	1970
CON11	521	294	2	173	114	139	179	215	352	1989
CON12	462	208	5	111	105	99	160	169	188	1507
CON13	345	231	3	118	80	124	108	126	237	1372
CON14	620	331	2	198	109	162	199	255	369	2245
CON15	389	208	6	121	85	85	116	143	195	1348
CON16	335	196	3	128	85	90	117	132	196	1282
CON17	662	392	11	223	122	163	299	221	432	2525
CON18	536	231	2	134	116	135	131	194	241	1720
CON19	264	158	0	83	64	76	80	98	179	1002
CON20	559	303	4	158	124	153	168	207	341	2017
CON21	716	357	1	271	127	172	250	273	307	2474
CON22	581	308	0	225	115	167	109	219	324	2048
CON23	778	352	4	238	166	133	235	331	429	2666

Code	Total nouns	Total verbs	Total phrasal verbs	Total adjectives	Total adverbs	Total conjunctions	Total determiners	Total prepositions	Total other	Total lexical units
EXP1	741	422	8	266	138	172	275	249	461	2732
EXP2	751	364	11	237	120	174	193	303	374	2527
EXP3	574	274	1	205	104	130	149	205	285	1927
EXP4	767	390	9	260	149	204	209	243	450	2681
EXP5	847	463	6	325	146	205	255	328	522	3097
EXP6	829	412	15	209	125	249	168	235	406	2648
EXP7	514	357	1	287	172	163	144	235	449	2322
EXP8	634	459	7	233	120	209	198	254	543	2657
EXP9	650	543	5	226	96	238	198	235	613	2804
EXP10	1006	535	10	312	129	329	313	442	624	3700
EXP11	622	476	4	236	173	239	213	243	543	2749
EXP12	582	348	2	257	102	129	138	158	357	2073
EXP13	491	254	1	156	91	74	114	174	242	1597
EXP14	566	336	11	183	125	129	138	229	475	2192
EXP15	543	358	6	177	133	135	144	204	372	2072
EXP16	770	388	5	259	196	175	195	279	423	2690
EXP17	650	418	4	212	137	169	158	212	427	2387
EXP18	656	336	4	255	153	134	140	230	377	2285
EXP19	659	415	4	258	186	243	199	262	474	2700
EXP20	665	393	6	306	140	172	154	201	398	2435
EXP21	587	349	5	209	93	154	156	200	343	2096
EXP22	660	486	2	259	161	220	226	248	456	2718
EXP23	504	359	2	227	119	137	141	219	403	2111

2. The following tables display the raw counts of MRW parts of speech produced under the data coding system used in chapter four and sections 5.3 - 5.5 of chapter five.

Code	MRW nouns	MRW verbs	MRW phrasal verbs	MRW adjectives	MRW adverbs	MRW conjunctions	MRW determiners	MRW prepositions	MRW other	MRW total
CON1	83	127	3	16	4	0	21	146	10	410
CON2	29	49	6	8	0	1	9	63	3	168
CON3	7	23	1	5	0	0	2	45	0	83
CON4	40	71	0	21	0	0	6	69	26	233
CON5	54	84	2	19	7	1	32	126	19	344
CON6	37	79	1	9	1	0	34	98	12	271
CON7	33	53	3	10	0	0	7	44	8	158
CON8	64	92	6	37	8	1	29	135	26	398
CON9	22	30	7	7	1	0	4	37	6	114
CON10	42	57	6	10	2	0	17	80	17	231
CON11	33	59	2	11	1	0	23	93	11	233
CON12	28	42	4	9	4	0	6	57	6	156
CON13	27	49	2	3	3	0	14	47	4	149
CON14	39	86	0	38	0	0	27	106	12	308
CON15	23	36	2	1	0	0	5	56	4	127
CON16	27	33	3	13	5	0	10	50	15	156
CON17	60	83	2	10	5	0	40	106	8	314
CON18	22	34	2	12	1	0	12	67	6	156
CON19	19	23	0	6	1	0	3	48	3	103
CON20	45	80	2	16	2	1	16	82	8	252
CON21	45	103	1	13	1	0	16	111	6	296
CON22	26	72	0	12	4	0	10	100	9	233
CON23	87	75	3	17	3	0	23	156	19	383

Code	MRW nouns	MRW verbs	MRW phrasal verbs	MRW adjectives	MRW adverbs	MRW conjunctions	MRW determiners	MRW prepositions	MRW other	MRW total
EXP1	48	103	6	33	2	0	16	94	23	325
EXP2	84	89	9	14	3	0	26	125	16	366
EXP3	32	87	0	22	0	0	2	84	2	229
EXP4	42	88	5	16	0	0	30	95	6	282
EXP5	57	119	3	30	4	0	15	156	19	403
EXP6	39	75	3	19	3	0	25	85	8	257
EXP7	32	62	1	25	5	1	17	77	18	238
EXP8	69	90	4	24	1	3	23	122	26	362
EXP9	65	92	5	33	2	1	14	105	30	347
EXP10	111	124	7	46	5	0	31	161	22	507
EXP11	62	106	4	16	5	1	15	102	30	341
EXP12	43	65	1	17	1	0	13	48	7	195
EXP13	30	64	0	10	3	0	18	85	20	230
EXP14	36	54	9	10	1	0	10	99	7	226
EXP15	46	60	3	11	1	0	18	94	19	252
EXP16	74	59	1	31	4	0	15	89	11	284
EXP17	23	80	2	13	2	0	20	76	18	234
EXP18	42	90	4	11	0	0	15	77	10	249
EXP19	40	98	1	22	5	2	17	112	12	309
EXP20	38	120	6	22	2	0	35	77	29	329
EXP21	31	56	2	14	5	0	16	89	4	217
EXP22	43	113	1	21	5	0	6	92	20	301
EXP23	44	70	2	23	3	0	8	88	9	247

3. The following tables display the raw counts of each category of MRW produced under the data coding system used in chapter four and sections 5.3 - 5.5 of chapter five.

Code	Indirect MRWs	Direct MRWs	Implicit MRWs	Total MRWs	Taught target MRWs	Untaught target MRWs	Combined target MRWs	OC non-target MRWS	Total lexical unit types	Taught target MRW types	Untaught target MRW types	Combined target MRW types
CON1	403	1	6	410	31	19	50	183	850	22	14	36
CON2	167	0	1	168	11	15	26	66	545	9	12	21
CON3	83	0	0	83	6	4	10	26	351	5	3	8
CON4	232	0	1	233	11	13	24	108	589	5	9	14
CON5	340	0	4	344	21	18	39	127	853	13	15	28
CON6	267	0	4	271	20	17	37	90	639	17	12	29
CON7	157	0	1	158	13	21	34	65	519	12	11	23
CON8	393	0	5	398	32	25	57	150	859	29	15	44
CON9	113	0	1	114	15	8	23	44	393	10	6	16
CON10	230	0	1	231	11	7	18	99	582	10	5	15
CON11	233	0	0	233	20	10	30	76	615	16	9	25
CON12	155	0	1	156	10	15	25	62	560	9	13	22
CON13	148	0	1	149	8	6	14	70	482	8	6	14
CON14	302	0	6	308	21	28	49	114	621	15	18	33
CON15	126	0	1	127	9	6	15	47	461	5	4	9
CON16	143	1	12	156	18	8	26	55	463	15	5	20
CON17	309	1	4	314	22	10	32	128	734	15	6	21
CON18	156	0	0	156	12	7	19	52	532	10	6	16
CON19	103	0	0	103	7	6	13	36	376	5	6	11
CON20	250	0	2	252	24	14	38	107	582	18	10	28
CON21	293	2	1	296	30	12	42	121	647	21	12	33
CON22	228	3	2	233	10	16	26	88	623	9	10	19
CON23	378	0	5	383	23	29	52	133	718	14	17	31

Code	Indirect MRWs	Direct MRWs	Implicit MRWs	Total MRWs	Taught target MRWs	Untaught target MRWs	Combined target MRWs	OC non-target MRWS	Total lexical unit types	Taught target MRW types	Untaught target MRW types	Combined target MRW types
EXP1	306	0	19	325	24	30	54	138	800	19	20	39
EXP2	353	4	9	366	36	15	51	148	721	28	13	41
EXP3	229	0	0	229	9	21	30	111	597	8	10	18
EXP4	279	1	2	282	29	10	39	112	691	21	8	29
EXP5	384	1	18	403	25	29	54	159	792	20	18	38
EXP6	257	0	0	257	27	23	50	89	731	21	8	29
EXP7	232	1	5	238	20	11	31	94	503	13	9	22
EXP8	343	0	19	362	41	8	49	139	697	25	6	31
EXP9	341	3	3	347	33	19	52	145	644	18	7	25
EXP10	499	0	8	507	45	35	80	213	911	25	12	37
EXP11	337	0	4	341	34	14	48	145	757	25	13	38
EXP12	191	0	4	195	24	15	39	88	546	13	8	21
EXP13	230	0	0	230	23	15	38	69	533	14	8	22
EXP14	223	3	0	226	23	6	29	81	646	16	6	22
EXP15	248	4	0	252	11	10	21	100	568	9	7	16
EXP16	279	2	3	284	28	20	48	121	756	15	16	31
EXP17	226	0	8	234	24	9	33	87	668	16	8	24
EXP18	248	0	1	249	23	12	35	112	549	11	9	20
EXP19	297	9	3	309	22	22	44	122	751	19	15	34
EXP20	319	0	10	329	25	33	58	130	617	18	21	39
EXP21	216	0	1	217	24	11	35	73	521	12	8	20
EXP22	301	0	0	301	34	19	53	130	657	25	13	38
EXP23	242	0	5	247	15	18	33	109	602	10	11	21

4. The following table displays the total number of lexical units and MRWs in each category for each writing topic.

Topic	Total lexical units		OC lexical units		CC lexical units		Total MRW		OC MRW		CC MRW		Taught target MRW		Untaught target MRW	
	Con.	Exp.	Con.	Exp.	Con.	Exp.	Con.	Exp.	Con.	Exp.	Con.	Exp.	Con.	Exp.	Con.	Exp.
1	2013	3257	1205	1937	808	1320	216	314	121	161	95	153	16	20	16	18
2	1934	3465	1105	2056	829	1409	209	395	109	226	100	169	11	32	11	33
3	2646	3260	1515	1848	1131	1412	315	396	177	228	138	168	21	29	21	29
4	2462	3781	1295	1973	1167	1808	357	620	220	380	137	240	18	46	12	14
5	2719	3505	1596	2047	1123	1458	313	321	185	167	128	154	36	20	31	13
6	2832	3536	1594	1938	1238	1598	335	342	136	147	199	195	6	12	9	24
7	2409	3279	1479	2003	930	1276	271	364	143	209	128	155	19	23	7	14
8	2450	3708	1367	2045	1083	1663	341	474	165	270	176	204	13	40	10	33
9	2480	3565	1348	2148	1132	1417	361	447	177	245	184	202	19	48	28	34
10	5957	6346	3415	3607	2542	2739	636	653	294	314	342	339	58	74	46	27
11	2688	3466	1585	2075	1103	1391	249	391	129	194	120	197	15	18	14	21
12	3712	3725	2174	2180	1538	1545	497	444	286	254	211	190	53	29	29	32
13	2909	3430	1758	2031	1151	1399	314	378	159	210	155	168	33	46	10	19
14	2475	3357	1458	2024	1017	1333	297	489	175	334	122	155	20	91	30	42
15	1718	2085	983	1157	735	928	198	235	69	107	129	128	14	9	8	14
16	1899	1983	1111	1225	788	758	248	275	114	160	134	115	17	36	13	16
17	864	1452	521	900	343	552	119	192	87	113	32	79	15	20	15	16

5. The following tables display the raw counts of parts of speech produced by each participant under the data coding system used in section 5.6 of chapter five and chapters 6 - 7.

Code	Total nouns	Total verbs	Total phrasal verbs	Total adjectives	Total adverbs	Total conjunctions	Total determiners	Total prepositions	Total other	Total lexical units
CON1	712	424	9	191	151	205	230	309	462	2693
CON2	554	291	6	191	113	143	121	170	265	1854
CON3	324	155	2	95	37	51	49	89	140	942
CON4	544	364	0	188	126	189	110	185	376	2082
CON5	749	302	2	185	167	160	237	322	342	2466
CON6	636	327	2	198	126	174	210	241	340	2254
CON7	375	246	3	122	87	102	106	119	250	1410
CON8	760	468	10	230	194	233	350	344	493	3082
CON9	317	195	7	112	78	62	69	93	185	1118
CON10	542	284	8	183	126	153	171	164	327	1958
CON11	510	294	2	173	125	139	169	204	352	1968
CON12	461	208	5	111	105	99	159	168	188	1504
CON13	337	231	3	118	86	124	105	121	237	1362
CON14	610	331	2	198	119	162	193	247	369	2231
CON15	385	208	6	121	89	85	113	140	195	1342
CON16	328	196	3	128	90	90	109	129	196	1269
CON17	645	391	11	224	135	163	275	207	431	2482
CON18	531	231	2	134	121	135	125	189	241	1709
CON19	258	158	0	83	70	76	70	92	179	986
CON20	551	303	4	158	132	153	155	199	341	1996
CON21	711	357	1	271	132	172	240	268	307	2459
CON22	575	308	0	225	120	167	99	214	324	2032
CON23	763	352	4	238	174	138	222	316	424	2631

Code	Total nouns	Total verbs	Total phrasal verbs	Total adjectives	Total adverbs	Total conjunctions	Total determiners	Total prepositions	Total other	Total lexical units
EXP1	739	422	8	264	140	172	273	247	461	2726
EXP2	748	364	11	237	122	174	191	303	374	2524
EXP3	562	273	1	198	116	130	141	193	285	1899
EXP4	764	390	9	260	151	204	207	243	450	2678
EXP5	839	463	6	320	153	205	250	321	522	3079
EXP6	827	412	15	209	127	249	168	233	406	2646
EXP7	509	357	1	282	177	163	139	230	449	2307
EXP8	632	459	7	233	122	209	198	252	543	2655
EXP9	643	543	5	226	99	238	195	235	613	2797
EXP10	995	535	10	305	140	329	303	432	624	3673
EXP11	615	476	4	234	174	239	212	240	543	2737
EXP12	569	348	2	255	115	129	132	149	357	2056
EXP13	486	254	1	156	96	74	111	172	242	1592
EXP14	557	336	11	181	128	129	135	227	475	2179
EXP15	536	358	6	172	140	135	138	198	372	2055
EXP16	767	388	5	259	199	175	194	276	423	2686
EXP17	646	418	4	210	141	169	155	209	427	2379
EXP18	648	336	4	252	161	134	137	222	377	2271
EXP19	656	415	4	257	188	244	198	259	473	2694
EXP20	660	393	6	305	145	172	150	199	398	2428
EXP21	584	349	5	207	96	154	153	198	343	2089
EXP22	659	486	2	259	163	220	226	246	455	2716
EXP23	503	359	2	226	121	137	141	217	400	2106

6. The following tables display the raw counts of MRW parts of speech produced under the data coding system used in section 5.6 of chapter five and chapters 6 - 7.

Code	MRW nouns	MRW verbs	MRW phrasal verbs	MRW adjectives	MRW adverbs	MRW conjunctions	MRW determiners	MRW prepositions	MRW other	MRW total
CON1	73	127	3	16	17	3	21	145	10	415
CON2	27	49	6	8	12	4	9	60	3	178
CON3	7	23	1	5	5	0	2	47	0	90
CON4	38	71	0	21	15	0	6	69	26	246
CON5	52	84	2	19	24	2	32	123	19	357
CON6	35	79	1	10	12	1	31	102	12	283
CON7	30	53	3	10	18	1	6	41	8	170
CON8	59	92	6	36	18	3	29	132	26	401
CON9	20	30	7	7	5	0	4	35	6	114
CON10	41	57	6	10	19	0	17	74	17	241
CON11	28	59	2	11	18	1	23	83	11	236
CON12	27	42	4	9	10	1	6	61	6	166
CON13	25	49	2	3	17	0	13	45	4	158
CON14	37	86	0	38	18	0	25	100	12	316
CON15	22	36	2	1	11	0	4	53	4	133
CON16	24	33	3	13	14	0	6	48	15	156
CON17	45	83	2	11	24	0	39	100	8	312
CON18	19	34	2	12	6	0	12	62	6	153
CON19	14	23	0	6	13	0	3	43	3	105
CON20	40	80	2	16	16	1	16	76	8	255
CON21	40	103	1	13	20	0	16	107	6	306
CON22	21	72	0	12	16	1	10	98	9	239
CON23	77	75	3	17	25	5	23	145	19	389

Code	MRW nouns	MRW verbs	MRW phrasal verbs	MRW adjectives	MRW adverbs	MRW conjunctions	MRW determiners	MRW prepositions	MRW other	MRW total
EXP1	46	102	6	33	6	0	16	94	23	326
EXP2	84	89	9	14	11	0	24	126	16	373
EXP3	24	87	0	22	13	0	2	74	2	224
EXP4	41	88	5	16	4	0	28	96	6	284
EXP5	52	119	3	30	21	0	15	150	19	409
EXP6	40	75	3	19	5	0	25	85	8	260
EXP7	27	62	1	25	18	1	17	76	18	245
EXP8	70	90	4	24	5	3	23	123	26	368
EXP9	65	92	5	33	9	1	11	106	30	352
EXP10	102	124	7	46	21	0	30	160	22	512
EXP11	56	106	4	16	15	1	15	102	30	345
EXP12	41	65	1	17	15	0	9	40	7	195
EXP13	30	64	0	10	11	0	15	83	20	233
EXP14	28	54	9	10	5	0	9	102	7	224
EXP15	41	60	3	11	14	0	17	89	19	254
EXP16	73	60	1	31	7	0	15	87	11	285
EXP17	21	80	2	13	13	0	19	78	18	244
EXP18	39	90	4	11	21	0	15	71	8	259
EXP19	38	98	1	22	8	3	17	110	12	309
EXP20	37	120	6	22	12	0	32	77	29	335
EXP21	29	56	2	14	10	0	15	87	4	217
EXP22	43	113	1	21	18	0	6	90	20	312
EXP23	44	70	2	23	9	0	8	87	9	252

7. The following tables display the raw counts of MRWs categorised by rhetorical function that were discussed in section 5.6 of chapter five.

Code	Evaluative: Emotive	Evaluative: Emphasis	Evaluative: General	Evaluative: Mitigation	All evaluative	Interpersonal: Advising	Interpersonal: Manipulative	All interpersonal	Textual: Reiteration	Textual: Structuring	All textual	OC Unmarked conventional
CON1	0	3	0	2	5	1	0	1	0	8	8	31
CON2	0	16	3	2	21	0	0	0	0	47	47	111
CON3	2	20	3	1	26	0	0	0	0	74	74	139
CON4	1	13	1	1	16	0	0	0	0	66	66	109
CON5	0	4	7	0	11	0	0	0	0	34	34	85
CON6	3	33	7	6	49	0	0	0	0	72	72	151
CON7	0	8	3	0	11	0	0	0	0	15	15	53
CON8	0	13	0	0	13	0	0	0	0	53	53	103
CON9	0	8	3	2	13	0	0	0	0	53	53	88
CON10	1	5	2	0	8	0	0	0	0	23	23	77
CON11	0	3	1	2	6	0	0	0	0	33	33	77
CON12	1	23	8	5	37	0	0	0	0	57	57	124
CON13	0	3	0	0	3	0	0	0	0	18	18	59
CON14	7	10	7	0	24	0	0	0	0	30	30	55
CON15	4	14	2	1	21	0	0	0	0	72	72	125
CON16	0	7	2	2	11	0	0	0	0	23	23	57
CON17	2	5	1	1	9	0	0	0	0	19	19	35
CON18	1	7	2	5	15	0	0	0	0	40	40	126
CON19	1	12	6	0	19	0	0	0	0	42	42	139
CON20	2	7	4	0	13	0	0	0	0	33	33	97
CON21	5	10	9	1	25	0	0	0	0	72	72	150
CON22	0	3	0	2	5	1	0	1	0	8	8	31
CON23	0	16	3	2	21	0	0	0	0	47	47	111

Code	Evaluative: Emotive	Evaluative: Emphasis	Evaluative: General	Evaluative: Mitigation	All evaluative	Interpersonal: Advising	Interpersonal: Manipulative	All interpersonal	Textual: Reiteration	Textual: Structuring	All textual	OC Unmarked conventional
EXP1	1	17	8	5	31	0	1	1	0	45	45	157
EXP2	2	8	3	2	15	0	0	0	0	50	50	183
EXP3	0	9	13	0	22	0	0	0	0	18	18	111
EXP4	2	11	4	0	17	0	0	0	0	39	39	133
EXP5	1	30	5	1	37	0	0	0	0	53	53	169
EXP6	4	10	5	1	20	0	0	0	0	37	37	119
EXP7	2	14	5	1	22	0	0	0	0	53	53	98
EXP8	7	8	5	0	20	1	0	1	1	59	60	167
EXP9	2	16	15	3	36	0	0	0	0	49	49	162
EXP10	4	21	3	0	28	0	1	1	0	73	73	256
EXP11	2	11	2	2	17	0	0	0	0	61	61	168
EXP12	1	7	5	1	14	0	0	0	0	31	31	111
EXP13	0	9	0	0	9	0	0	0	0	43	43	98
EXP14	2	5	1	1	9	0	0	0	0	25	25	93
EXP15	1	7	1	1	10	0	0	0	0	49	49	106
EXP16	5	17	10	4	36	0	0	0	0	31	31	132
EXP17	2	5	4	0	11	0	0	0	0	53	53	107
EXP18	0	5	4	0	9	0	0	0	0	44	44	135
EXP19	0	14	11	0	25	0	0	0	0	36	36	138
EXP20	3	14	7	0	24	1	0	1	0	74	74	161
EXP21	0	12	5	1	18	1	0	1	0	24	24	87
EXP22	2	16	9	3	30	0	0	0	0	38	38	154
EXP23	0	17	9	1	27	0	0	0	0	24	24	116

Code	Evaluative types	Textual types	OC unmarked conventional types
CON1	15	19	125
CON2	4	11	57
CON3	4	6	23
CON4	14	13	53
CON5	19	24	100
CON6	15	21	72
CON7	6	14	59
CON8	33	21	83
CON9	7	8	39
CON10	10	14	51
CON11	12	13	57
CON12	8	14	59
CON13	6	12	53
CON14	19	18	67
CON15	3	10	29
CON16	17	10	46
CON17	16	16	82
CON18	10	7	38
CON19	7	9	23
CON20	12	16	71
CON21	15	12	78
CON22	11	12	56
CON23	17	17	84

Code	Evaluative types	Textual types	OC unmarked conventional types
EXP1	21	13	99
EXP2	12	15	91
EXP3	9	9	58
EXP4	11	10	75
EXP5	25	14	98
EXP6	15	9	69
EXP7	14	16	56
EXP8	14	16	95
EXP9	15	12	70
EXP10	16	20	119
EXP11	12	18	97
EXP12	11	12	59
EXP13	5	9	58
EXP14	8	11	60
EXP15	7	13	57
EXP16	25	10	87
EXP17	9	17	70
EXP18	7	13	64
EXP19	15	12	92
EXP20	15	14	89
EXP21	11	8	53
EXP22	19	11	77
EXP23	16	13	67

Appendix K. Rhetorical function scatter plot data

Sizes of output corpora used in log-likelihood calculations

Control output corpus	
Total	43830
Low	10554
High	17503

Experimental output corpus	
Total	56972
Low	18482
High	20687

Log-likelihood calculations for rhetorical function scatter plots

The following notes apply to all tables below.

Note: Log-likelihood calculations for conditions based on output frequencies in each corpus relative to the size of the whole corpus.

Note: Log-likelihood calculations for ability based on frequencies in output by learners in the low/high ability groups based on TOEIC scores relative to the size of the combined low/high TOEIC group output (low TOEIC = 29,036 tokens; high TOEIC = 38,190)

a = Combined output frequency from low or high level learners in both conditions.

* = Significant at the 0.05 level; ** = 0.01 level; *** = 0.001 level; **** = 0.0001 level

Metaphorical adjectives used to perform an evaluative function

MRW	Control freq.	Experimental freq.	Low TOEIC freq.[a]	High TOEIC freq.[a]	Log-likelihood (condition)	Log-likelihood (ability)
big	23	48	16	34	3.66	2.63
clear	6	1	2	2	-5.39*	-0.07
comfortable	6	7	1	5	-0.04	1.93
dark	3	9	0	8	1.77	9.05**
deep	6	4	1	3	-1.10	0.57
free	1	5	1	2	1.96	0.12
great	18	25	10	24	0.05	2.74
hard	5	43	21	20	25.32****	-1.07
high	26	38	21	22	0.21	-0.55
large	14	8	9	8	-3.61	-0.65
long	43	40	18	43	-2.32	4.84*
low	9	2	2	7	-6.84**	1.74
natural	12	6	3	9	-3.92*	1.72
short	4	7	2	6	0.23	1.15
strong	18	20	13	17	-0.23	0.00

Metaphors used to perform an evaluative function

MRW	Control freq.	Experimental freq.	Low TOEIC freq.[a]	High TOEIC freq.[a]	Log-likelihood (condition)	Log-likelihood (ability)
big (adj)	23	48	16	31	3.66	1.64
boom (n)	8	11	2	4	0.01	0.24
boom (v)	0	6	0	5	6.87**	5.65*
challenge (n)	9	6	3	3	-1.65	-0.11
clear (adj)	6	1	2	2	-5.39*	-0.07
comfortable (adj)	6	7	1	5	-0.04	1.93
dark (adj)	3	9	0	8	1.77	9.05**
deep (adj)	6	4	1	3	-1.10	0.57
deeply (adv)	7	8	6	4	-0.06	-1.14
dominate (v)	10	2	2	7	-8.13**	1.74
dramatically (adv)	6	2	0	5	-3.28	5.65*
expand (v)	3	7	9	1	0.77	-9.74**
free (adj)	1	5	1	2	1.96	0.12
great (adj)	18	25	10	21	0.05	1.56
hard (adj)	5	43	21	20	25.32****	-1.07
hard (adv)	6	12	6	6	0.77	-0.22
heart (n)	4	5	5	2	0.00	-2.28
high (adj)	26	38	21	21	0.21	-0.79
large (adj)	14	8	9	8	-3.61	-0.65
long (adj)	43	40	18	43	-2.32	4.84*
low (adj)	9	2	2	7	-6.84**	1.74
natural (adj)	12	6	3	9	-3.92*	1.72
power (n)	9	1	0	5	-9.63**	5.65*
pressure (n)	1	7	4	2	3.63	-1.34
short (adj)	4	7	2	6	0.23	1.15
star (n)	3	8	3	3	1.24	-0.11
strong (adj)	18	20	13	16	-0.23	-0.03
strongly (adv)	3	7	0	7	0.77	7.92**
surround (v)	5	0	0	5	-8.31**	5.65*
treasure (n)	1	6	4	2	2.77	-1.34

Metaphors used to perform a text structuring function

MRW	Control freq.	Experimental freq.	Low TOEIC freq.[a]	High TOEIC freq.[a]	Log-likelihood (condition)	Log-likelihood (ability)
according to (prep)	24	18	13	16	-3.15	-0.03
and so on (adv)	51	33	30	11	-10.05**	-15.12***
at first (adv)	19	23	15	13	-0.05	-1.22
at least (adv)	4	8	2	8	0.52	2.40
by the way (adv)	6	2	1	2	-3.28	0.12
in addition (adv)	65	25	21	43	-30.45****	2.89
in case of (prep)	5	0	1	3	-8.31**	0.57
in conclusion (adv)	17	7	6	9	-7.33**	0.06
in contrast (adv)	5	6	1	4	-0.02	1.20
in fact (adv)	12	13	9	8	-0.21	-0.65
in order to (conj)	12	1	1	7	-14.08***	3.57
in short (adv)	10	5	3	8	-3.27	1.19
in summary (adv)	8	7	15	0	-0.59	-25.19****
in terms of (prep)	11	4	0	12	-5.49*	13.57***
in that (conj)	6	0	2	2	-9.97**	-0.07
it (pron)	25	21	6	25	-2.19	7.89**
no longer (adv)	8	3	3	6	-3.86*	0.37
on the other hand (adv)	63	45	32	38	-9.59**	-0.18
thanks to (prep)	10	12	2	8	-0.03	2.40
that (det)	51	28	10	35	-14.18***	8.70**
that (pron)	66	74	13	82	-0.76	38.72****
their (pron)	1	14	8	7	10.29**	-0.62
them (pron)	9	17	3	18	0.85	8.17**
these (det)	98	105	61	72	-1.89	-0.39
these (pron)	34	10	21	15	-20.88****	-3.32
these days (adv)	13	22	10	15	0.58	0.10
they (pron)	6	50	22	21	28.92****	-1.10
this (det)	192	247	104	174	-0.01	3.84
this (pron)	86	167	56	118	9.49**	8.85**
those (det)	13	10	3	11	-1.57	2.93

Metaphorical nouns classified as unmarked conventional

MRW	Control freq.	Experimental freq.	Low TOEIC freq.[a]	High TOEIC freq.[a]	Log-likelihood (condition)	Log-likelihood (ability)
age (n)	14	8	8	12	-3.61	0.08
aspect (n)	15	5	3	9	-8.20**	1.72
atmosphere (n)	1	9	2	5	5.43*	0.64
birthplace (n)	6	3	1	3	-1.96	0.57
case (n)	0	5	0	1	5.72*	1.13
character (n)	1	7	1	3	3.63	0.57
chart (n)	4	10	7	4	1.32	-1.86
dream (n)	4	9	4	2	0.89	-1.34
element (n)	25	15	9	9	-5.83*	-0.34
end (n)	0	9	3	1	10.30**	-1.67
fashion (n)	0	6	0	4	6.87**	4.52*
feature (n)	1	5	0	3	1.96	3.39
feeling (n)	21	35	17	24	0.83	0.05
field (n)	8	1	0	6	-8.19**	6.79**
form (n)	12	12	5	8	-0.41	0.12
hybrid (n)	10	13	6	6	0.00	-0.22
idol (n)	3	21	7	11	10.88***	0.14
image (n)	12	16	5	9	0.00	0.32
impact (n)	2	8	6	4	2.45	-1.14
import (n)	0	5	0	1	5.72*	1.13
impression (n)	7	7	2	7	-0.24	1.74
industry (n)	5	3	0	6	-1.17	6.79**
link (n)	5	0	3	1	-8.31**	-1.67
look (n)	4	5	0	6	0.00	6.79**
market (n)	17	12	8	10	-2.67	-0.01
mixture (n)	0	6	0	3	6.87**	3.39
movement (n)	9	2	1	10	-6.84**	6.29*
part (n)	22	15	8	21	-3.80	3.02
point (n)	46	28	28	25	-10.41**	-1.98
role (n)	8	3	0	5	-3.86*	5.65*
sense (n)	5	1	3	3	-4.06*	-0.11
side (n)	4	8	1	5	0.52	1.93
society (n)	0	5	2	2	5.72*	-0.07
spirit (n)	1	9	1	6	5.43*	2.72
taste (n)	6	9	6	3	0.07	-2.01
thing (n)	217	275	138	183	-0.08	0.01
way (n)	50	57	18	56	-0.46	11.45***
world (n)	8	20	12	5	2.65	-5.21*

Metaphorical verbs classified as unmarked conventional

MRW	Control freq.	Experimental freq.	Low TOEIC freq.[a]	High TOEIC freq.[a]	Log-likelihood (condition)	Log-likelihood (ability)
absorb (v)	24	6	7	13	-16.80****	0.56
accept (v)	14	13	5	17	-0.76	4.04*
add (v)	16	29	18	11	1.17	-4.17*
adopt (v)	8	11	12	4	0.01	-6.68***
appear (v)	27	35	14	18	0.00	0.00
attract (v)	9	11	3	12	-0.02	3.60
born (v)	23	20	15	18	-1.73	-0.07
bring (v)	24	7	1	20	-14.85***	16.26****
come (v)	61	99	41	58	1.89	0.13
connect (v)	10	10	9	4	-0.34	-3.59
copy (v)	1	18	5	6	14.37***	-0.02
develop (v)	49	26	14	26	-14.49***	1.12
disappear (v)	6	14	8	9	1.54	-0.10
enter (v)	3	21	10	2	10.88***	-8.24**
evolve (v)	12	13	10	8	-0.21	-1.11
face (v)	13	13	5	14	-0.45	2.33
feel (v)	53	89	39	72	2.22	2.99
find (v)	26	20	9	22	-3.15	2.64
follow (v)	5	13	6	12	1.89	0.73
get (v)	13	26	7	20	1.68	3.47
give (v)	28	23	17	23	-2.68	0.01
go (v)	5	12	5	10	1.43	0.61
grow (v)	2	12	4	8	5.54*	0.49
have (v)	280	407	208	241	2.09	-1.79
hold (v)	61	73	36	46	-0.23	-0.02
identify (v)	3	12	1	11	3.68	7.24**
imitate (v)	7	12	5	7	0.35	0.01
import (v)	4	30	12	9	16.27****	-1.65
introduce (v)	45	23	14	15	-14.18***	-0.30
keep (v)	23	36	20	26	0.49	0.00
look (v)	23	23	18	18	-0.79	-0.67
lose (v)	33	46	28	28	0.09	-1.05
make (v)	62	105	46	82	2.78	2.79
mix (v)	7	14	9	8	0.90	-0.65
rise (v)	11	4	5	6	-5.49*	-0.02
say (v)	6	11	1	15	0.47	11.16***
see (v)	16	17	3	18	-0.33	8.17**
show (v)	45	54	29	43	-0.16	0.25
spend (v)	4	30	12	8	16.27****	-2.28
spread (v)	32	41	19	17	0.00	-1.33
suit (v)	0	16	0	12	18.31****	13.57***
take (v)	24	29	13	22	-0.07	0.53
take in (ph v)	16	17	10	3	-0.33	-6.14*
take place (v)	5	10	6	9	0.64	0.06

Appendix L. Culture metaphor input and output data

1. The table below displays the sizes of input and output corpora.

Condition	Input corpus (tokens)	Output corpus (lexical units)
Control	245,363	43,830
Experimental	297,507	56,972

2. The table below displays the raw input frequencies of each culture metaphor theme that were presented in Figure 6.2 and Table 6.1.

Metaphor theme	Control			Experimental		
	Taught	Untaught	Total	Taught	Untaught	Total
CULTURES ARE LIVING THINGS	136	18	154	196	16	212
CULTURES ARE SPACES	101	4	105	223	21	244
CULTURES ARE SUBSTANCES	56	21	77	154	25	179
CULTURES ARE CONSTRUCTIONS	69	0	69	66	0	66
CULTURES HAVE MANY PARTS	82	7	89	180	10	190
CULTURES ARE POSSESSIONS	57	16	73	124	29	153
Total	**501**	**66**	**567**	**943**	**101**	**1044**

3. The following tables display the raw input and output frequencies of individual target words in each culture metaphor theme. The data in Tables 6.2 - 6.7 and 6.13 - 6.18 were based on these counts.

CULTURES ARE LIVING THINGS

Target word	Input frequency		Output frequency	
	Control	Experimental	Control	Experimental
Taught target forms				
alive (adj)	2	6	0	1
evolution (n)	0	6	2	3
evolve (v)	13	24	12	13
grow (v)	8	38	1	8
hybrid (adj)	52	75	13	14
hybrid (n)	61	47	13	13
Untaught target forms				
ancestor (n)	0	0	1	1
birth (n)	0	1	0	2
birthplace (n)	2	4	6	3
born (v)	0	0	23	23
bring up (ph. v)	0	0	0	1
grow up (ph. v)	0	0	0	6
growth (n)	11	5	0	1
habitat (n)	0	0	0	1
recover (v)	0	0	0	2
revitalize (v)	0	0	0	2
survival (n)	1	2	0	0
survive (v)	4	4	1	0

Target word	Input frequency		Output frequency	
	Control	Experimental	Control	Experimental
Taught target forms				
come into (ph. v)	24	65	5	35
enter (v)	1	10	2	19
export (n)	7	15	0	0
export (v)	0	3	1	1
import (n)	5	7	0	3
import (v)	18	28	4	33
inside (prep)	2	3	3	1
outside (adv)	3	9	0	0
outside (n)	11	12	0	1
outside (prep)	13	23	0	0
outsider (n)	2	10	0	0
surround (v)	2	12	1	0
take in (ph. v)	13	16	17	15
transcend (v)	0	10	0	0
Untaught target forms				
access (v)	0	0	1	0
borderline (n)	0	0	0	1
bring into (ph. v)	0	3	4	0
contain (v)	1	6	6	0
deport (v)*	0	0	0	1
go into (ph. v)	0	3	0	3
intake (v)*	1	0	1	0
leave (v)	0	1	0	0
outside (adj)	0	5	0	0
penetrate (v)	1	1	0	2
take into (ph. v)	1	1	0	0
within (prep)	1	0	0	0

Note: * = Marked forms by learners

Target word	Input frequency		Output frequency	
	Control	Experimental	Control	Experimental
Taught target forms				
absorb (v)	15	23	20	5
flow (n)	3	16	1	1
flow (v)	0	1	0	2
fusion (n)	3	14	0	2
mix (n)	0	5	0	0
mix (v)	5	21	9	21
mixture (n)	21	15	0	6
shape (n)	3	8	3	3
shape (v)	1	17	1	3
spread (v)	5	34	35	40
Untaught target forms				
combination (n)	2	2	0	0
combine (v)	5	3	7	3
derive (v)	0	0	2	0
flood (v)	0	0	0	1
form (n)	14	19	13	12
form (v)	0	0	1	1
permeate (v)	0	0	1	0
refine (v)	0	0	0	2
refined (adj)	0	1	0	0
spread out (ph. v)*	0	0	4	1

Note: * = Marked forms by learners

CULTURES ARE CONSTRUCTIONS

Target word	Input frequency		Output frequency	
	Control	Experimental	Control	Experimental
Taught target forms				
construct (v)	6	10	4	4
fit (v)	1	2	0	0
fit into (ph. v)	4	7	1	4
make up of (ph. v)	0	2	0	0
manufacture (v)	56	43	0	6
(un)damaged (adj)	2	2	0	1
Untaught target forms				
break (v)*	0	0	1	1
destroy (v)	0	0	1	0
destruction (n)	0	0	1	0
reconstruct (v)	0	0	1	0
restoration (n)	0	0	2	0

Note: * = Marked forms by learners

CULTURES HAVE MANY PARTS

Target word	Input frequency		Output frequency	
	Control	Experimental	Control	Experimental
Taught target forms				
aspect (n)	18	30	15	5
component (n)	1	10	2	4
divide (v)	2	4	0	1
element (n)	28	52	25	15
ingredient (n)	4	13	0	0
part (n)	29	71	13	8
Untaught target forms				
add (v)	5	7	15	26
piece (n)	2	3	0	2

Target word	Input frequency		Output frequency	
	Control	Experimental	Control	Experimental
Taught target forms				
borrow (v)	14	37	0	2
claim (v)	1	13	3	2
lose (v)	22	31	13	31
share (v)	9	22	0	4
shared (adj)	0	1	0	0
steal (v)	0	5	0	1
take (v)	11	15	11	16
Untaught target forms				
accept (v)	4	8	16	12
adopt (v)	1	1	7	11
adoption (n)	1	1	0	0
exchange (n)	0	3	1	0
exchange (v)	0	0	0	1
give (v)	0	1	1	2
hand down (ph. v)	0	0	2	0
have (v)	3	4	11	15
inherit (v)	0	0	1	2
keep (v)	1	4	12	10
pass (v)*	0	0	0	1
possess (v)	3	4	0	1
receive (v)*	0	0	3	6
retain (v)	3	3	0	1
save (v)	0	0	0	2

Note: * = Marked forms by learners

4. The following tables display the forms in which culture metaphors appeared in input. The data in Figure 6.5 and Table 6.10 were based on these counts.

Control condition

Metaphor theme	Form of input					Total
	Aural: video	Aural: teacher talk	Aural: explicit teaching	Written: text	Written: explicit teaching	
CULTURES ARE LIVING THINGS	14	76	44	10	10	154
CULTURES ARE SPACES	22	71	4	6	2	105
CULTURES ARE SUBSTANCES	32	34	6	4	1	77
CULTURES ARE CONSTRUCTIONS	7	44	8	10	0	69
CULTURES HAVE MANY PARTS	26	45	2	15	1	89
CULTURES ARE POSSESSIONS	18	42	3	10	0	73
Total	**119**	**312**	**67**	**55**	**14**	**567**

Experimental condition

Metaphor theme	Form of input					Total
	Aural: video	Aural: teacher talk	Aural: explicit teaching	Written: text	Written: explicit teaching	
CULTURES ARE LIVING THINGS	11	78	97	13	13	212
CULTURES ARE SPACES	24	107	78	23	12	244
CULTURES ARE SUBSTANCES	30	68	63	11	7	179
CULTURES ARE CONSTRUCTIONS	6	29	19	10	2	66
CULTURES HAVE MANY PARTS	27	87	47	19	10	190
CULTURES ARE POSSESSIONS	23	67	47	13	3	153
Total	**121**	**436**	**351**	**89**	**47**	**1044**

5. The table below displays the raw output frequencies of each culture metaphor theme that were presented in Figures 6.7 - 6.8 and Table 6.12.

Metaphor theme	Control			Experimental		
	Taught	Untaught	Total	Taught	Untaught	Total
CULTURES ARE LIVING THINGS	41	31	72	52	42	94
CULTURES ARE SPACES	33	12	45	108	7	115
CULTURES ARE SUBSTANCES	69	28	97	83	20	103
CULTURES ARE CONSTRUCTIONS	5	6	11	15	1	16
CULTURES HAVE MANY PARTS	55	15	70	33	28	61
CULTURES ARE POSSESSIONS	27	54	81	56	64	120
Total	**230**	**146**	**376**	**347**	**162**	**509**

6. The table below displays the normalised frequencies of target culture metaphors appearing in input and output in the two conditions. The normalised frequencies of the same metaphors as estimated from COCA are also displayed. These correlation values in Table 6.20 were based on these data.

Word	Normalised MRW input frequencies		Normalised MRW frequency	Normalised MRW output frequencies	
	Con.	Exp.	COCA	Con.	Exp.
alive (adj)	0.082	0.202	0.174	0.000	0.175
evolution (n)	0.000	0.202	0.121	0.103	0.524
evolve (v)	0.530	0.807	0.227	0.615	2.273
grow (v)	0.326	1.278	1.080	0.051	1.399
hybrid (adj)	2.119	2.522	0.059	0.666	2.448
hybrid (n)	2.486	1.580	0.049	0.666	2.273
come into (ph v)	0.978	2.185	0.200	0.256	6.119
enter (v)	0.041	0.336	0.631	0.103	3.322
export (v)	0.000	0.101	0.018	0.051	0.175
import (n)	0.204	0.235	0.006	0.000	0.524
import (v)	0.734	0.941	0.016	0.205	5.769
inside (prep)	0.082	0.101	0.237	0.154	0.175
outside (n)	0.448	0.403	0.009	0.000	0.175
surround (v)	0.082	0.403	0.148	0.051	0.000
take in (ph v)	0.530	0.538	0.116	0.871	2.622

(continued)

Word	Normalised MRW input frequencies		Normalised MRW frequency	Normalised MRW output frequencies	
	Con.	Exp.	COCA	Con.	Exp.
absorb (v)	0.611	0.773	0.142	1.025	0.874
flow (n)	0.122	0.538	0.225	0.051	0.175
flow (v)	0.000	0.034	0.144	0.000	0.350
fusion (n)	0.122	0.471	0.022	0.000	0.350
mix (v)	0.204	0.706	0.203	0.461	3.671
mixture (n)	0.856	0.504	0.058	0.000	1.049
shape (n)	0.122	0.269	0.190	0.154	0.524
shape (v)	0.041	0.572	0.237	0.051	0.524
spread (v)	0.204	1.143	0.306	1.794	6.993
construct (v)	0.245	0.336	0.197	0.205	0.699
fit into (ph v)	0.163	0.235	0.043	0.051	0.699
manufacture (v)	2.282	1.446	0.017	0.000	1.049
aspect (n)	0.734	1.009	0.699	0.769	0.874
component (n)	0.041	0.336	0.006	0.103	0.699
divide (v)	0.082	0.134	0.128	0.000	0.175
element (n)	1.141	1.748	0.688	1.282	2.622
part (n)	1.182	2.387	3.871	0.666	1.399
borrow (v)	0.571	1.244	0.052	0.000	0.350
claim (v)	0.041	0.437	1.045	0.154	0.350
lose (v)	0.897	1.042	2.959	0.666	5.420
share (v)	0.367	0.740	1.131	0.000	0.699
steal (v)	0.000	0.168	0.192	0.000	0.175
take (v)	0.448	0.504	9.429	0.564	2.797

Appendix M. Longitudinal analysis data

1. The following tables display the number of lexical units and target MRWs produced by learners in three 5-week periods.

Code	Lexical units produced			Taught target MRWs produced			Untaught target MRWs produced			Combined target MRWs produced		
	Weeks 1-5	Weeks 6-10	Weeks 11-15	Weeks 1-5	Weeks 6-10	Weeks 11-15	Weeks 1-5	Weeks 6-10	Weeks 11-15	Weeks 1-5	Weeks 6-10	Weeks 11-15
CON1	626	1084	983	6	6	9	1	3	8	7	9	17
CON2	502	693	659	3	4	2	3	2	6	6	6	8
CON3	236	348	358	0	4	0	0	0	3	0	4	3
CON4	651	759	672	1	6	2	2	4	1	3	10	3
CON5	522	849	1095	0	3	6	4	2	2	4	5	8
CON6	601	889	764	5	2	5	1	6	4	6	8	9
CON7	438	356	549	1	2	7	0	3	5	1	5	12
CON8	907	1092	1083	6	2	5	3	2	1	9	4	6
CON9	369	328	421	4	3	6	0	1	1	4	4	7
CON10	549	737	672	2	0	4	0	0	2	2	0	6
CON11	581	651	736	1	5	9	1	1	2	2	6	11
CON12	347	674	483	3	0	4	0	4	0	3	4	4
CON13	438	331	593	0	2	3	3	3	0	3	5	3
CON14	599	875	757	1	1	2	2	1	13	3	2	15
CON15	318	533	491	1	2	4	1	1	4	2	3	8
CON16	478	341	450	1	2	1	3	1	0	4	3	1
CON17	524	976	839	1	3	5	1	1	4	2	4	9
CON18	422	624	663	2	5	0	0	1	0	2	6	0
CON19	252	381	353	0	1	0	1	1	1	1	2	1
CON20	550	760	686	2	10	6	1	6	0	3	16	6
CON21	603	922	934	5	9	8	2	1	2	7	10	10
CON22	501	775	756	0	2	5	2	0	1	2	2	6
CON23	544	954	1133	4	5	9	4	5	2	8	10	11

Code	Lexical units produced			Taught target MRWs produced			Untaught target MRWs produced			Combined target MRWs produced		
	Weeks 1-5	Weeks 6-10	Weeks 11-15	Weeks 1-5	Weeks 6-10	Weeks 11-15	Weeks 1-5	Weeks 6-10	Weeks 11-15	Weeks 1-5	Weeks 6-10	Weeks 11-15
EXP1	738	1083	905	1	5	5	1	4	9	2	9	14
EXP2	742	903	879	0	9	10	1	3	3	1	12	13
EXP3	613	645	641	1	0	1	1	2	1	2	2	2
EXP4	747	1071	860	3	2	15	1	0	5	4	2	20
EXP5	927	1113	1039	1	2	6	2	2	5	3	4	11
EXP6	724	1111	811	0	7	12	0	6	2	0	13	14
EXP7	810	742	755	6	3	3	0	1	5	6	4	8
EXP8	879	923	853	6	6	11	3	0	4	9	6	15
EXP9	1005	882	910	1	5	15	1	0	1	2	5	16
EXP10	1082	1360	1231	0	9	12	1	3	6	1	12	18
EXP11	843	824	1070	2	3	14	2	0	6	4	3	20
EXP12	592	806	658	3	5	7	0	4	8	3	9	15
EXP13	385	642	565	2	8	8	1	0	7	3	8	15
EXP14	573	788	818	1	6	13	0	0	0	1	6	13
EXP15	682	659	714	0	4	1	1	1	0	1	5	1
EXP16	765	908	1013	0	3	3	1	1	3	1	4	6
EXP17	796	782	801	1	7	11	1	0	1	2	7	12
EXP18	517	927	827	0	12	6	0	4	6	0	16	12
EXP19	820	958	916	3	4	7	2	3	1	5	7	8
EXP20	763	822	843	2	2	8	7	4	9	9	6	17
EXP21	599	806	684	3	7	3	1	1	3	4	8	6
EXP22	949	917	850	3	7	16	2	0	5	5	7	21
EXP23	653	688	765	1	1	3	0	2	2	1	3	5

Appendix N. Files sampled from BAWE

The files samples from the BAWE corpus for the analyses in chapter seven.

ID	Title	Module	Year of program	Gender	Words	Lexical units
3001j	What continuities can one trace between nineteenth century anthropology and subsequent developments in the twentieth century?	Social Anthropology: Theories and Philosophy	2	f	1789	1656
3014a	Compare the 'functionalisms' of Malinowski and Radcliffe-Brown	Social Anthropology: Theories and Philosophy	2	f	2440	1852
3014b	Consider the view that Gender should be included as one of the core analytical categories in social anthropology.	Social Anthropology: Theories and Philosophy	2	f	2842	2073
3016c	The Dichotomies of Nature-Culture/Nature-Society: A review of the recent critical discussions	Social Anthropology: Theories and Philosophy	2	m	2806	2137
3027a	Anthropology- Library exercise	Introduction to Social Anthropology	1	f	1243	1091
3053a	How does Japanese Anime portray women's characters, as 'weak and vulnerable' or 'strong and heroic'?	Japan at play	3	f	1829	1674
3055a	Write an essay using the ethnography, 'Traveller Gypsies', to show how ideas about dirt and cleanliness, pollution and/or taboo may be related to systems of classification	Introduction to Social Anthropology	1	f	1525	1472
3088b	Religion' is sometimes a difficult concept to apply to beliefs and practices....	Introduction to Social Anthropology	1	f	1358	1327
3098a	Write an essay, using ethnography or other sources about a particular people, to show how ideas about dirt...	Introduction to Social Anthropology	1	m	1214	1106
3099a	Book Review essay - Geisha, by Liza Dalby	Introduction to Japanese Society and Culture	1	m	1093	1076
3126a	Book Review	Social Anthropology Theory 1	2	f	1055	826
3126b	Discuss, with examples, how the attitudes of different people to gifts and exchange may reveal broader information about their social organisation.	Introduction to Social Anthropology	1	f	2211	1523

ID	Title	Module	Year of program	Gender	Words	Lexical units
3126c	Anomie	Understanding Society 2: Transformations	1	f	684	523
3126d	Gender identities assessment	Gender and Society	2	f	804	737
3126e	A critical review of a sociological research study of your choice that employs a single research method either quantitative or qualitative	Researching the social world	2	f	1559	1304
3135a	Critical review of "Our land was a forest: An Ainu Memoir" - Kayano Shigeru	Minorities and Marginalities: Class and Conflict in Japan	2	f	989	801
3135b	How does the ethnography reflect/comment upon Japanese culture/society?	Introduction to Japanese Society and Culture	1	f	1016	758
3135c	Discuss some of the ways in which a society manages to maintain a degree of social and moral order amongst its members. Try to use informal examples as well as formal ones, and to provide ethnographic illustrations of your ideas.	Introduction to Social Anthropology	1	f	2232	2012

Appendix O. Error codes used in the study

The error coding system used in the study was based on the set of error tags developed for the National Institute of Information and Communications Technology Japanese Learner of English (NICT JLE) corpus.

Nouns

Error	Example (from NICT JLE Corpus)	Error category for this study
Noun inflection	There are four childrens playing the snowball.	Grammatical
Number of noun	One of the lady has a dog.	Grammatical
Noun case	So it's good [...] chance to visit around Shibuya's cafe.	Grammatical
Countability of noun	When I was a child, I was surrounded by the many natures...	Grammatical
Complement of noun	I don't have friends going to movie theatre.	Phraseological
Noun: lexis	I went to supermarket to buy some food. I choose some materials and I paid the fee.	Lexical

Verbs

Error	Example (from NICT JLE Corpus)	Error category for this study
Verb inflection	I growed some vegetable in the garden.	Grammatical
Subject-verb disagreement	The man who were standing in front of me hit me.	Grammatical
Verb form	I was very busy this morning because I have to came here.	Grammatical
Verb tense	...after I left her home, I tour around the Hiroshima city...	Grammatical
Verb aspect	I was having a dog almost fifteen years ago.	Grammatical
Verb voice	When I was a high school student, I was belonged to a kyudo club.	Grammatical
Usage of finite/infinite verb	We looking forward to visit that restaurant.	Grammatical
Verb negation	...I'm sorry I haven't time to go out for your party.	Grammatical
Question	What you use the computer for?	Grammatical
Complement of verb	...I put on the suit, I didn't like it. So I would like you to exchange another one.	Phraseological
Verb: lexis	...please compare, watching a movie in a theater and renting a video at home.	Lexical

Adjectives

Error	Example (from NICT JLE Corpus)	Error category for this study
Adjective inflection	My room is more dirty.	Grammatical
Usage of positive / comparative / superlative of adjective	I found that the kids in those schools are much much better, behave good than Japanese.	Grammatical
Complement of adjective	Was it sad, or were you happy for coming back?	Phraseological
Adjective: lexis	I guess fifteen thousand is very big price for me.	Lexical

Adverbs

Error	Example (from NICT JLE Corpus)	Error category for this study
Adverb inflection	I have been studying more hard recently.[a]	Grammatical
Usage of positive / comparative / superlative of adverb	...sometimes, I eat dinner in center of Tokyo. So I came back home later.	Grammatical
Adverb position	...we have a chorus competition recently. So we're practicing hard singing.	Phraseological
Adverb: lexis	...there are a lot of trees. But, my neighborhood, that is not so much trees.	Lexical

a = Invented example - NICT JLE corpus contains no examples of this tag

Others

Error	Example (from NICT JLE Corpus)	Error category for this study
Collocation: lexis[b]	Ekiden is a big game for runners.[c] (*game* should be *event*)	Lexical
Collocation: phrase[b]	Japanese anime is one of the biggest cultures.[c]	Phraseological
Misordering of words	And this young man didn't hurt anything, but he complained about his cellular phone broken...	Phraseological
Unknown type errors	...it's hard to drive in a car because the right side drive is dangerous and hesitate to drive. [...] And, for my ancient time, so, someday, I'll try.	Other
Japanese English	...drinking something from a pet bottle...	Other

b = Codes adopted for this study. The code in the NICT JLE tagset is 'collocation.'

c = Examples based on learner output in this study.

Appendix P. Frequency and part of speech data

1. The table below displays the sizes of corpora used for the analyses in sections 7.3 and 7.4.

Corpus	Size (lexical units)
Control output	43,830
Experimental output	56,972
BAWE sample	23,937

2. The table below displays the number of open-class MRWs in the control and experimental output corpora and the BAWE sample that fall into each of a range of frequency bands in COCA. The analyses in Figure 7.1 and Table 7.1 were based on these data.

Corpus	MRW tokens grouped by COCA frequency (per 10,000 words)						Total
	<1.0	1.0-1.99	2.0-2.99	3.0-3.99	4.0-4.99	5.0+	
BAWE	939	392	160	89	108	446	2134
Control	1114	302	225	135	186	1018	2980
Experimental	1436	391	293	158	157	1448	3883

3. The following tables display individual words that were found to significantly favour one corpus over another (log-likelihood: ± 3.84) with minimum Bayes Factor effect size of 2.00. Metaphors must also have been produced by at least 25% of participants.

Frequency: 5+ per 10,000 words in COCA

Metaphor (POS)	1. Control 2. BAWE		1. Experimental 2. BAWE		1. Control 2. Experimental	
	Log-likelihood	Effect size	Log-likelihood	Effect size	Log-likelihood	Effect size
have (v)	34.66****	23.54^{+++}	52.70****	41.46^{+++}	-2.22	-9.30
see (v)	-60.59****	49.47^{+++}	-76.72****	65.48^{+++}	0.33	-11.19
come (v)	10.10**	-1.03	18.88****	7.59^{++}	-1.89	-9.63
thing (n)	118.44****	107.31^{+++}	120.61****	109.32^{+++}	0.09	-11.43
give (v)	-7.28**	-3.84	-17.95****	8.67^{++}	2.68	-8.84
big (adj)	10.27**	-0.85	21.75****	10.45^{+++}	-3.66	-7.87
show (v)	-11.85***	0.72	-15.61****	4.35^{+}	0.16	-11.36

Note: Negative log-likelihood values indicate underuse in corpus 1 relative to corpus 2 (positive values indicate overuse)

** = Significant at the 0.01 level; *** = 0.001 level; **** = 0.0001 level

+ = Positive evidence against null hypothesis; ++ = Strong evidence; +++ = Very strong evidence

Frequency: 4.0 - 4.99 per 10,000 words in COCA

Metaphor (POS)	1. Control 2. BAWE		1. Experimental 2. BAWE		1. Control 2. Experimental	
	Log-likelihood	Effect size	Log-likelihood	Effect size	Log-likelihood	Effect size
system (n)	-57.06****	45.94^{+++}	-64.07****	52.80^{+++}	-0.02	-11.50
bring (v)	13.81***	2.69^{+}	0.90	-10.40	15.59****	4.07^{+}

Note: Negative log-likelihood values indicate underuse in corpus 1 relative to corpus 2 (positive values indicate overuse)

*** = Significant at the 0.001 level; **** = 0.0001 level

+ = Positive evidence against null hypothesis; +++ = Very strong evidence

Frequency: 3.0 - 3.99 per 10,000 words in COCA

Metaphor (POS)	1. Control 2. BAWE		1. Experimental 2. BAWE		1. Control 2. Experimental	
	Log-likelihood	Effect size	Log-likelihood	Effect size	Log-likelihood	Effect size
area (n)	-29.14****	18.01^{+++}	-34.10****	22.80^{+++}	0.00	-11.52
long (adj)	22.28****	11.15^{+++}	13.60***	2.31^{+}	2.68	-8.84
power (n)	-17.69****	6.57^{++}	-36.16****	24.88^{+++}	3.51	-8.01
lose (v)	16.84****	5.71^{+}	19.28****	7.98^{++}	-0.06	-11.46
lead (v)	-9.93**	-1.19	-15.70****	4.41^{+}	0.54	-10.98

Note: Negative log-likelihood values indicate underuse in corpus 1 relative to corpus 2 (positive values indicate overuse)

** = Significant at the 0.01 level; *** = 0.001 level; **** = 0.0001 level

+ = Positive evidence against null hypothesis; ++ = Strong evidence; +++ = Very strong evidence

Frequency: 2.0 - 2.99 per 10,000 words in COCA

Metaphor (POS)	1. Control 2. BAWE		1. Experimental 2. BAWE		1. Control 2. Experimental	
	Log-likelihood	Effect size	Log-likelihood	Effect size	Log-likelihood	Effect size
spend (v)	-0.72	-10.40	6.16*	-5.14	-16.27****	4.75^{+}
low (adj)	-5.13*	-5.99	-21.28****	9.99^{++}	6.84**	-4.68
sense (n)	-6.08*	-5.05	-18.35****	7.06^{++}	4.06*	-7.46
hard (adj)	-3.68	-7.44	5.70*	-5.60	-26.25****	14.72^{+++}
field (n)	-1.44	-9.68	-13.90***	2.61^{+}	8.19**	-3.33
develop (v)	29.99****	18.87^{+++}	8.70**	-2.60	14.49***	2.96^{+}

Note: Negative log-likelihood values indicate underuse in corpus 1 relative to corpus 2 (positive values indicate overuse)

* = Significant at the 0.05 level; ** = 0.01 level; *** = 0.001 level; **** = 0.0001 level

+ = Positive evidence against null hypothesis; ++ = Strong evidence; +++ = Very strong evidence

Frequency: 1.0 - 1.99 per 10,000 words in COCA

Metaphor (POS)	1. Control 2. BAWE		1. Experimental 2. BAWE		1. Control 2. Experimental	
	Log-likelihood	Effect size	Log-likelihood	Effect size	Log-likelihood	Effect size
view (n)	-18.82****	7.70^{++}	-27.44****	16.15^{+++}	0.65	-10.87
form (n)	-19.87****	8.74^{++}	-27.73****	16.45^{+++}	0.41	-11.11
clear (adj)	-1.85	-9.27	-17.05****	5.75^{+}	9.99**	-1.53
approach (n)	-48.44****	37.31^{+++}	-65.77****	54.47^{+++}	1.67	-9.86
order (n)	-33.30****	22.18^{+++}	-38.97****	27.67^{+++}	0.00	-11.52
present (v)	-31.22****	20.10^{+++}	-36.54****	25.24^{+++}	0.00	-11.52
apply (v)	-27.06****	15.93^{+++}	-25.15****	13.86^{+++}	-1.14	-10.38
structure (n)	-18.80****	7.67^{++}	-29.23****	17.93^{+++}	1.67	-9.86
maintain (v)	-43.71****	32.58^{+++}	-51.15****	39.85^{+++}	0.00	-11.52

Note: Negative log-likelihood values indicate underuse in corpus 1 relative to corpus 2 (positive values indicate overuse)

** = Significant at the 0.01 level; **** = 0.0001 level

+ = Positive evidence against null hypothesis; ++ = Strong evidence; +++ = Very strong evidence

Frequency: less than 1.0 per 10,000 words in COCA

Metaphor (POS)	1. Control 2. BAWE		1. Experimental 2. BAWE		1. Control 2. Experimental	
	Log-likelihood	Effect size	Log-likelihood	Effect size	Log-likelihood	Effect size
reflect (v)	-15.24****	4.11^{+}	-14.03***	2.74^{+}	-0.26	-11.26
in addition (adv)	42.83****	31.70^{+++}	8.15**	-3.15	30.45****	18.93
born (v)	20.04****	8.92^{++}	14.03***	2.73^{+}	1.73	-9.79
element (n)	-3.20	-7.92	-15.54****	4.26^{+}	5.83*	-5.69
introduce (v)	31.66****	20.54^{+++}	10.26**	-1.04	14.18***	2.66^{+}
context (n)	-24.98****	13.85^{+++}	-29.23****	17.93^{+++}	0.00	-11.52
demonstrate (v)	-6.08*	-5.05	-14.94***	3.65^{+}	2.24	-9.29
spread (v)	27.89****	16.76^{+++}	28.76****	17.46^{+++}	0.00	-11.52
point out (ph. v)	-20.81****	9.69^{++}	-24.36****	13.06^{+++}	0.00	-11.52
perception (n)	-16.65****	5.53^{+}	-19.49****	8.19^{++}	0.00	-11.52
link (v)	-10.47**	-0.66	-14.03***	2.74^{+}	0.14	-11.38
at first (adv)	16.56****	5.43^{+}	16.14****	4.83^{+}	0.05	-11.47
live (adj)	1.74	-9.38	21.05****	9.74^{++}	-22.60****	11.08^{+++}
on the other hand (adv)	36.74****	25.62^{+++}	16.43****	5.13^{+}	9.59**	-1.94
illustrate (v)	-18.80****	7.67^{++}	-29.23****	17.93^{+++}	1.67	-9.86
boundary (n)	-14.57***	3.45^{+}	-17.05****	5.75^{+}	0.00	-11.52
highlight (v)	-20.81****	9.69^{++}	-24.36****	13.06^{+++}	0.00	-11.52
absorb (v)	20.92****	9.79^{++}	4.21*	-7.09	16.80****	5.28^{+}
and so on (adv)	44.45****	33.32^{+++}	23.15****	11.85^{+++}	10.05**	-1.47
import (v)	3.49	-7.64	21.05****	9.74^{++}	-16.27****	4.75
idol (n)	3.49	-7.64	14.73***	3.42^{+}	-8.64**	-2.88
hybrid (adj)	13.94***	2.82^{+}	9.82**	-1.48	1.17	-10.35
gothic (adj)	1.74	-9.38	25.96****	14.66^{+++}	-29.78****	18.26^{+++}
lolita (n)	14.82**	3.69^{+}	32.27****	20.97^{+++}	-7.34**	-4.18

Note: Negative log-likelihood values indicate underuse in corpus 1 relative to corpus 2 (positive values indicate overuse)

* = Significant at the 0.05 level; ** = 0.01 level; *** = 0.001 level; **** = 0.0001 level

+ = Positive evidence against null hypothesis; ++ = Strong evidence; +++ = Very strong evidence

4. The following tables display the number of open-class MRW types produced by each learner in the control and experimental condition and in each text of the BAWE sample that fall into a range of frequency bands in COCA. The analyses in Figure 7.2 and Tables 7.2 and 7.3 were based on these data.

Code	MRW types grouped by COCA frequency (per 10,000 words)					
	<1.0	1.0-1.99	2.0-2.99	3.0-3.99	4.0-4.99	5.0+
CON1	66	21	16	3	7	30
CON2	26	3	5	4	6	21
CON3	13	1	1	2	4	7
CON4	30	8	3	2	4	25
CON5	63	19	11	5	5	25
CON6	39	13	12	3	7	20
CON7	31	12	6	2	4	14
CON8	50	19	10	8	6	30
CON9	22	3	5	3	2	12
CON10	31	8	6	2	3	16
CON11	31	6	7	3	5	19
CON12	36	5	7	5	8	10
CON13	31	10	7	0	3	13
CON14	31	17	7	6	7	24
CON15	11	5	4	4	1	12
CON16	39	4	2	5	1	15
CON17	36	21	9	4	3	33
CON18	16	5	8	2	5	15
CON19	12	1	4	2	5	10
CON20	37	11	7	5	7	18
CON21	42	8	9	7	7	25
CON22	36	7	7	3	2	16
CON23	55	17	7	7	7	15

Code	MRW types grouped by COCA frequency (per 10,000 words)					
	<1.0	1.0-1.99	2.0-2.99	3.0-3.99	4.0-4.99	5.0+
EXP1	49	22	11	5	8	27
EXP2	49	17	13	6	3	20
EXP3	27	12	7	3	6	16
EXP4	39	5	5	6	4	29
EXP5	62	21	8	3	6	29
EXP6	36	8	10	4	3	25
EXP7	41	6	5	3	1	20
EXP8	56	14	8	6	3	27
EXP9	31	12	7	6	9	22
EXP10	67	25	13	5	5	27
EXP11	49	19	11	6	5	25
EXP12	36	6	6	6	2	18
EXP13	33	9	4	5	1	16
EXP14	42	6	7	1	2	13
EXP15	32	8	5	3	4	20
EXP16	53	20	13	3	5	20
EXP17	33	9	8	7	7	20
EXP18	40	7	3	2	3	22
EXP19	51	11	12	2	4	30
EXP20	47	21	7	2	7	27
EXP21	24	11	6	5	3	19
EXP22	50	12	8	5	4	24
EXP23	29	15	12	9	5	15

Code	MRW types grouped by COCA frequency (per 10,000 words)					
	<1.0	1.0-1.99	2.0-2.99	3.0-3.99	4.0-4.99	5.0+
3001j	87	28	15	4	3	12
3014a	68	26	8	9	5	26
3014b	62	30	8	5	7	13
3016c	102	27	10	4	6	14
3027a	17	7	4	5	0	11
3053a	26	19	10	5	3	25
3055a	51	12	7	1	7	14
3088b	37	19	7	5	7	18
3098a	42	13	7	3	6	14
3099a	26	6	7	2	2	13
3126a	30	12	1	1	0	13
3126b	36	16	5	6	4	22
3126c	26	7	6	3	1	11
3126d	17	9	3	0	0	8
3126e	29	11	6	4	3	22
3135a	42	13	5	1	1	9
3135b	32	8	6	2	5	11
3135c	39	23	9	4	6	16

5. The table below displays the number of lexical units (MRW and all) from each open-class part of speech in the control and experimental output corpora and the BAWE sample. The analyses in Tables 7.4 and 7.5 and Figures 7.3 - 7.5 were based on these data.

Part of speech	Total (MRWs & non-MRWs)			MRWs		
	BAWE	Control	Experimental	BAWE	Control	Experimental
Noun	6000	12178	15144	776	801	1071
Verb	3251	6624	9134	979	1440	1964
Phrasal verb	89	92	129	43	58	79
Adjective	2033	3877	5507	259	304	483
Adverb	1202	2703	3214	83	353	270

Appendix Q. Collocation and colligation data

1. The table below displays the sizes of corpora used for the analyses in sections 7.3 and 7.4.

Corpus	Size (lexical units)
Control output	43,830
Experimental output	56,972
Control input	245,363
Experimental input	297,507

2. The following tables display the number of collocation and colligation features produced by learners and the number of those uses that contained an error. The analyses in section 7.5 (Tables 7.6 - 7.7, 7.15, 7.17 - 7.18 and 7. 20) were based on these data.

Code	Total lexical units	Literal adj/n collocations	% literal adj/n collocation	MRW adj/n collocations	% MRW adj/n collocation	MRW adj/n collocation errors	% MRW adj/n collocation errors
CON1	2693	2	0.07%	2	0.07%	0	0.00%
CON2	1854	0	0.00%	1	0.05%	0	0.00%
CON3	942	0	0.00%	1	0.11%	0	0.00%
CON4	2082	1	0.05%	4	0.19%	2	0.10%
CON5	2466	3	0.12%	7	0.28%	0	0.00%
CON6	2254	0	0.00%	4	0.18%	0	0.00%
CON7	1410	0	0.00%	1	0.07%	0	0.00%
CON8	3082	1	0.03%	14	0.45%	1	0.03%
CON9	1118	3	0.27%	0	0.00%	/	/
CON10	1958	2	0.10%	2	0.10%	0	0.00%
CON11	1968	0	0.00%	7	0.36%	0	0.00%
CON12	1504	4	0.27%	0	0.00%	/	/
CON13	1362	1	0.07%	1	0.07%	0	0.00%
CON14	2231	3	0.13%	14	0.63%	2	0.09%
CON15	1342	1	0.07%	0	0.00%	/	/
CON16	1269	0	0.00%	4	0.32%	0	0.00%
CON17	2482	4	0.16%	2	0.08%	1	0.04%
CON18	1709	2	0.12%	5	0.29%	0	0.00%
CON19	986	0	0.00%	1	0.10%	0	0.00%
CON20	1996	4	0.20%	5	0.25%	0	0.00%
CON21	2459	1	0.04%	5	0.20%	2	0.08%
CON22	2032	3	0.15%	1	0.05%	0	0.00%
CON23	2631	5	0.19%	1	0.04%	1	0.04%

Code	Total lexical units	Literal adj/n collocations	% literal adj/n collocation	MRW adj/n collocations	% MRW adj/n collocation	MRW adj/n collocation errors	% MRW adj/n collocation errors
EXP1	2726	0	0.00%	10	0.37%	3	0.11%
EXP2	2524	1	0.04%	4	0.16%	0	0.00%
EXP3	1899	0	0.00%	4	0.21%	3	0.16%
EXP4	2678	1	0.04%	6	0.22%	2	0.07%
EXP5	3079	2	0.06%	11	0.36%	4	0.13%
EXP6	2646	2	0.08%	5	0.19%	0	0.00%
EXP7	2307	0	0.00%	3	0.13%	1	0.04%
EXP8	2655	0	0.00%	6	0.23%	2	0.08%
EXP9	2797	2	0.07%	8	0.29%	1	0.04%
EXP10	3673	0	0.00%	12	0.33%	3	0.08%
EXP11	2737	3	0.11%	3	0.11%	0	0.00%
EXP12	2056	1	0.05%	3	0.15%	0	0.00%
EXP13	1592	0	0.00%	0	0.00%	/	/
EXP14	2179	2	0.09%	2	0.09%	0	0.00%
EXP15	2055	1	0.05%	6	0.29%	2	0.10%
EXP16	2686	0	0.00%	6	0.22%	2	0.07%
EXP17	2379	4	0.17%	3	0.13%	0	0.00%
EXP18	2271	0	0.00%	3	0.13%	1	0.04%
EXP19	2694	2	0.07%	3	0.11%	1	0.04%
EXP20	2428	0	0.00%	8	0.33%	3	0.12%
EXP21	2089	1	0.05%	4	0.19%	0	0.00%
EXP22	2716	3	0.11%	5	0.18%	1	0.04%
EXP23	2106	0	0.00%	8	0.38%	2	0.09%

Code	Total lexical units	MRW v/n collocations	% MRW v/n collocations	MRW v/n collocation errors	% MRW v/n collocation errors	MRW verb colligations	% MRW verb colligations	MRW verb colligation errors	% MRW verb colligation errors
CON1	2693	8	0.30%	2	0.07%	6	0.22%	2	0.07%
CON2	1854	7	0.38%	3	0.16%	4	0.22%	2	0.11%
CON3	942	3	0.32%	3	0.32%	4	0.42%	3	0.32%
CON4	2082	3	0.14%	0	0.00%	/	/	/	/
CON5	2466	8	0.32%	3	0.12%	4	0.16%	2	0.08%
CON6	2254	8	0.35%	5	0.22%	6	0.27%	4	0.18%
CON7	1410	/	/	/	/	3	0.21%	2	0.14%
CON8	3082	3	0.10%	1	0.03%	5	0.16%	2	0.06%
CON9	1118	2	0.18%	1	0.09%	14	1.25%	5	0.45%
CON10	1958	3	0.15%	3	0.15%	1	0.05%	0	0.00%
CON11	1968	7	0.36%	4	0.20%	6	0.30%	3	0.15%
CON12	1504	6	0.40%	3	0.20%	5	0.33%	2	0.13%
CON13	1362	5	0.37%	0	0.00%	3	0.22%	1	0.07%
CON14	2231	5	0.22%	1	0.04%	/	/	/	/
CON15	1342	5	0.37%	2	0.15%	7	0.52%	6	0.45%
CON16	1269	5	0.39%	2	0.16%	3	0.24%	2	0.16%
CON17	2482	6	0.24%	1	0.04%	1	0.04%	0	0.00%
CON18	1709	4	0.23%	2	0.12%	1	0.06%	1	0.06%
CON19	986	/	/	/	/	1	0.10%	0	0.00%
CON20	1996	7	0.35%	3	0.15%	9	0.45%	3	0.15%
CON21	2459	11	0.45%	5	0.20%	12	0.49%	4	0.16%
CON22	2032	9	0.44%	3	0.15%	3	0.15%	2	0.10%
CON23	2631	3	0.11%	1	0.04%	7	0.27%	3	0.11%

Code	Total lexical units	MRW v/n collocations	% MRW v/n collocations	MRW v/n collocation errors	% MRW v/n collocation errors	MRW verb colligations	% MRW verb colligations	MRW verb colligation errors	% MRW verb colligation errors
EXP1	2726	14	0.51%	5	0.18%	6	0.22%	1	0.04%
EXP2	2524	13	0.52%	4	0.16%	8	0.32%	4	0.16%
EXP3	1899	1	0.05%	0	0.00%	3	0.16%	1	0.05%
EXP4	2678	14	0.52%	3	0.11%	6	0.22%	2	0.07%
EXP5	3079	7	0.23%	1	0.03%	2	0.06%	0	0.00%
EXP6	2646	1	0.04%	0	0.00%	10	0.38%	8	0.30%
EXP7	2307	6	0.26%	0	0.00%	3	0.13%	1	0.04%
EXP8	2655	5	0.19%	1	0.04%	7	0.26%	0	0.00%
EXP9	2797	6	0.21%	0	0.00%	7	0.25%	1	0.04%
EXP10	3673	14	0.38%	7	0.19%	5	0.14%	1	0.03%
EXP11	2737	4	0.15%	1	0.04%	8	0.29%	0	0.00%
EXP12	2056	7	0.34%	4	0.19%	8	0.39%	7	0.34%
EXP13	1592	13	0.82%	6	0.38%	4	0.25%	3	0.19%
EXP14	2179	7	0.32%	4	0.18%	4	0.18%	2	0.09%
EXP15	2055	4	0.19%	2	0.10%	2	0.10%	1	0.05%
EXP16	2686	3	0.11%	2	0.07%	2	0.07%	1	0.04%
EXP17	2379	6	0.25%	2	0.08%	9	0.38%	3	0.13%
EXP18	2271	7	0.31%	2	0.09%	3	0.13%	2	0.09%
EXP19	2694	5	0.19%	2	0.07%	4	0.15%	0	0.00%
EXP20	2428	23	0.95%	14	0.58%	5	0.21%	5	0.21%
EXP21	2089	4	0.19%	3	0.14%	11	0.53%	9	0.43%
EXP22	2716	23	0.85%	9	0.33%	9	0.33%	1	0.04%
EXP23	2106	/	/	/	/	5	0.24%	1	0.05%

Code	Total lexical units	MRW noun colligations	% MRW noun colligations	MRW noun colligation errors	% MRW noun colligation errors
CON1	2693	7	0.26%	1	0.04%
CON2	1854	/	/	/	/
CON3	942	/	/	/	/
CON4	2082	3	0.14%	1	0.05%
CON5	2466	4	0.16%	0	0.00%
CON6	2254	3	0.13%	2	0.09%
CON7	1410	2	0.14%	0	0.00%
CON8	3082	4	0.13%	1	0.03%
CON9	1118	/	/	/	/
CON10	1958	2	0.10%	0	0.00%
CON11	1968	7	0.36%	4	0.20%
CON12	1504	4	0.27%	1	0.07%
CON13	1362	2	0.15%	1	0.07%
CON14	2231	1	0.04%	0	0.00%
CON15	1342	/	/	/	/
CON16	1269	/	/	/	/
CON17	2482	5	0.20%	0	0.00%
CON18	1709	3	0.18%	2	0.12%
CON19	986	1	0.10%	1	0.10%
CON20	1996	1	0.05%	1	0.05%
CON21	2459	2	0.08%	1	0.04%
CON22	2032	2	0.10%	0	0.00%
CON23	2631	9	0.34%	5	0.19%

Code	Total lexical units	MRW noun colligations	% MRW noun colligations	MRW noun colligation errors	% MRW noun colligation errors
EXP1	2726	/	/	/	/
EXP2	2524	/	/	/	/
EXP3	1899	/	/	/	/
EXP4	2678	1	0.04%	1	0.04%
EXP5	3079	2	0.06%	0	0.00%
EXP6	2646	1	0.04%	1	0.04%
EXP7	2307	/	/	/	/
EXP8	2655	6	0.23%	1	0.04%
EXP9	2797	4	0.14%	2	0.07%
EXP10	3673	4	0.11%	3	0.08%
EXP11	2737	2	0.07%	2	0.07%
EXP12	2056	/	/	/	/
EXP13	1592	2	0.13%	1	0.06%
EXP14	2179	2	0.09%	0	0.00%
EXP15	2055	2	0.10%	0	0.00%
EXP16	2686	/	/	/	/
EXP17	2379	1	0.04%	0	0.00%
EXP18	2271	6	0.26%	5	0.22%
EXP19	2694	1	0.04%	1	0.04%
EXP20	2428	2	0.08%	2	0.08%
EXP21	2089	/	/	/	/
EXP22	2716	3	0.11%	0	0.00%
EXP23	2106	/	/	/	/

3. The following tables display the number of uses of each metaphorical verb analysed in section 7.5.3 in the usage patterns produced by learners in both conditions. The analysis in Table 7.16 was based on these data.

MRW verb	Most frequent pattern	Uses in each pattern			Total uses	Most frequent pattern uses	Less frequent pattern uses
		Passive	Transitive	Intransitive			
evolve	Intransitive	3	3	19	25	19	6
take in	Transitive	4	23	0	27	23	4
absorb	Transitive	0	25	2	27	25	2
mix	Intransitive	2	14	7	23	7	16
spread	Intransitive	10	16	47	73	47	26
lose	Transitive	0	53	8	61	53	8
					Total	**174**	**62**

MRW verb	Errors in each pattern			Total errors	Most frequent pattern errors	Less frequent pattern errors
	Passive	Transitive	Intransitive			
evolve	3	2	8	13	8	5
take in	3	10	0	13	10	3
absorb	0	8	2	10	8	2
mix	1	9	4	14	4	10
spread	8	5	26	39	26	13
lose	0	7	7	14	7	7
				Total	**63**	**40**

4. The table below displays the number of uses of [part] nouns in three colligation patterns in the learner output corpora and COCA. The analysis in Table 7.19 was based on these data.

Corpus	Pre-modifying adj.	Head of prepositional phrase	With verb collocate	Total uses
COCA	89853	276818	125293	491964
Control	26	38	20	62
Experimental	14	22	12	39
Combined learners	40	60	32	101

Appendix R. Metaphor source feedback form data

1. The following tables display the number of feedback forms received from each learner and the number of times each source was selected. The analysis in Table 7.21 was based on these data.

Control condition

Number	Metaphor forms received	Sources of metaphor					
		L2 dictionary	L1 dictionary	Translated	Used words from class	Knew language already	Other
CON1	8	1	4	2	1	2	0
CON2	2	0	1	0	0	1	0
CON3	4	0	4	0	0	0	1
CON4	5	0	0	3	1	2	1
CON5	6	0	4	4	0	2	0
CON6	6	3	1	2	0	6	1
CON7	5	0	4	1	0	4	0
CON8	6	1	0	1	0	0	4
CON9	1	0	0	1	0	0	0
CON10	3	2	0	1	0	1	0
CON11	6	1	1	4	1	1	0
CON12	3	3	0	0	0	0	0
CON13	6	0	2	2	1	4	0
CON14	7	0	0	6	0	2	0
CON15	3	0	3	0	1	1	0
CON16	7	0	6	1	0	2	1
CON17	4	1	2	2	0	4	0
CON18	3	1	1	1	0	0	0
CON19	3	0	1	2	0	0	0
CON20	8	0	3	3	0	3	0
CON21	5	0	0	2	3	2	0
CON22	2	1	1	0	0	1	0
CON23	8	1	1	5	0	6	0

Note: Learners were able to select more than one source for a given metaphor. Therefore, the number of sources does not equal the number of forms received by each learner.

Experimental condition							
		Sources of metaphor					
Number	Metaphor forms received	L2 dictionary	L1 dictionary	Translated	Used words from class	Knew language already	Other
EXP1	5	1	2	1	1	1	1
EXP2	8	0	2	1	5	0	0
EXP3	4	1	1	3	1	2	0
EXP4	8	0	6	2	2	6	0
EXP5	9	1	4	2	1	2	0
EXP6	4	1	2	1	0	2	0
EXP7	6	1	0	2	2	1	0
EXP8	7	0	1	0	5	2	0
EXP9	7	0	1	4	0	1	1
EXP10	7	0	3	3	2	1	2
EXP11	7	0	0	0	5	3	0
EXP12	5	0	1	1	3	2	1
EXP13	5	0	0	3	3	3	0
EXP14	5	2	4	0	0	3	0
EXP15	3	0	1	1	0	2	0
EXP16	9	0	6	3	0	3	3
EXP17	6	1	0	4	3	2	0
EXP18	7	1	1	6	0	0	0
EXP19	6	0	0	2	4	1	1
EXP20	7	1	2	1	3	2	0
EXP21	4	1	2	1	0	1	0
EXP22	8	0	1	3	2	4	1
EXP23	5	0	3	0	3	0	0

Note: Learners were able to select more than one source for a given metaphor. Therefore, the number of sources does not equal the number of forms received by each learner.

Appendix S. L1 influence data

1. The table below display verbs that are used metaphorically in Japanese in collocations with *bunka* (culture) and their collocation measures[47]. These values were used to identify L1 translation equivalents of metaphorical verbs in English.

English	Japanese	MI score	t-score
absorb	吸収 (*kyuushuu*)	5.14	28.61
accept; receive	受け入れる (*ukeireru*)	5.64	40.73
adapt	溶け込む (*tokekomu*)	6.32	18.34
adopt	取り上げる (*toriageru*)	4.13	20.68
advance	進歩 (*advance*)	6.46	30.17
bring up; grow up	育む (*hagukumu*)	8.82	50.76
combine	結び付ける (*musubitsukeru*)	5.04	12.57
come into	入り込む (*hairikomu*)	4.22	13.18
connect	繋がり (*tsunagari*)	4.90	21.67
copy	模倣 (*mohou*)	6.64	15.46
develop	発展 (*hatten*)	7.43	80.25
disappear	消え去る (*kiesaru*)	4.45	6.40
enter	入る (*hairu*)	2.11	37.12
erase	抹殺 (*massatsu*)	6.20	12.79
evolve	進化 (*shinka*)	4.79	25.39
export	輸出 (*yushutsu*)	5.22	22.86
feel	感ずる (*kanzuru*)	3.59	56.02
fit	合う (*au*)	3.01	24.93
flow	流れ込む (*nagarekomu*)	5.24	10.67
form	形成 (*keisei*)	6.64	51.03
grow	育つ (*sodatsu*)	5.97	40.88
have	持つ (*have*)	4.22	100.65
import	輸入 (*yunyuu*)	4.51	24.56
introduce	渡来 (*torai*)	8.08	19.62
lose	失う (*ushinau*)	4.44	31.91
make	作る (*tsukuru*)	3.22	60.12
mix	混在 (*konzai*)	6.79	17.70
reflect	反映 (*hanei*)	5.52	26.32

(continued)

[47] Collocation measures were calculated based on the number of times the collocate verb appeared within four spaces of bunka.

English	Japanese	MI score	t-score
shape	形作る (*katachidzukuru*)	6.51	13.67
share	共有 (*kyouyuu*)	5.37	32.09
spread	広める (*hiromeru*)	7.33	29.45
take in	取り入れる (*toriireru*)	6.57	43.63

2. The table below displays the normalised frequencies for metaphorical verbs in the study corpora. The analysis in Table 7.25 was based on these data.

English	Japanese	Normalised frequencies (per 10,000 words)					
		Control output	Experimental output	jptenten11 corpus	COCA	Control input	Experimental input
absorb	吸収 (*kyuushuu*)	5.48	1.05	0.25	0.23	0.69	1.34
accept; receive	受け入れる (*ukeireru*)	4.11	4.21	0.35	1.17	0.37	0.71
adapt	溶け込む (*tokekomu*)	0.23	0.00	0.04	0.29	1.79	1.51
adopt	取り上げる (*toriageru*)	1.83	1.93	0.28	0.55	1.51	1.51
advance	進歩 (*advance*)	0.23	0.53	0.11	0.30	0.04	0.44
bring up; grow up	育む (*hagukumu*)	0.68	1.23	0.06	0.72	0.77	0.74
combine	結び付ける (*musubitsukeru*)	1.60	0.53	0.05	0.66	0.77	0.27
come into	入り込む (*hairikomu*)	0.68	5.09	0.11	0.65	1.06	1.41
connect	繋がり (*tsunagari*)	2.28	1.76	0.17	0.60	2.20	1.88
copy	模倣 (*mohou*)	0.23	3.16	0.02	0.12	1.59	1.24
develop	発展 (*hatten*)	11.18	4.56	0.38	2.06	1.47	2.12
disappear	消え去る (*kiesaru*)	1.37	2.46	0.02	0.57	0.49	0.81
enter	入る (*hairu*)	0.68	3.69	5.49	1.31	0.69	1.41
erase	抹殺 (*massatsu*)	0.00	0.18	0.02	0.09	0.00	0.07
evolve	進化 (*shinka*)	2.74	2.28	0.26	0.29	0.53	0.81
export	輸出 (*yushutsu*)	0.23	0.18	0.15	0.10	0.04	0.20
feel	感ずる (*kanzuru*)	12.09	15.62	3.14	6.86	10.31	9.31
fit	合う (*au*)	0.91	1.05	1.02	0.69	0.69	0.61
flow	流れ込む (*nagarekomu*)	0.00	0.35	0.03	0.28	0.24	0.27
form	形成 (*keisei*)	0.46	0.18	0.27	1.07	0.65	0.50
grow	育つ (*sodatsu*)	0.46	2.11	0.28	2.77	1.63	2.99
have	持つ (*have*)	63.88	71.44	6.18	121.30	132.46	141.61
import	輸入 (*yunyuu*)	0.91	5.27	0.29	0.10	0.41	0.40

(continued)

English	Japanese	Normalised frequencies (per 10,000 words)					
		Control output	Experimental output	jptenten11 corpus	COCA	Control input	Experimental input
introduce	渡来 (*torai*)	10.27	4.04	0.01	0.77	0.77	1.21
lose	失う (*ushinau*)	7.53	8.07	0.52	3.29	2.73	3.03
make	作る (*tsukuru*)	14.15	18.43	4.93	20.98	26.08	28.64
mix	混在 (*konzai*)	1.60	2.46	0.03	0.51	0.86	1.14
reflect	反映 (*hanei*)	0.46	0.70	0.16	0.99	0.69	1.28
shape	形作る (*katachidzukuru*)	0.68	0.53	0.02	0.34	0.12	0.57
share	共有 (*kyouyuu*)	0.00	0.88	0.26	1.41	6.11	4.81
spread	広める (*hiromeru*)	7.30	7.20	0.06	0.68	0.33	1.24
take in	取り入れる (*toriireru*)	3.65	2.98	0.21	0.08	0.61	0.71

Note: Words with a frequency of zero were not included in the correlations calculated with this data.

3. The table below displays the normalised frequencies for metaphorical verb + [culture] collocations in the study corpora. The analysis in Table 7.26 was based on these data.

English	Japanese	Normalised frequencies (per 10,000 words)					
		Control output	Experimental output	jptenten11 corpus	COCA	Control input	Experimental input
absorb	吸収 (*kyuushuu*)	2.05	0.70	0.0008	0.0009	0.16	0.13
accept; receive	受け入れる (*ukeireru*)	0.68	1.05	0.0017	0.0009	0.00	0.00
adapt	溶け込む (*tokekomu*)	1.14	0.00	0.0003	0.0006	0.04	0.03
adopt	取り上げる (*toriageru*)	0.46	0.88	0.0005	0.0013	0.04	0.03
advance	進歩 (*advance*)	0.00	0.18	0.0009	0.0003	0.00	0.00
bring up; grow up	育む (*hagukumu*)	0.23	0.88	0.0025	0.0002	0.00	0.00
combine	結び付ける (*musubitsukeru*)	0.46	0.35	0.0002	0.0007	0.00	0.00
come into	入り込む (*hairikomu*)	0.23	1.93	0.0002	0.0000	0.04	0.30
connect	繋がり (*tsunagari*)	0.23	0.18	0.0005	0.0013	0.04	0.00
copy	模倣 (*mohou*)	0.00	0.53	0.0002	0.0000	0.00	0.00
develop	発展 (*hatten*)	1.83	0.35	0.0063	0.0017	0.00	0.00
disappear	消え去る (*kiesaru*)	0.00	0.35	0.0000	0.0008	0.00	0.03
enter	入る (*hairu*)	0.23	1.23	0.0023	0.0010	0.00	0.24
erase	抹殺 (*massatsu*)	0.00	0.18	0.0002	0.0003	0.00	0.00
evolve	進化 (*shinka*)	0.00	0.18	0.0007	0.0013	0.00	0.07
export	輸出 (*yushutsu*)	0.23	0.00	0.0005	0.0003	0.00	0.00
feel	感ずる (*kanzuru*)	0.00	0.35	0.0036	0.0003	0.00	0.00
fit	合う (*au*)	0.00	0.35	0.0008	0.0013	0.08	0.17
flow	流れ込む (*nagarekomu*)	0.00	0.18	0.0001	0.0001	0.00	0.00
form	形成 (*keisei*)	0.23	0.00	0.0026	0.0013	0.00	0.00
grow	育つ (*sodatsu*)	0.00	0.18	0.0017	0.0010	0.00	0.00
have	持つ (*have*)	2.97	4.04	0.0110	0.0137	0.20	0.17

(continued)

English	Japanese	Normalised frequencies (per 10,000 words)					
		Control output	Experimental output	jptenten11 corpus	COCA	Control input	Experimental input
import	輸入 (*yunyuu*)	0.68	2.11	0.0006	0.0002	0.00	0.07
introduce	渡来 (*torai*)	1.14	0.88	0.0004	0.0013	0.00	0.00
make	作る (*tsukuru*)	1.37	0.70	0.0044	0.0025	0.04	0.03
lose	失う (*ushinau*)	1.83	2.98	0.0011	0.0010	0.29	0.57
mix	混在 (*konzai*)	0.00	1.23	0.0003	0.0010	0.00	0.13
reflect	反映 (*hanei*)	0.23	0.18	0.0007	0.0016	0.08	0.07
shape	形作る (*katachidzukuru*)	0.23	0.18	0.0002	0.0040	0.00	0.27
share	共有 (*kyouyuu*)	0.00	0.35	0.0010	0.0035	0.12	0.20
spread	広める (*hiromeru*)	1.14	0.18	0.0009	0.0009	0.00	0.10
take in	取り入れる (*toriireru*)	0.91	1.05	0.0019	0.0001	0.04	0.07

Note: Collocations with a frequency of zero were not included in the correlations calculated with this data.