# An Investigation into the Role of Analogy in Instructed Second Language Acquisition




Andrew Harris

Linacre College

University of Oxford

A thesis submitted for the degree of

*Doctor of Philosophy*

Hilary Term 2013

# Acknowledgements

From proposal to submission of this thesis the most influential person by far has been my supervisor, Catherine Walter. With humour and forensic rigour, she has sympathetically guided half-baked ideas, meticulously analysed a seemingly endless number of drafts and courteously suggested on countless occasions revisions to everything. This work is so much better because of your help. Also at Oxford, thanks go to Ernesto Macaro and Vicky Murphy on the transfer and confirmation panels, whose brief appearance in the history of this project provided me with many nudges in the right direction.

I would also like to thank students and staff in the ELC department at London Metropolitan University, without whose assistance (staff) and patience (students) in the embryonic stages of this project, life would have been a lot harder. I'd also like to extend my thanks to the staff and students in the LC department at the University of East London. Accommodating me was a great help.

Thanks also go to Pat Healy for taking the time to understand NVivo and my obscure coding systems, and Cécile for the unwavering encouragement and support through thin and thick.

Andy Harris, February 2013

# Statement of Originality

The work contained in this thesis is entirely my own work, and has not been submitted, in part or in whole, in connection with any other course of study or degree.

Word Count: 73,000
Page Count: 285

Signed:

$20^{th}$ February 2013

# Abstract

Usage-based approaches to Second Language Acquisition (SLA) hold that grammatical development can at least in part be explained by a trajectory from initially entrenched formulaic chunks of linguistic material through partially abstract low-scope patterns to abstract constructions. Although earlier empirical research found little to support such a trajectory, more recent research suggests that when the methodological focus is on the breakdown of tokens not on the acquisition of abstract morphology formulaic chunks do seem to develop into at least low-scope patterns. What is not clear from these later studies, however, is why users select tokens from the environment, how these selections are repurposed to meet communicative needs and what grammatical development may be the result of such repurposing.

Drawing on insights into usage-based approaches to SLA in general, and specifically predictions that the analogical processing of similar tokens can explain in part a trajectory from tokens in the linguistic environment to abstract constructions, this study investigates whether intervention in the form of task demands and written task productions can drive the selection and repurposing of task-relevant input, and whether such repurposing may influence the development of past-counterfactual constructions in instructed SLA.

The study uses a classroom-based pre-test/post-test quasi-experimental design with pre-sessional university students in intact classes (n=92). Out of the three groups in the study, one group were exposed to instances of past counterfactuals which were identical to the forms needed for task completion in terms of function and lexical items (Literal Group). A second group were exposed to instances which called for the same function but different lexical items for successful task completion (Analogous Group), while a third group were not exposed to input.

Results show that the analogous processing of task-relevant tokens led to positive and significant gains on most acceptability judgment and production test measures. Regression models further show that the selection of tokens explained very little variance, while analogical processing, operationalised as the repurposing of tokens to fit task demands, explained a significant amount of the variance in the same models. These findings highlight the importance of analogical processing in SLA and the significant implications such processing has for cognitive accounts of SLA and for second language teaching.

# Contents

**List of Figures** **vii**

**List of Tables** **xi**

**1 Introduction** **1**

**2 Literature Review** **5**

2.1 Usage and a trajectory from formula to construction . . . . . 6

2.1.1 A property theory as a continuum of abstractness . . . . 9

2.1.1.1 Formulas and low-scope patterns . . . . . 9

2.1.1.2 Construction grammars . . . . . 16

2.1.1.3 Summary . . . . . 20

2.1.2 Evidence for a trajectory . . . . . 22

2.1.2.1 Evidence for a trajectory in naturalistic SLA . 22

2.1.2.1.1 Formulas and child SLA . . . . . 22

2.1.2.1.2 Summary . . . . . 30

2.1.2.1.3 Formulas and adult SLA . . . . . 31

2.1.2.1.4 Discussion . . . . . 36

2.1.2.2 Evidence for a trajectory in instructed SLA . . 42

2.1.2.2.1 R.Ellis (1984) . . . . . 42

2.1.2.2.2 Myles et al. (1998, 1999) . . . . . 44

2.1.2.2.3 Eskildsen (2012) . . . . . 47

2.1.3 Summary . . . . . 50

2.2 Processes in a usage-based trajectory . . . . . 51

2.2.1 Motivation for the breakdown of formulas . . . . . 52
2.2.1.1 Motivation from the linguistic environment . . 52
2.2.1.2 Motivation from communicative need . . . . . 55
2.2.2 Processes acting on tokens in the input: categorisation, analogy and similarity . . . . . 58
2.2.2.1 Categorisation of tokens . . . . . 58
2.2.2.2 The role of analogy and similarity in the categorisation of tokens . . . . . 61
2.2.2.3 Structure mapping and schema abstraction . . 63
2.2.2.4 A developmental perspective on the role of similarity and analogy . . . . . 65
2.2.2.5 Analogy and similarity in language learning . . 66
2.2.2.6 A career of similarity in language learning . . . 69
2.2.3 Summary . . . . . 72
2.3 The current research . . . . . 74
2.3.1 Summary and research space . . . . . 74
2.3.2 Research questions and hypotheses . . . . . 78

**3 Methodology** **81**
3.1 Research design . . . . . 81
3.1.1 Participants . . . . . 83
3.2 Materials . . . . . 85
3.2.1 Initial piloting of intervention materials and testing instuments . . . . . 85
3.2.2 Intervention materials . . . . . 86
3.2.2.1 The linguistic target: past counterfactual constructions . . . . . 86
3.2.2.1.1 Past counterfactuals: abstract and concrete . . . . . 86

3.2.2.1.2 A hypothetical trajectory for past counterfactuals . . . 87
3.2.2.1.3 Meaning of past counterfactuals . . . . 88
3.2.2.1.4 Cross-linguistic variation . . . . . 88
3.2.2.2 Task materials . . . . . 95
3.2.2.2.1 Establishing the task demands . . . . 96
3.2.2.2.2 Analogous and literal input . . . . . 97
3.2.2.2.3 Task production . . . . . 99
3.2.3 Testing instruments . . . . . 99
3.2.3.1 Operationalisation of acquisition . . . . . 99
3.2.3.2 Test materials . . . . . 100
3.2.3.2.1 Acceptability judgment test . . . . . 100
3.2.3.2.2 The production test . . . . . 102
3.2.4 Procedure . . . . . 103
3.3 Data scoring . . . . . 106
3.3.1 Scoring the tests . . . . . 107
3.3.1.1 The acceptability judgment test . . . . . 107
3.3.1.2 The production test . . . . . 107
3.3.2 Scoring selections from the input . . . . . 112
3.3.3 Scoring repurposing operations . . . . . 113
3.3.4 Stimulated recall interviews . . . . . 116
3.3.5 Coding the stimulated recall interviews . . . . . 118
3.3.5.1 Task demands and the selection of tokens from the input . . . . . 118
3.3.5.2 Explicit reference to patterns in the input and learner production . . . . . 119
3.3.5.3 Knowledge of the meaning of counterfactual constructions . . . . . 120
3.3.6 Inter-rater reliability . . . . . 120

**4 Findings** **123**

4.1 Quantitative Analyses . . . . . 123

4.1.1 Selection of tokens from the input . . . . . 123

4.1.2 Exact and repurposed use of selected tokens . . . . . 126

4.1.3 Treatment effects on acceptability judgment and production tests . . . . . 129

4.1.3.1 Acceptability judgment test . . . . . 130

4.1.3.2 Production test . . . . . 133

4.1.4 Effects of exact use and repurposed use on tests . . . . . 137

4.1.4.1 Predictors of acceptability judgment test scores 138

4.1.4.2 Predictors of production test scores . . . . . 140

4.1.4.3 Overall proficiency as a predictor of test scores 141

4.1.5 The effects of cross-linguistic variation across L1s . . . . 143

4.1.5.1 Overall proficiency as a confounding variable . 146

4.1.6 Summary of statistical analyses . . . . . 147

4.2 Qualitative studies . . . . . 150

4.2.1 Case Studies . . . . . 150

4.2.1.1 Case study one . . . . . 151

4.2.1.2 Case study two . . . . . 157

4.2.1.3 Case study three . . . . . 162

4.2.1.4 Case study four . . . . . 167

4.2.1.5 Case study five . . . . . 173

4.2.1.6 Discussion . . . . . 177

4.2.2 Stimulated recall interviews . . . . . 178

4.2.2.1 Selection of the input . . . . . 180

4.2.2.2 Processing of selected instances . . . . . 183

4.2.2.3 Knowledge of past counterfactuals . . . . . 187

4.2.2.4 Discussion . . . . . 191

**5 Discussion** **193**

5.1 Did the intervention promote analogical processing? . . . . . 194

5.2 The role of task demands in the breakdown of tokens . . . . . 202

5.3 Does analogical processing lead to acquisition? . . . . . 205

5.4 Did analogical processing lead to abstract constructional knowledge? . . . . . 209

5.5 Implications for cognitive accounts of SLA . . . . . 215

5.6 Pedagogical implications . . . . . 223

5.7 Limitations and future directions . . . . . 226

**6 Conclusions** **231**

**References** **237**

**Appendices**

**A Consent form and biographical questionnaire for the study** **259**

**B Ethics approval confirmation for the study (CUREC)** **261**

**C Pre-study information for participants** **263**

**D Production test** **265**

**E Acceptability judgment test** **267**

**F Materials for the Analogous Group (Group One, Task One)** **271**

**G Materials for the Literal Group (Group Two, Task One)** **275**

**H Materials for the Control Group (Group Three, Task One)** **279**

**I Data for the partial and full selection of result and condition clauses from the input** **283**

# List of Figures


2.1 Formulas and the creation of the lexicon in FLA . . . . . 40
2.2 Formulas and the creation of the lexicon in adult SLA . . . . . 41
2.3 Distribution of subject don't/no verb patterns over time . . . . 48
2.4 Native and non-native use of lexical verbs in Verb-Locative (VL), Verb-Object Locative (VOL) and Verb-Object-Object (VOO) constructions . . . . . 54
2.5 Categorising usage events . . . . . 58
2.6 Categorisation of and Schematisation from two tokens . . . . . 60
2.7 Extension and schematisation of usage events . . . . . 60
2.8 Mapping objects and relations from a base domain to a target domain . . . . . 62
2.9 Structural alignment and similarity between scenes . . . . . 63
2.10 Structural alignment in the analogical processing of tokens . . . 67
2.11 Schematisation based on attribute similarity . . . . . 68
2.12 High and low variance presentation of category members . . . . 70
2.13 Mapping of form and function in the analogical processing of past counterfactual tokens with shared attributes . . . . . 76
2.14 Predicted schematisation of low-scope patterns based on the mapping of form and function in the analogical processing of past counterfactual tokens with shared attributes . . . . . 76
2.15 Literal processing of selected input tokens (base) and output tokens (target) . . . . . 78

3.1 Summary of Research Design . . . . . 82
3.2 Summary of research design with task types . . . . . 96
3.3 Task rubric for Analogous, Literal and Control groups (Task One) 97
3.4 Task rubric with a sample of analogous pre-task input . . . . . 98
3.5 Task rubric for Group Two with a sample of literal input. . . . . 98
3.6 A six-point cline of acceptability in the acceptability judgment test . . . . . 102
3.7 A cline between naturally-occurring and experimentally-elicited samples . . . . . 103
3.8 Production test rubric . . . . . 104
3.9 Task procedure . . . . . 105
3.10 Determining target-like utterances . . . . . 110
3.11 Scoring the selections of past counterfactual condition and result clauses . . . . . 113

4.1 Full and partial selection of result and condition clauses from the input . . . . . 125
4.2 Mean number of repurposed result clauses in task production . . 127
4.3 Mean number of repurposed condition clauses in task production 128
4.4 Exact use of result and condition clauses in task production . . 129
4.5 Pre and post-test scores on the acceptability judgment test for the Analogous, Literal and Control groups . . . . . 130
4.6 Profile plots for Analogous, Literal and Control groups . . . . . 131
4.7 Q-Q plot of acceptability judgment test scores, showing outliers three or more standard deviations from the mean. . . . . 132
4.8 Use of result clauses in production test . . . . . 134
4.9 Use of condition clauses in production test . . . . . 135
4.10 Mean pre-test/post-test scores on the acceptability judgment test for the three L1 groups . . . . . 144

4.11 Mean pre-test/post-test scores on production test for the three L1 groups (result clauses) . . . . . 145
4.12 Mean pre-test/post-test scores on production test for the three L1 groups (condition clauses) . . . . . 146

5.1 Mapping of a base on to a target . . . . . 196
5.2 Mapping base selections on to targets (exact) . . . . . 201
5.3 Mapping base selections on to targets (repurposed) . . . . . 202
5.4 Mapping of a base on to a target . . . . . 204
5.5 Selections from the input, repurposed output and hypothetical low-scope past counterfactual patterns . . . . . 212
5.6 From small to large differences between base and target . . . . . 227

# List of Tables


2.1 Example developmental trajectories for tokens in the input . . . 10
2.2 Examples of lexicalised sentence stems . . . . . 13
2.3 Summary of formulas, low-scope patterns and alternative mechanisms . . . . . 15
2.4 Examples of the caused-motion construction with meaning assigned to the entire construction . . . . . 18
2.5 A continuum of constructions from concrete to abstract . . . . . 19
2.6 Degree of linguistic productivity/concreteness and its effects on a trajectory from fomula to low-scope pattern to construction . 21
2.7 Paul's production question forms . . . . . 27
2.8 Percentage of formulaic speech for five subjects over a nine-month period . . . . . 28
2.9 Nora's breakdown of formulaic speech over a nine-month period 29
2.10 Wes' accuracy order for nine grammatical morphemes in obligatory contexts . . . . . 33
2.11 Examples of formulaic phrases produced by Wes . . . . . 34
2.12 Standard and non-standard forms used by 'K' and advanced ESL learners . . . . . 36
2.13 Frequencies of the *comment t'appelles-tu*? chunk with other uses of *s'appeler*, by round . . . . . 46
2.14 Type-token rations for Carlos' SUBJ AUX-dopresent/past tense NEG VERB construction . . . . . 48

3.1 Dependent and independent variables . . . . . 83
3.2 Biographical data of participants in the study . . . . . 84
3.3 Summary of cross-linguistic differences in past-couterfactuals with issues for SLA . . . . . 95
3.4 Number and percentage of seeded past counterfactuals in the input. . . . . 99
3.5 Scoring of acceptability judgment test . . . . . 107
3.6 Coding of production test for condition and result clauses (targetlike and interlanguage measures). . . . . 112
3.7 Coding for the use of selections from input in task performance 115
3.8 A priori questions for stimulated recall interviews . . . . . 117
3.9 Inter-rater reliability of coding schemes . . . . . 121

4.1 Research questions and hypothesis related to the selection of partial and full condition and result clauses from the input . . . 124
4.2 Descriptive statistics for partial and full selection of result clauses.124
4.3 Descriptive statistics for partial and full selection of condition clauses. . . . . 124
4.4 Research questions and hypothesis related to the exact and repurposed use of result and condition clauses in task production . 126
4.5 Descriptive statistics for the exact and repurposed use of result clauses. . . . . 128
4.6 Descriptive statistics for the exact and repurposed use of condition clauses. . . . . 128
4.7 Research questions and hypothesis related to the development of past-counterfactuals . . . . . 129
4.8 Descriptive statistics for pre/post-test scores on the acceptability judgment test. . . . . 130
4.9 Between-group differences in gain scores on the acceptability judgment test . . . . . 133

4.10 Descriptive statistics for the use of result clauses in pre- and post-production test . . . . . 134
4.11 Descriptive statistics for the use of condition clauses in the production test . . . . . 135
4.12 Within group differences for the production test . . . . . 136
4.13 Between-group differences in gain scores on the production test 137
4.14 Hierarchical regression analysis of the selection and use of result clauses as predictors of gain scores on acceptability judgment test 139
4.15 Hierarchical regression analysis of the selection and use of condition clauses as predictors of gain scores on acceptability judgment test . . . . . 140
4.16 Hierarchical regression analysis of the selection and use of condition clauses as predictors of gain scores on the production test 141
4.17 Hierarchical regression analysis of the selection and use of result clauses as predictors of gain scores on the production test . . . . 141
4.18 Regression analysis of IELTS score as a predictor of gain scores on the acceptability judgment test . . . . . 142
4.19 Regression analysis of IELTS score as a predictor of gain scores on the production of result clauses in the production test . . . . 142
4.20 Regression analysis of IELTS score as a predictor of gain scores on the production of condition clauses in the production test . . 142
4.21 Research questions and hypothesis related to the development of past-counterfactuals . . . . . 143
4.22 Descriptive statistics for pre- and post-test scores on the acceptability judgment test for Verb, Particle and Pragmatic L1 groups. 144
4.23 Descriptive statistics for pre- and post-test scores on the production test for Verb, Particle and Pragmatic L1 groups (result clauses). . . . . 145

4.24 Descriptive statistics for pre- and post-test scores on the production test for Verb, Particle and Pragmatic L1 groups (condition clauses). . . . . 146
4.25 Statistical profile for ST14 for production and acceptability judgment tests. . . . . 154
4.26 Pre-test, post-test and task production for ST14 . . . . . 155
4.27 Statistical profile for ST4 for production and acceptability judgment tests. . . . . 159
4.28 Pre-test, post-test and task production for ST4 . . . . . 160
4.29 Statistical profile for ST7a for production and acceptability judgment tests. . . . . 164
4.30 Pre-test, post-test and task production for ST7a . . . . . 165
4.31 Statistical profile for ST8 for production and acceptability judgment tests. . . . . 169
4.32 Pre-test, post-test, task production and selections from the input for ST8 . . . . . 170
4.33 Statistical profile for ST7b for production and acceptability judgment tests. . . . . 174
4.34 Pre-test, post-test, task production and selections from the input for ST7b . . . . . 175
4.35 Summary of stimulated recall interviews (raw counts) . . . . . 179

5.1 Research questions and hypotheses related to the exact and repurposed use of selected result and condition clauses in task production . . . . . 195
5.2 Examples of exact repetitions of selected tokens . . . . . 196
5.3 Examples of repurposing operations on selected tokens. . . . . 198
5.4 Summary of the statistical differences between the use of exact clauses by the Analogous and Literal groups . . . . . 203

5.5 Summary of the statistical differences between the use of repurposed clauses by the Analogous and Literal groups . . . . . 203
5.6 Research questions and hypothesis related to the development of past-counterfactuals . . . . . 206
5.7 Summary of the statistical differences between the Analogous, Literal and Control groups on the acceptability judgment test . 206
5.8 Summary of the statistical differences between the Analogous, Literal and Control groups on the production test . . . . . 208
5.9 Research question related to the development of abstract past-counterfactual constructions . . . . . 209
5.10 Type and token frequency and the Involvement Load Hypothesis 222
5.11 Task sequence from the entrenchment of formulas to the development of patterns . . . . . 229

I.1 Raw counts and percentages for partial and full selection of condition and result clauses from the input for Analogous Group (3 tasks). . . . . 284
I.2 Raw counts and percentages of partial and full selection of condition and result clauses from the input for Literal Group (3 tasks). . . . . 285

# Conventions

Text in upper case refers to the meaning of a token, formula or construction, e.g. X CAUSES Y TO MOVE $Z_{PATH/LOC}$.

Text in italics refers to examples of phrases, technical terms and added emphasis in citations e.g. *If it hadn't been for the kids, I'd have gone.*

Categories such as NP, VP and Subj are in standard font, unless used in combination with concrete expressions. In this case, italics are used, e.g. *Give* $Obj_1$ $Obj_2$.

Subscripted categorial symbols refer to the limited scope of the slot. e.g. in a NP slot, subscripting the NP in *If it'd* $_{[VPNP]}$ indicates that only a restricted number of nouns are used in the slot.

In cited material, the original conventions are preserved.

# Chapter 1

# Introduction

How tokens in the linguistic environment lead to the development of structure is one of the key questions in First Language Acquisition (FLA), Second Language Acquisition (SLA), and cognitive science in general. In language acquisition, one can approach the question from the perspective that linguistic tokens in the input are impoverished sources of information for the development of structure and that a highly-complex rule system is not possible unless one invokes domain-specific linguistic devices (e.g. Chomsky, 1965; Eubank and Gregg, 2002; Fodor and Crowther, 2008; Hawkins, 2008; Pinker, 1989, 2004). An alternative view is that recourse to domain-specific devices is not necessary, as the frequency of tokens in the environment and general non-linguistic processes that act upon these tokens is sufficient to explain development (e.g. Bybee, 2006; Ellis, 2002; MacWhinney, 2004; Tomasello, 2003, 2004). As a first approximation, this latter approach needs to specify at least three attributes to explain a trajectory from token to structure. The first of these is the type and token frequency in the linguistic environment that act as data for acquisitional processes (Bybee, 2007; Ellis, 2002, 2009; Ellis and Ferreira-Junior, 2009b); the second is a description of the internal linguistic system at both the end point of acquisition and intermediate points, e.g. a 'taxonomic network of constructions' (Croft, 2001, p.25). The final characteristic is the processes that act as interfaces between tokens in the linguistic environment and the internal linguistic system. The general aim of this thesis is to explore the latter,

although it is also necessary to describe in some detail the possible changes to the internal system that are brought about by an interaction between tokens in the input and the processes that act upon them. To elaborate on this aim, the current study draws on usage-based accounts of language acquisition, and more specifically the role analogy may play in the development of structure from tokens.

Usage-based approaches offer a broad account of a link between tokens in the linguistic environment and the status of the internal system (e.g. Barlow and Kemmer, 2000; Bybee, 2006, 2010; Robinson and Ellis, 2008; Tomasello, 2000, 2003). In describing the internal system, a usage-based perspective on language knowledge rejects the distinction between words and rules, instead opting for a 'structured inventory of conventional symbolic units' (Langacker, 1987, p.73) with form-function units forming a continuum between concrete formulas and abstract constructions (e.g. Cowie, 1998; Culicover, 1999; Goldberg, 2003; Howarth, 1998; Nattinger and DeCarrico, 1992; Pawley and Syder, 1983). Unlike minimalist accounts of language such as those described by UG, usage-based approaches tend to employ a highly-redundant 'maximalistic language model' (Tummers et al., 2005, p.228) in which abstract constructions and exemplars of these constructions co-occur in the mind of the user.

Usage-based approaches maintain that the learning of such a maximalistic system 'result[s] from [the] accumulated experience with language across the totality of usage events' (Tomasello, 2000, p.61). The initial stage of learning is the repeated detection within usage events of concrete linguistic tokens that are attached to specific intentions, while abstract systems emerge from 'the collaboration of the memories of all of the utterances in a learner's entire history of language use and the frequency-biased abstraction of regularities within them' (Ellis and Larsen-Freeman, 2009, p.92). From this perspective, the learning trajectory for an internal system of symbolic units of varying degrees of abstraction is from the initial learning of concrete formulas, through the

development of low-scope patterns to abstract constructional patterns (Ellis, 2002; Tomasello, 2003).

A number of studies to date support this usage-based trajectory from formulaic sequences to rule-like behaviour in FLA (e.g. Dabrowska, 2001; Dabrowska, 2006; Lieven et al., 2003), and child SLA (Huang and Hatch, 1978; Wong Fillmore, 1976, and to a lesser extent Hakuta 1974). Although such a proposed trajectory has limited empirical support from earlier longitudinal studies in adult SLA (e.g. Hanania and Gradman, 1977; Shapira, 1978), suggesting that the initial learning and subsequent analysis of formulas is superseded by a much more powerful creative construction process (Krashen and Scarcella, 1978), the emphasis of these studies was chiefly focused on the acquisition of fairly abstract morphemes, with acquisition operationalised as the correct suppliance of morphemes in 80-90% of obligatory contexts. More recently, the methodological emphasis has shifted from the correct use of morphemes to the tracking of formulae through initial developmental stages (e.g. Eskildsen, 2009, 2012; Myles, Hooper and Mitchell, 1998; Myles, Mitchell and Hooper, 1999). Although far from conclusive, this shift has provided evidence that at least some learners establish a repertoire of formulaic chunks that show little flexibility, which they then break down, or 'unpack', to allow for more flexible use in a larger number of contexts.

The main focus of this study is on the motivation for the selection and use of tokens, and the processes that allow for a progression from tokens to formulas to more abstract units. Studies within the broad area of Task-Based Learning suggest that task demands are a catalyst for the selection of task-relevant tokens from the input (Robinson, 2005; Samuda, 2001), while more in terms of the development of already-learnt formulas, Myles at al. (1998; 1999, see also Wray, 2002) hypothesise that the breakdown of formulas is at least in part driven by the need to express functions which are beyond the capability of already-learnt formulas. Taken together, these studies suggest that task demands lead to the selection of tokens and that the selection and subsequent processing of

the tokens may lead to the development of more abstract patterns. However, the link between functional task demands and the mechanisms that allow for the development of patterns is unclear. Both Bybee (2010) and Tomasello (2003) suggest that one key mechanism in the development from formula to construction is the categorisation of unanalysed or partly analysed formulas based on principles of similarity and analogy. Drawing on Gentner's work on analogical reasoning (Gentner, 1983; Gentner and Markman, 1997, 2005), and specifically analogical encoding (Gentner et al., 2003, 2004), this thesis focuses on the initial stages of the functionally-driven breakdown of formulas by investigating the role that task demands and analogy have in the development from formula to low-scope patterns.

The literature review expands on the above overview by first considering a synchronic view of the abstractness of the linguistic system, and a diachronic trajectory from formula to low-scope pattern to abstract construction. More precisely, *Section: 2.1.1 A property theory as a continuum of abstractness* examines evidence for a continuum between formulas which show very limited abstraction from tokens in the environment to highly-productive abstract constructions (Croft, 2001; Croft and Cruse, 2004; Goldberg, 1995, 2006; Kay, 1997; Langacker, 1987), and subsequently considers the extent to which the degree of abstractness interacts with a proposed trajectory from formula to low-scope pattern to abstract construction (Ellis, 2002; Tomasello, 2003). *Section 2.1.2: Evidence for a trajectory* reviews to what extent a trajectory from formula to low-scope pattern to construction is supported by empirical studies in naturalistic child and adult second language acquisition and in instructed second language contexts.

Following on from the evidence for a trajectory, *Section 2.2: Processes in a Usage-Based Trajectory* focuses on the motivation for progression along the trajectory, and the categorisation processes that may interface between tokens in the linguistic environment and a developing internal system. The main focus of this section is how tokens in the input may be compared to produce

structure based on analogy and similarity (e.g. Gentner, 1983; Gentner and Markman, 1997; Gentner and Medina, 1997), and how the relative importance of these processes may vary depending on the stage of development between formula and construction. The literature review finally brings together the notions outlined above to develop research questions that consider the development of past counterfactual constructions from the analogous processing of past counterfactual tokens in the input.

# Chapter 2

# Literature Review

The main focus of this first section of the literature review is the proposal that there exists a typical developmental trajectory in SLA from formula to low-scope pattern to construction (N. Ellis, 2002). To discuss the extent of such a trajectory, it is first useful to distinguish between 'what is being acquired' and 'how it is acquired'. Drawing on Cummins (1983), Gregg (2003a,b) labels the 'what is being acquired' a *property theory*, and contrasts this with a *transition theory*. A property theory explains what it is that is known, while a transition theory attempts to explain how knowledge progresses from point A to point B.

After first outlining in more detail a trajectory in both FLA and SLA from a usage-based perspective, *Section 2.1.1: A property theory as a continuum of abstractness* considers a property theory that describes concrete formulas, low-scope patterns and abstract constructions, and proposes how a continuum between formulas and abstract constructions may impact on a developmental trajectory. *Section 2.1.2: Evidence for a trajectory* then considers to what extent the transition, or trajectory, from formula to low-scope pattern to construction is supported by empirical studies of naturalistic second language acquisition by both children and adults and in instructed second language contexts.

## 2.1 Usage and a trajectory from formula to construction

*Usage-based* accounts of language acquisition hold that the grammatical end state is the 'cognitive organization of one's experience with language' (Bybee, 2006, p.711). Language acquisition within such an approach is seen as gradual and piecemeal (Ellis, 2002; Tomasello, 2003), with structure emerging from experience, or usage events, within a complex linguistic and social environment. Without recourse to domain-specific linguistic devices such as those assumed by UG accounts, the burden on such accounts is that a highly complex mental grammar is formed from the input via simple domain-general internal processes that act upon it. FLA and SLA from this perspective involve the selection of unanalysed chunks of linguistic information, or *tokens*, the retention of tokens as *formulas* in the linguistic system, and subsequent analysis of these formulas, supported by frequent exposure, into more abstract *constructions*.

In the area of FLA, Peters (1985) provides substantial data in support of the *extraction* of tokens from the speech environment and the subsequent segmentation and extraction of morphosyntactic frames, at least in the initial stages of FLA. Similarly, Tomasello (2003) argues for the concept of *joint attention* between care-giver and child in the extraction of tokens from the environment, i.e. both the care-giver and mother share the same perceptual information and intentions, which allows the child to isolate tokens in the environment. These extracted tokens provide evidence upon which a set of pattern-finding skills such as categorisation based on similarity and analogy can act to create more abstract constructions. While there are distinct similarities with the research by Peters (1985), Tomasello (2003) suggests that the trajectory from token to abstract construction is not only evident in the initial stages of FLA, but it is also evident at later stages of development (for similar perspectives, see Barlow and Kemmer, 2000; Bybee, 2006, 2010; Langacker, 2000; Tomasello, 2000).

There is a growing body of evidence that supports Tomasello's view of a trajectory. Empirical studies in FLA provide evidence for the proximity of child utterances to the utterances of their caregivers and argue that a child's early knowledge of language is based on imitation or simple lexicalised frames, which are created via very simple processes such as juxtaposition and superposition of imitated formulas (Cameron-Faulkner et al., 2003; Dabrowska, 2001; Dabrowska and Lieven, 2005; Lieven et al., 2003, 1997).

Based largely on work in FLA, N.Ellis (2002) proposes that in SLA '[t]he typical route of emergence of constructions is from formula, through low-scope pattern, to construction' (p.143), i.e. tokens extracted from the input such as *give it me* become *formulas* in the linguistic system. With repeated exposure, these formulas develop into low-scope patterns such as *give it X*, which further develop into abstract schematic patterns such as V $\text{Obj}_1$ $\text{Obj}_2$. While Ellis' proposal posits a single process from formulas to abstract constructions, an alternative perspective is that of Krashen and Scarcella (1978) who opt for a dual-process approach:

> Prefabricated routines may evolve into patterns, but at the same time, independently, the creative construction process develops. (p.284)

In this account, formulas (termed *routines* by Krashen and Scarcella, 1978) develop into low-scope patterns, as in N.Ellis' account, but a separate process also exists for the development of what Krashen and Scarcella (1978) term *creative construction* or *propositional speech*. Although Krashen and Scarcella give little indication of what propositional speech or creative construction is or how it develops, in other works by Krashen (e.g. Dulay et al., 1982) the assumption is that there exists an innate Universal Grammar (UG), which is at the core of this creative construction.

> To say that humans are creative with respect to language means that they do not simply imitate what they hear. In fact, they often

> use sentences they have never heard before. This central aspect of language acquisition is believed to be rooted in innate and universal structural properties of the mind.
>
> Thus we use the term **creative construction** to refer to the subconscious process by which language learners gradually organize the language they hear, according to rules that they construct to generate sentences. The form of the rules is determined by mental mechanisms responsible for human language acquisition and use. These mechanisms appear to be innate. (Dulay et al., 1982, p.11, emphasis in original)

Formulas and low-scope patterns, from this perspective, are 'essentially and fundamentally different' (Krashen and Scarcella, 1978, p.298) from propositional language and are peripheral to core rule-based mechanisms. The notion that the core of language acquisition is fundamentally different is central to claims that abstract characteristics of grammar are not learnable from the input alone and therefore the input acts as evidence that *triggers* innate parameters (Gregg, 2003a; White, 2003). While UG is the most detailed property theory by far, very little is known about the relationship between the input and the hypothesised triggering of parameters. Gregg (2003a), while arguing against a trajectory from token to abstract construction without recourse to innate principles and parameters, sums up the status of research into triggering in SLA:

> . . . very little is even thought to be known about what specific stimuli in the input could act as the trigger for the setting of a specific parameter in a specific language. And in SLA, discussion of triggering, and of parsing in general, is close to non-existent. (Gregg, 2003a, p.848)

With so little evidence for the ways in which the input may trigger parameters, the line taken in this thesis is that although triggering may play a role in SLA, investigating a trajectory from formula to low-scope to construction is worthwhile in order to appreciate the extent to which it may be less peripheral than Krashen and Scarcella (1978) claim. With this in mind, *Section 2.1.1:*

*A property theory as a continuum of abstractness* reviews evidence for a synchronic description of language that takes into account concrete formulas, low scope patterns and abstract constructions. The literature review then builds on this description by reviewing evidence for a trajectory from token to formula, token to low-scope pattern and finally from token to constructional pattern.

### 2.1.1 A property theory as a continuum of abstractness

A usage-based perspective on language knowledge rejects the distinction between words and rules, instead opting for form-function units which form a continuum between concrete formulas and abstract constructions (Cowie, 1998; Culicover, 1999; Ellis, 2008; Goldberg, 2003; Howarth, 1998; Jackendoff, 2002; Nattinger and DeCarrico, 1992; Pawley and Syder, 1983). With such a continuum, a developmental trajectory will be minimal for more idiomatic expressions, whereas highly abstract constructions will follow N.Ellis' 'formula to low-scope pattern to construction' trajectory. For example, as a first approximation, Table 2.1 shows hypothetical trajectories for three constructions. The first in the table is the highly idiomatic *A case in point*, which shows a minimal trajectory from input tokens to formula. This end of the continuum is populated by concrete expressions such as individual lexical items and highly idiomatic expressions that show no productivity. At the other extreme there exist highly productive structures such as the caused-motion construction. Due to the abstract nature of this construction, a usage-based approach predicts a trajectory from fixed formula, such as *She pushed it off the top* to low-scope pattern *Subj pushed it off X* to fully abstract construction Subj V Obj $\text{Obl}_{path/loc}$. Between these two extremes of productivity lie low-scope patterns that see varying degrees of productivity and therefore a limited trajectory from formulas. The following sections review evidence for the claim that there exists a continuum from concrete expressions to abstract constructions by first reviewing formulas and low-scope patterns, and then extending the discussion to more abstract constructions.

**Table 2.1:** *Example developmental trajectories for tokens in the input*

| Example Input | Degree of Abstraction from the Input | | |
|---|---|---|---|
| | Formula | Low-Scope Pattern | Abstract Construction |
| A case in point | A case in point | NA | NA |
| The bigger the better | The bigger the better | The Xer, the Yer | NA |
| She pushed it off the top | She pushed it off the top | Subj pushed it off X | Subj V Obj $\text{Obl}_{pathlloc}$ |

#### 2.1.1.1 Formulas and low-scope patterns

An example of the more concrete end of the continuum is that of lexical sequences larger than the individual word which show little or no variability. Lyons (1968), for example, uses the term *ready-made utterances* to describe unanalysed expressions tied to specific situations with idiomatic interpretations (1-4).

1. How do you do?
2. Rest in peace.
3. Easy come, easy go.
4. All that glitters is not gold.

(Lyons, 1968, pp.177-8)

Similarly, Yorio (1980) lists idioms and routine formulas as two main types of *conventionalized expressions*, describing routine formulas as 'highly conventionalized pre-patterned expression[s] whose occurrence is tied to a more or less standardized communication situation' (p.434), giving the example of *It's not what you think* as a utterance that is tied to a compromising situation. Hakuta (1974, and also Krashen and Scarcella 1978) refers to the same phenomena as *prefabricated routines*, which he describes as memorised unanalysed wholes.

Fillmore et al. (1988) propose the categories *Unfamiliar pieces unfamiliarly arranged* and *Familiar pieces, unfamiliarly arranged* as examples of limited

productivity. The former category consists of idioms that contain lexical items that are unique to a certain idiom, for example the *kith* in *kith and kin* and the *main* in *with might and main* (1988, p.506). The second category is composed of lexical items that exist elsewhere in the language, but are idiomatically arranged. Examples of these include *all of a sudden* and *in point of fact* (Fillmore et al., 1988, p.508, see also Nattinger and DeCarrico 1992). Due to their highly idiomatic nature, the expressions outlined in the accounts by Lyons (1968), Yorio (1980) and Fillmore et al. (1988) above exhibit little productivity and therefore, unless overgeneralised, would not progress along a 'formula-low-scope pattern-construction' trajectory.

A third category cited by Fillmore et al. (1988) is *Familiar pieces, familiarly arranged:* Items in this category adhere to more general combinatorial principles, and consist of constituents that are found elsewhere in the language. They are, however, semantically idiomatic, e.g. *pull someone's leg* and *tickle the ivories*. Although many such semantically idiomatic expressions would seem to have no productivity and therefore show no progression from formula to a more abstract construction, Barlow (2000) argues that even the most fixed idiomatic expressions show sophisticated productivity. He gives the example of *Make hay while the sun shines*. The first three examples (5, 6 and 7) show a certain amount of productivity after the temporal adverb *while*. In the final two examples (8 and 9), the only fixed items are *make* and *hay*; the other elements are productive.

5. The Chiefs aren't the only NFL team that hasn't made hay while their quarterback shined.
6. Now is the season of savers' discontent, though borrowers should make hay while the warmer climate remains.
7. Hewden Stuart's plant hire business made hay while the rest of the construction industry suffered from a drought.
8. One of the areas where the group has made hay in recent years.
9. ... This field is wide open to Labour. If Tony Blair cannot make hay in such political sunshine, how will he fare when winter comes?

(Barlow, 2000, pp.337-8)

The differences in productivity between 'kith and kin' and 'make hay' can be seen as a distinction between *substantive* and *formal idioms* (Fillmore et al., 1988). Substantive or *lexically-filled* idioms are fully specified, e.g. there is no variability in the idioms *all of a sudden* (**everything of a sudden, *all of a quick*) and *in point of fact* (**in point of truth*), nor in *kith and kin*. Formal or *lexically-open* idioms can be partially substantive and partially open, such as in partially-productive phrases *The Xer, the Yer* and *pull one's leg*. In the example *The Xer, the Yer*, *the* is substantive, but X and Y are lexically open. The example *pull one's leg* has the substantive elements *pull* and *leg*, but is lexically open in terms of whose leg, i.e. *pull* $NP_{pro}$ *leg*.

A number of other authors make the distinction between substantive and formal idioms, although using different terminology. Lyons (1968), for example, makes a clear distinction between the *ready-made* utterances mentioned previously and *sentence schemata* (pp.177-8). While the former he categorises as unanalysed expressions, sentence schemata (Examples 10 and 11) and phrase schemata (Example 12) have some degree of productivity in one or more of the 'slots'. Similarly, Hakuta (1974, see also Krashen and Scarcella 1978 for the same terminology) makes a distinction between *prefabricated formulas* which see no variability and *prefabricated patterns* which see some paradigmatic variability.

10. What's the use of . . . ing?
11. Down with . . . !
12. for . . . 's sake.

(Lyons, 1968, pp.177-8)

Pawley and Syder (1983) consider that the idea of idiomaticity extends much further beyond what are usually considered idioms. To illustrate this, they use the example *I want to marry you* and several paraphrases (Pawley and Syder,

1983, p.186). The paraphrases in 13-20 are all grammatical possible sentences in English. However, *I want to marry you* is the idiomatic term for the concept. The others, while possible, are not usually performed[1].

13. I wish to be wedded to you.
14. I desire you to become married to me.
15. Your marrying me is desired by me.
16. My becoming your spouse is what I want.
17. I want marriage with you.
18. What is desired by me is to wed you.
19. I, who am speaking, want to marry you, whom I am addressing.
20. It is my wish that I become married to you. etc. etc.

(Pawley and Syder, 1983, p.186)

Pawley and Syder (1983) argue that idiomaticity 'rests to a considerable extent on knowledge of a body of *sentence stems* which are *institutionalized* or *lexicalized*' (p.191). In line with Fillmore et al.'s distinction between substantive and formal idioms, these lexicalised sentence stems are not concrete rigid expressions, but partially abstract frames that allow for degrees of productivity (See Table 2.2).

Sinclair (1991) likewise puts significant emphasis on idiomaticity by making the distinction between *the open-choice principle* and *the idiom principle*. The open-choice principle can be seen as a productive grammar, with a 'very large number of complex choices' (Sinclair, 1991, p.109) at junctures between words, phrases or clauses. The idiom principle, on the other hand, is defined as 'a large number of semi-preconstructed phrases that constitute single choices' (p.110). Sinclair sees the idiom principle as the 'primary model' in the construction and interpretation of texts, with the open-choice principle called upon at junctures between semi-preconstructed phrases. Wray (2002), while using different terminology, offers a similar distinction between what she terms *language grammar*

[1]A Google search of the phrase *My becoming your spouse is what I want* shows 32 hits, all of which are references to Pawley and Syder (1983), while *I want to marry* you shows over 13 million hits (March 2012).

**Table 2.2:** *Examples of lexicalised sentence stems (Pawley and Syder, 1983)*

| **Lexicalised sentence frame** | **Example** |
|---|---|
| NP be-TENSE sorry to keep-TENSE you waiting | I'm sorry to keep you waiting.<br>I'm so sorry to have kept you waiting.<br>Mr X is sorry to keep you waiting all this time. |
| NP tell - TENSE the truth | Tell the truth!<br>Jo seldom tells the truth.<br>I wish you had told me the truth. |
| Why do-TENSE n't NPi pick on someone PROi-gen own size! | Why doesn't that bully pick on someone his own size?<br>Why don't you pick on somebody your own size? |
| Who (the EXPLET) do-PRES NPi think PROi be-PRES! | Who the hell do you think you are?<br>Who does that woman think she is?<br>Who the fuck do you think you are?<br>Who does George Bush think he is? |

and *formulaic sequences*, the former she describes as 'the ability to create and understand entirely novel strings' (Wray, 2002, p.12) and the latter as:

> a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar. (p.9)

Wray's distinction between formulaic sequences and language grammar seems representative of a number of positions that make the distinction between on the one hand formulas and low scope patterns, and on the other a productive grammar. As can be seen in Table 2.3, early descriptions such as those of Krashen and Scarcella (1978) and Lyons (1968) put very little emphasis on formulas and low-scope patterns, and instead placed the main emphasis on a generative rule-based grammar. While maintaining a distinction between what Wray refers to as formulaic sequences and the language grammar, Pawley and Syder (1983), Wray (2002) and Sinclair (1991) put significantly more emphasis on the idiomaticity of language.

**Table 2.3:** *Summary of formulas, low-scope patterns and alternative mechanisms*

| Author | Formula | Low-Scope Pattern | Other generative principles |
|---|---|---|---|
| Lyons (1968) | Ready-Made Utterances | Phrase and Sentence Schemata | Formula and low-scope patterns play peripheral roles. Generative grammar is core. |
| Hakuta (1974) | Prefabricated Routines | Prefabricated patterns | Prefabs are at 'the surface structure level', suggesting a secondary deeper mechanism |
| Krashen and Scarcella (1978) | Prefabricated Routines | Prefabricated patterns | Prefabs expedite communication, but propositional/creative language is core |
| Fillmore et al. (1988) | Substantive idioms | Formal idioms | Later work by Fillmore argues for no division between idioms and propositional language (Kay and Fillmore, 1999) |
| Pawley and Syder (1983) | Memorised expressions | Lexicalised sentence stems | 'Productive rules' (p.219), but lexicalised sentences stems provide 'the main building blocks of fluent connected speech' (p.214) |
| Sinclair (1991) | The idiom principle (primary) | | The open-choice principle (secondary) |
| Wray (2002) | Formulaic sequences | | The language grammar, i.e. 'the ability to create and understand entirely novel strings' (Wray, 2002, p.12) |

Based on this overview of fixed or semi-fixed expressions, in this thesis, individual elements or sequences of elements in the linguistic environment, either written or spoken, are referred to as *tokens*. If a token is memorised but unanalysed, it is termed a *formula*, while memorised and partially abstract sequences of elements will be termed *low-scope patterns*. In terms of the 'Other generative principles' in Table 2.3 above, the following section discusses the role of construction grammars in describing the more productive end of the continuum, and its relationship to formulas and low-scope patterns.

#### 2.1.1.2 Construction grammars

Although the above characterisation of low-scope patterns and formulas suggests a considerable role for these in any description of the linguistic system, many descriptions of language systems separate out a maximally productive *core* of combinatorial principles, and a *periphery* of marked exceptions (Chomsky, 1965, 1972, 1995; Pinker, 1998). The emphasis in such a dual system has often been the generative core, with the exceptions relegated to the 'unassimilable residue of odd facts, exceptions, and quirky constructions' (Riemsdijk and Williams, 1986, p.175).

The issue with separating the core and the periphery, however, is the extent to which low-scope patterns are productive and therefore the extent to which this productivity can extend into core areas of grammatical productivity. Langacker (1987) proposes that any distinction between a core rule system and lexical items is a 'fallacy', and strongly suggests a continuum of productivity from formulas at one end to abstract constructions at the other. This idea of a continuum from formula to construction is shared by a set of closely related cognitive-functional theories which see constructions as central elements (Croft and Cruse, 2004; Fillmore et al., 1988; Goldberg, 1995, 2006; Kay, 1997; Lakoff, 1987; Langacker, 1987, 2000).

A central tenet of construction grammars is that elements of a language are seen as symbolic units of form and meaning at all levels of complexity and

granularity, including basic morphemes, words and idioms and partially and fully productive phrasal patterns. An item is given *construction* status if it is not predictable from more abstract constructions. In Goldberg's terms (1995) a construction is:

> . . . a pairing of form with meaning/use such that some aspects of the form or some aspect of the meaning/use is not strictly predictable from the component parts or from other constructions already established to exist in the language. (p.4)

With this description in mind, individual lexical items and fully substantive (i.e. non-productive and unpredictable) idioms, for example, are arbitrary correspondences between form and meaning, which are not wholly predictable, and therefore can be considered constructions (Fillmore et al., 1988; Jackendoff, 2002). The same status would also seem to apply to the lexicalised sentence stems such as *NP be-TENSE sorry to keep-TENSE you waiting* and *Why do-TENSE n't NPi pick on someone PROi-gen own size!* described by Pawley and Syder (1983) and the formal idioms described by Fillmore et al. (1988).

As championed by Langacker (2000), proponents of construction grammars go further than emphasising the extent of formulas and low-scope patterns by dispensing with *any* distinction between lexical item and a core rule system[2], arguing that what are often considered rules are in fact abstract constructions situated at the productive end of a continuum of productivity from formula to construction. This can be seen in Goldberg's analysis of the abstract *caused-motion* construction (Examples 21, 22 and 23).

21. He pushed the woman out of the window
22. She pulled the car along the road
23. She sneezed the foam off the cappuccino (Goldberg, 1995)

[2]Although formulas and low-scope patterns are considered constructions, in this thesis the term 'construction' will be reserved for abstract patterns, while the term 'formula' and 'low-scope' pattern will be retained.

Goldberg argues that the semantic meaning of this construction comes from both the elements that constitute the construction (the lexical items, also considered constructions in their own right) and the construction as a whole. This division is not easy to distinguish in such examples as 21 and 22, for which it could be argued that the default meaning of the verb includes an object and a path. However, in Example 23, *sneezed* is normally an intransitive verb, and therefore its listing in the lexicon is unlikely to be sufficient in the interpretation of the meaning of the whole sentence. For a full interpretation, Goldberg (1995) specifies a construction which pairs a meaning of physical or metaphorical TRANSFER with the formal elements SUBJ VP OB OBL (Table 2.4).

**Table 2.4:** *Examples of the caused-motion construction with meaning assigned to the entire construction (Based on Goldberg 1995)*

| **Meaning and Form** | **Example** |
|---|---|
| X CAUSES Y TO MOVE | He laughed the teacher out of the class |
| $Z_{PATH/LOC}$ | She punched the cat off the wall |
| $\Updownarrow$ | He pulled the car along the road |
| $SUBJ\ VP\ OB\ OBL$ | He moved the fish off the boat |
| | She put the five pound note into her pocket |
| | She pulled the cork out of the bottle |
| | They dragged the man into the house |

As illustrated in Table 2.5, the caused-motion construction can be seen as an exemplar of the abstract end of a productivity continuum, with formulas at the opposite end. Here, morphemes, words, complex words and filled idioms are fully substantive constructions. Partially-filled idioms and the covarational-conditional structure (*The Xer the Yer*) are substantive in parts (e.g. *jog* and *memory* and *The... er, the... er*), and lexically-open, or schematic in parts (e.g. the NP in *jog someone's memory*). Finally, the caused-motion construction, ditransitive and passive constructions contain no substantive morphological material, and are therefore fully schematic constructions. The end result of developing such a continuum between the idiosyncratic and the fully schematic

is the claim that all of the language can be described within a constructionist framework (Croft, 2001; Croft and Cruse, 2004; Goldberg, 1995, 2006; Kay, 1997; Langacker, 1987).

**Table 2.5:** *A continuum of constructions from concrete to abstract (Goldberg 2003)*

| Abstractness | Construction | Form/Example | Function |
|---|---|---|---|
| Concrete | Morpheme | e.g. anti-, pre-, -ing | |
| Concrete | Word | e.g. Avocado, anaconda | |
| Concrete | Complex Word | e.g. Daredevil, shoo-in | |
| Concrete | Idiom (filled) | e.g. Going great guns | |
| Partially abstract | Idiom (partially filled) | e.g. Jog (someone's) memory | |
| Partially abstract | Covariationai-Conditional construction | Form: The Xer the Yer (e.g. The more you think about it, the less you understand) | Meaning: linked independent and dependent variables |
| Fully abstract | Ditransitive (double-object) construction | Form: Subj V Obj1 Obj2 (e.g. He gave her a Coke; He baked her a muffin) | Meaning: transfer (intended or actual) |
| Fully abstract | Passive | Form: Subj aux VPpp (PPby) (e.g. The armadillo was hit by a car) | Discourse function: to make undergoer topical and/or actor non-topical |

In terms of the development of such constructions, both Culicover (1999) and Jackendoff (2002) have suggested that the learning of individual linguistic items in the L1 is similar to the learning of abstract generalisations, as both are considered lexical items. The difference is that the concrete linguistic material in an individual item is replaced by a variable in the abstract construction. If this is the case, then the learning of individual lexical items will have a minimal trajectory (from token to formula), while a low-scope pattern will have

a greater trajectory from token to formula to low-scope pattern, and finally an abstract construction such as the caused motion construction will involve the 'full' trajectory from formula through low-scope pattern to (caused motion) construction.

#### 2.1.1.3 Summary

Construction grammars originally offered a description of idioms that had proved difficult for mainstream approaches. More recently, however, they have been extended to move beyond the periphery to cover more central aspects of language (for example, argument structure constructions). Whether construction grammars can offer the grammatical depth described by Universal Grammar (UG) is a contentious point (the degree of agreement between Goldberg and Jackendoff in Goldberg and Jackendoff (2004) highlights how far a constructionist approach could go), but what they do offer is a useful description of an end point of a trajectory which takes into account substantive, fully instantiated form-function mappings through semi-productive and semi-schematic morphological and phrasal elements to fully schematic and productive schemas.

This view of the representation of language is inextricably linked to the trajectory from formula to construction. As approximated at the beginning of this section, at one end of the continuum the trajectory is minimal, as there is very limited abstraction from tokens in the input (minimal in the sense that a stored type is always an abstraction from tokens in the linguistic environment), i.e. words, inflexible lexical phrases and highly idiomatic language do not develop into patterns. Low-scope patterns such as partially-filled idioms take a trajectory from formula to pattern, while abstract constructions at the extreme end of the spectrum, e.g. caused-motion constructions, develop along a trajectory from formula through low-scope pattern to construction. This interaction between a construction's degree of abstractness and a trajectory can

be seen in Table 2.6, which summarises the interaction into three broad trajectories. The first trajectory is from token to formula, the second is from token to formula to low-scope pattern, and finally from token to formula to low-scope pattern to construction. This latter trajectory corresponds to N.Ellis' formula to low-scope pattern to construction (2002).

**Table 2.6:** *Degree of linguistic productivity/concreteness and its effects on a trajectory from fomula to low-scope pattern to construction*

| Author | Input Tokens | Degree of Abstraction from the Input | | | | |
|---|---|---|---|---|---|---|
| | | Formula | → | Low-Scope Pattern | → | Abstract Construction |
| Lyons (1968) | *Rest in peace* | → *Rest in peace* | | NA | | NA |
| Yorio (1980) | *It's not what you think* | → *It's not what you think* | | NA | | NA |
| Fillmore et al. (1988) | *kith and kin, with might and main* | → *kith and kin, with might and main* | | NA | | NA |
| Lyons (1968) | *Down with the King!* | → *Down with the King!* | → | *Down with . . . !* | | NA |
| Fillmore et al. (1988) | *I'm pulling your leg* | → *I'm pulling your leg* | → | *subj pull NP leg* | | NA |
| Pawley and Syder (1983) | *I'm sorry to keep you waiting* | → *I'm sorry to keep you waiting* | → | *NP be-TENSE sorry to keep-TENSE you waiting* | | NA |
| Goldberg (1995) | *She pushed it off the top* | → *She pushed it off the top* | → | *Subj pushed OB1 off Obl* | → | Subj VP OB OBL |
| Goldberg (2003) | *He's called John* | → *He's called John* | → | *Subj Aux called/known /done (PPby)* | → | Subj aux VPpp (PPby) |

### 2.1.2 Evidence for a trajectory

This section considers empirical support for the development from formula to low-scope pattern to construction. To do this, data is presented mainly from longitudinal studies in SLA with both children and adults in naturalistic and classroom contexts.

#### 2.1.2.1 Evidence for a trajectory in naturalistic SLA

**2.1.2.1.1 Formulas and child SLA** In a longitudinal study of the untutored language development of 'Uguisu', a 6-year old Japanese girl temporarily resident in the United States, Hakuta (1974) focused on the progression of both *prefabricated routines* and *prefabricated patterns* in Uguisu's speech.[3] Hakuta traced three patterns over a 15-month period - the copula *be*, the phrase *do you* in question forms and *how to* in embedded *how*-questions. He found that Uguisu displayed rule like behaviour at a local level, i.e. variability of certain slots in an otherwise fixed expression. Her use of copula *be*, for example, did not display systemic rule-like behaviour such as number agreement, but there was evidence of productivity at the local level in patterns such as *this is*, *these are*, *it's*, *I'm* and *NP is*. This restricted use was also apparent in the use of *do you* question forms (Examples 24-27), which in the second month of the study Uguisu used in certain contexts but did not not extend the structure to the third person (Examples 28-30).

24. Do you know?
25. How do you do it?
26. Do you have coffee?
27. Do you want this one?

[3]Recall that prefabricated routines show no variability, whereas prefabricated patterns show limited variability in slots within a pattern, cf. the distinction made by Lyons (1968, pp.177-8) between 'ready made utterances', and the more open 'sentence schemata', and also between Fillmore et al. (1988)'s distiction between 'substantive' or lexically-filled idioms and 'formal' or lexically-open idioms. Note that in this thesis, prefabricated routines are classified as formulas, and prefabricated patterns as low-scope patterns.

28. What do you doing, this boy?
29. What do you do it, this, froggie?
30. What do you doing?

(Hakuta, 1974, p.293)

In terms of the development of questions, although Uguisu did not successfully develop third person reference in question forms for the duration of the study, there was evidence for the development of past tense morphology. At the start of the study, *do you* was used exclusively in the present tense. In month 4, the past tense appears but in the context of *where*, at month 5 in the context of *what*, and in month 6 with *how*. Hakuta attempts to explain these findings by suggesting that Uguisu utilised a 'strategy of learning on the surface structure level: learning through rote memorization of segments of speech *without knowledge of the internal structure of those speech segments*' (p.287, emphasis added). This argument suggests a secondary process that develops the 'internal structure of... segments', in line with the dual process account of Krashen and Scarcella (1978). However, although Hakuta describes Uguisu's interlanguage as lacking in internal structure, it does depend what is meant by this. In Hakuta's terms, structure consists of abstract elements such as third-person *-s*. However, if we describe structure as low-scope patterns, Uguisu did show considerable knowledge, such as *NP X*, *This is X* and her knowledge of past tenses frames. On these terms, Uguisu's development provides some evidence for the a trajectory from initial formulas to low-scope patterns, at least for these patterns.

More difficult to explain is the development of *how to* in embedded *how*-questions in the same study. In months 3 to 6 of the study, Uguisu produced formulas with little variability, such as *I know how to do it* (p.295). These formulas disappeared from Uguisu's production, replaced by non-conventional *how to* phrases such as *I know how to you write this* (p.295). Later on, in months 11 to 15, conventional use could be seen again in 50% of utterances. As

Hakuta notes, this leads to the question of why original formulas may disappear when more creative construction is evident:

> ... we would expect that if she [Uguisu] were relying solely on the surface patterns of her peer speech, she would stay with what she had. But she did not; gradually the proportion of correctly used *how to* forms went from nearly 100 percent in months III to VI, to a zero to 50 percent correctness in months XI to XV. (p.295)

If there was a development from formula to low-scope patterns to abstract constructions, this U-shaped development is difficult to explain. The disappearance of formulas could be due to a secondary creative process acting independently to pattern development, as in Krashen and Scarcella (1978)'s dual-process explanation. Based on Hakuta's data, it is difficult to speculate, but one possibility for the apparent disappearance of the routines could be Hakuta's methodology. Identifying unanalysed formulas in production data is relatively simple compared with identifying low-scope routines/patterns, as these low-scope patterns share very similar structure to more creative constructions, and therefore these patterns may have been classed as creative construction (cf. Krashen and Scarcella, 1978). For example, *I know how to...* could be considered a formula which is juxtaposed on to erroneous forms such as *... you write this*. In Hakuta's study, this type of production did not contribute to the findings, as the emphasis was on correctness. Speculating, if the emphasis had been the use of formulas, then there is a possibility that the findings would have shown a trajectory from rigid formula anchored to one context to the extended use of the formula in different contexts and with different cotexts. As the data stands, however, there would seem to be development of formulas into patterns such as *NP X*, *This is X* and past tense question forms, but to what extent there is evidence for a secondary process of creative construction is unclear.

A study which gives more support for a trajectory is that of Huang and Hatch (1978), who followed the formula development of a Taiwanese boy called 'Paul Chen'. Huang tracked Paul's production of interrogatives over a 4½

month period while he was enrolled in a play school with English-speaking children. For the first month, Paul's productions were unanalysed imitations of utterances he had heard, often overextending their function to inappropriate contexts. In subsequent months, Paul moved from these imitations to more complex productions. Table 2.7 shows Paul's development broken into three stages. In stage one, words from carers are echoed (X, at the top of Table 2.7), questions are imitated ($A_1$), and creative speech ($C_1$) is limited to rising intonation and two-word combinations to signal questions. In the second stage there were fewer imitated questions, and more creative constructions, notable the use of the chunks *where's* and *what's*, with the contracted copula attached. These constructions seem to be more productive versions of imitated questions from the input in Stage 1 and 2, seen in the 'Key aspects of Input' column. In stage 3, correctly inverted *Can*-questions ($C_a$) became productive based it seems on the inverted *Can*-questions in the input, as Paul did not go through a stage of producing non-inverted *can*-statements first. Interestingly, however, Huang and Hatch note that Paul did not generalise from inverted Copula questions (e.g. *Are you ready?*). For these types of questions, Paul went from two-word constructions to statements to inversion questions.

Huang and Hatch's study is important, as it carefully documents the input and output and therefore can make tentative correlations between tokens in the input, imitation of these tokens and more creative construction. Questions such as *Can you?*, *Where's*, *What's* were frequent in the input, which seemed to lead first to Paul imitating them and secondly to him breaking them down and using them more productively. As Huang and Hatch note, why *these* questions seemed to break down, while other questions took a different route is open to question.

Again, one could conclude, as Krashen and Scarcella (1978) do, that the development of formulas into low-scope patterns is an entirely different process to the development of propositional speech. The *Can*, *What* and *Where* questions support the prefabricated trajectory, but the development of other

question forms such as the copula does not, at least not on the basis of the data in Huang and Hatch (1978). The development of these latter questions could be attributed to Krashen and Scarcella's 'language acquisition process', but what the data in Huang and Hatch (1978) seems to show is that while at times there appears to be a trajectory from the input to imitated formulas to creative construction, at other times, the process is more complex and not directly readable from this type of data.

**Table 2.7:** *Paul's production question forms (Extended from Huang and Hatch, 1978, p.127, to include key aspects of input)*

| **Stage 1** | **Development** | **Key aspects of Input** |
|---|---|---|
| X:[1] | Table? This? Walk? Wash hands? Two cat? | Do you want to X? |
| $A_1$ | Question repetition | e.g. go home, help us, eat that |
| | What do you like? How are you doing? What are you doing? What's my name? | Can you X? e.g. kick it, drive a truck, say 'teacher' |
| $B_1$ | Questions learned via imitation | Where's X? |
| | Are you ready? Are you going too? Hi, how are you? What? What's that? What's your name? Where's Bobby? Which way? | |
| $C_1$ | Rule application. (Q =rising intonation) | |
| | This+++slipper? joe, this+++ball? Ball+++doggy? Okay? Fish+++see? | |
| **Stage 2** | | |
| $A_2$ | Question repetition | What's this/that? |
| | Which one? Where are you going? | Is this a X? |
| $B_2$ | Questions learned via imitation | Do you want X? |
| | What's this? Whose is this? This one? That one? What now? | Can you X? |
| $C_3$ | Rule application (Q =rising intonation; where's, what's, whose are question markers, "okay" is a tag question marker) | What are you. . . ing? Are you. . . ing? |
| | a This+++orange? This+++you? This+++not truck? This+++yours book? | Who/Whose questions |
| | b Where's Kenny? Where's pen? Where's car? Where's Teddy's car? Whose truck? | |
| | c This is my book? This is jacket? This is flower? | |
| **Stage 3** | | |
| $B_3$ | May I be excused? | *what* questions |
| $C_3$ | Rule application (addition of can questions, how many, color, and *be*-inversion) | e.g. What's this/that? |
| | a Can K have some juice? Teddy, can I play? Can I write my name? Jim, can you play with the ball? | What's the baby doing? What is it that you want? |
| | b You put the belt on? You see a flag? B is a boy? You want this one? | *would*-questions. |
| | c Is this yours house? Are you a good boy? What am I doing? | e.g. Would you give me one please? |

*immediate repetition speech, but later used in appropriate situations

Arguably the most detailed study of a trajectory from formula to creative construction in child SLA is that of Wong Fillmore (1976), who documented the use and breakdown of formulas of 5 Spanish-speaking children over the first nine months of their residence in the USA. As can be seen from Table 2.8, she found extensive use of formulas over the full nine months, with the first time period seeing the greater number[4].

**Table 2.8:** *Percentage of formulaic speech for five subjects over a nine-month period (simplified from Fillmore, 1976)*

| **Time** | **Nora** | **Ana** | **Alej** | **Jesus** | **Juan** |
|---|---|---|---|---|---|
| Time 1 (Oct-Nov) | 53% | 58% | 79% | 87% | 100% |
| Time 2 (Dec-Jan) | 52% | 58% | 70% | 84% | 93% |
| Time 3 (Feb-March) | 40% | 55% | 65% | 87% | 88% |
| Time 4 (April-May) | 37% | 66% | 63% | 78% | 81% |

Importantly, Fillmore presents data that suggested that the learning of formulas led to the development of patterns. For example, Table 2.9 shows the trajectory of the phrase *How do you do dese?*, initially learnt as a chunk by 'Nora' (Time 2). The first piece of evidence that the formula was used creatively is at Time 3-4, which saw Nora juxtapose *How do you do dese?* with further items to extend its meaning, e.g. *How do you do dese flower power?*. This was followed by flexibility in the verb phrase slot (e.g. *How do you make a little...*, *How do you like...*, *How do you gonna do...*) and also flexibility in the auxiliary (*do/did*). In the final stage of the study (Time 4), Nora seemed to have freed up *How* and showed increased flexibility in the auxiliary slot (*do/does/did*), as well as adding alternative verbs such as *cut*, *make* and *work*.

The trajectory offered by Wong Fillmore (1976) is worth quoting at length as it is very much in line with the trajectories proposed by N.Ellis (2002) and

[4]Wong Fillmore (1976, p.641) notes that if subjects or predicatives were changed in the production, it was then classed as productive (not formulaic), even if the subject, e.g. *dese one* and the predicate *put it right here* seemed formulaic. This suggests that if a more liberal interpretation of formulas was adopted in which formulas were not solely interpreted as sentence frames but also, for example, as subjects, the number of formulaic phrases in production would be higher.

**Table 2.9:** *Nora's breakdown of formulaic speech over a nine-month period (Wong Fillmore, 1976, p.646)*

| **Structure** | **Examples** |
|---|---|
| **Time 2** | |
| $Wh_{[F]}$[a] How do you do dese? | How do you do dese? |
| **Time 3-4** | |
| $Wh_{[Fx]}$[b] | How do you do dese? |
| How do you do dese (X)/ X= NP, PP[c] | How do you do dese September por mañana? |
| | How do you do dese flower power? |
| | How do you do dese little tortillas? |
| | How do you do dese in English? |
| $Wh_{[Fx]}$: | How do you make a little gallenas? (='bollenas') |
| How do/did you X/X=VP | How do you like to be a cookie cutter? (='How would you...') |
| | How do you like to be a sharks? |
| | How do you make the flower? |
| | How do you gonna make dese? |
| | How do you gonna do dese in English? |
| | How did you make it? |
| | How did you lost it? |
| **Time 4** | |
| $Wh_{[Fx]}$ | How do cut it? How do make it? |
| How do/does/did X/X=clause | How does this color is? |
| | How did dese work? (='How does this work?') |
| $Wh_{[S]}$[b] HOW is freed, preposed | Because when I call him, how I put the number? (='How will I dial his number?') |
| | How you make it? |
| | How will take off paste? |

a: F denotes a fully formulaic phrase with no variables
b: Fx denotes a partly formulaic phrase; S, a productive (non-formulaic) sentence
c: X denotes a category, e.g. in 'How did you do dese X/X=NP, PP, 'How did you do dese' is followed by the category 'X', and that category is a noun phrase or a prepositional phrase

Tomasello (2003). The first prerequisite in the trajectory in the categorisation of two similar forms:

> [The analysis of formulaic expressions] involves noticing which parts of these formulaic expressions are like parts of other utterances the learner knows or hears, or noticing variations of these utterances in the speech of others. For example, Nora had in her speech repertory two related formulas: *I wanna play wi' dese* and *I don' wanna do dese*. No doubt the similarity of these expressions allowed her to discover that the constituents following wanna were interchangeable, and that she could say "I don' wanna play wi' dese" and "I wanna do dese." (Wong Fillmore, 1976, pp.644-5)

The next stage builds on this categorisation to produce low-scope patterns:

> As soon as she realized that these phrases were interchangeable, she was on her way to discovering that similar phrases could be inserted. At that time these formulas became formulaic frames with analysed slots: *I wanna X/X=VP* and *I don't wanna X/X=VP*. [...] (ibid.)

The next stage is very much one of analysis in which individual constituents could be freed from the formula:

> [T]hrough a gradual process of figuring out which of the formulas could be varied, children were able to free the constituents in many of them to become units in productive constructions. In the above formulas, the phrase *play wi' X/X=NP* became a formulaic verb phrase unit which could be used in the verb phrase slot of other frames such as Le's [*sic*] X/X=VP (e.g. "Le's play wi' that one") and in productive constructions as well: *NP + VP*, "She's play wi' dese." Thus the analytical process carried out on formulas yielded formulaic frames with abstract slots representing constituent types which could substitute in them, and it also freed constituent parts of the formulas to function in other constructions either as formulaic units or as wholly analyzed items. (ibid.)

The final stage sees considerably more productivity, with the concrete constituents disappearing leaving an abstract 'frame':

> Finally, when all of the constituents of the formula have become freed from the original construction, what the learner has left is an abstract structure consisting of a pattern or rule by which he can construct like utterances. (ibid.)

**2.1.2.1.2 Summary** The pattern that emerges from these longitudinal studies with children learning English as a second language is that formulaic phrases are prevalent throughout the course of acquisition, and are particularly visible in the first stages. It is also apparent that at least some formulaic phrases develop into patterns through processes such as juxtaposition or substitution.

The existence of a secondary process, in addition to the analysis of formulas, is a moot point. Hakuta (1974) questions why formulas would disappear, which

could be interpreted as evidence for a secondary process developing, as put forward by Krashen and Scarcella (1978). However, distinguishing between creative construction and formula development is methodologically difficult, as the surface shape of the propositional utterances are often unlike the original formulas (Paul in Huang and Hatch (1978) going from formulaic phrase to two-word constructions, for example). Whether this is due to a secondary process or the lack of more fine-grained methods of analysis is unclear. What seems clearer is that there is evidence for a trajectory in the production data of children learning a second language at least with some constructions (e.g. *How do you do dese?* in Wong Fillmore, 1976).

**2.1.2.1.3 Formulas and adult SLA** An early study of naturalistic adult SLA was that of Hanania and Gradman (1977), who sampled the communicative speech of a 19-year old Saudi Arabian woman called 'Fatmah', over a period of 18 months in the US. Fatmah knew no English at the start of the study and in the early stages her speech 'consisted mainly of memorized items that are commonly used in social contexts or with children.' (p.78). Hanania and Gradman (1977) comment that '[t]he use of these expressions, however, does not imply that she recognized the individual words within them, or that she was able to use the words in new combinations' (ibid.). While this suggests that the initial stage of acquisition is formulaic, the subsequent phases in Fatmah's development do not show signs of these chunks breaking down into more flexible patterns. It would seem that later on in the study, Fatmah produced two-term constructions (e.g. name+location), and later longer utterances. These two-word and longer utterances included such expressions as *No English* (Period 1, out of a total of 6 periods of the 18-month study), *Tahani good girl* (Period 2), *Mommy coming* (Period 4), *Go wash the hand* (Period 5) and *The dog eating* (Period 6).

Although there were what appeared to be formulaic phrases occurring throughout the study, it is difficult to read into the data a trajectory from formula to

pattern. Krashen and Scarcella (1978), commenting directly on Hanania and Gradman's study, take this as evidence that 'propositional language [in SLA] appears to be independent of gestalt[5] language' (p.288). Although this may be the case, Hanania and Gradman (1977) explicitly state that the focus of their analysis was morpheme development, not formulas, and therefore the formulas that were documented were representative only of the fact that they actually existed in Fatmah's speech, and not a systematic study of formulas and low-scope patterns in Fatmah's productions (a point also made by Wray, 2002). Nevertheless, the non-standard forms that Fatmah produced such as *No English* do appear to be single lexical items glued together with some form of incipient interlanguage grammar, not emerging formulas.

A similar longitudinal case study is that of Shapira (1978), who traced the morpheme development of 'Zoila', a 25-year old Guatemalan woman working as a housekeeper in the US. Over a period of 18 months, ten morphemes were tracked[6], with acquisition operationalised as the suppliance of a grammatically correct form in 80-85% of obligatory contexts. Shapira found 'practically no evidence for her progress in [Zoila's] acquisition of English grammar' (p.253). There was, however, some evidence that formulas were becoming productive. The formula *I don't know* was used early on in the study as a set phrase, while later on it appeared to become more flexible by allowing a number of verbs, in addition to *know*, into the VP slot, including *see*, *care*, *do*, *think* and *say*. Apart from the one example in the data of *I don't know*, however, there is no other evidence that formulas developed into low-scope patterns. This is not surprising, however, as the emphasis in Shapira's work, as in the study by Hanania and Gradman (1977), was on morpheme development. If one takes a view of language in which formulas are built up into low scope patterns, with abstract systems emerging much later, then the 80-85% threshold for

[5]Krashen and Scarcella (1978) draw on Peters' (1985) distinction between language that is analysed and that which is gestalt, the latter referring to unanalysed formulas.

[6]Plural *-s*, third person *-s*, plural *be*, copula *be+*, *V+be, -ing, be+NEG, do+NEG*, yes/no question, *wh-* question and possessive *-s*.

acquisition in *all* obligatory contexts could well cloud the development of local patterns, which are not as yet system wide. Notwithstanding this caveat, the data in Shapira's study, as in that of Hanania and Gradman, suggest that there is both little grammatical development and therefore little use of formulas in grammatical development.

Particularly relevant data for any transition from formula to low-scope patterns is in a study by Schmidt (1983) of a 33-year old Japanese speaker called 'Wes'. Schmidt tracked Wes' communicative competence over a 3-year period during his transition to permanent residence in Hawaii. Wes had had no instruction in English prior to his move, and his initial ability in English was considered minimal. In terms of his grammatical competence, operationalised here as the suppliance of morphemes in 90% of obligatory contexts, Wes made little progress over the three years on his production of articles, possessives and regular past tenses (See Table 2.10).

**Table 2.10:** *Wes' accuracy order for nine grammatical morphemes in obligatory contexts (Schmidt 1983)*

| | July 1978 | November 1980 |
|---|---|---|
| Copula BE | Acquired, present only | No change |
| Progressive ING | Acquired (?[a]) | No change |
| Auxilliary BE | Acquired (?) | No change |
| Past irregular | 25% | 55% |
| Plural | 5% | 43%/33%[b] |
| 3rd singular | 0% | 21% |
| Article | 0% | 19%/6% |
| Possessive | 0% | 8% |
| Past regular | 0% | 0% |

a) the question mark denotes uncertainty due to oversuppliance of the form in non-obligatory contexts
b) the lower figure is due to the removal of instances that were deemed formulaic

Even when it appeared that acquisition had taken place (with Copula *be*, Progressive *-ing* and *be* auxilliary), the figures in Table 2.10 refer to suppliance in an obligatory context *and* suppliance in non-obligatory contexts. From Schmidt's data, Wes was producing *-ing* and auxilliary *be* in a large number of

non-standard utterances mapping on to a large number of non-standard meanings. Futhermore, although the plural and article system seemed to greatly improve, a significant number of these are part of high-frequency phrases, e.g. *n years ago*, *a little (bit) X*. This lack of development was also the case in other areas of grammar:

> Except in routine, formulaic utterances (do you have time? are you busy?), Wes has no subject-verb inversion in questions (ah, you has keys? when Tim is coming?), no dummy or pseudo-subjects (because here is nothing for "because there is nothing here"), no relative clauses (you know before people bought my painting for "people who bought my paintings before"; he saw before I sent painting for "he saw the painting I sent before"), and no passives (this is cannot help for "this cannot be helped"). (p.149)

In contrast to Wes' lack of grammatical development, he showed a good knowledge of formulaic phrases (such as those listed in Table 2.11), and had gained 'considerable control' (p.154) over their sociolinguistic use.

**Table 2.11:** *Examples of formulaic phrases produced by Wes (Schmidt 1983, p.150)*

| | | |
|---|---|---|
| Hi! How's it? | So, what's new? | I beg your pardon |
| What did you say your name was? | You know what ... | I tell you. |
| You know what I mean? | What can I do? | Whaddya want? |
| Whaddya mean? | Whaddya doing? | Do you remember? |
| I told you before. | Something like that. | I dunno why. |
| That's great, terrific! | Sure, I would love to. | Can you imagine? |
| Thank you VERY much. | Thank you calling. | |

Wes also showed some ability in 'freeing up' elements of the formulas for use in other utterances. Early in the study, Wes would produce the formula *shall we go?*, but not extend it to contexts where it would be appropriate, for example he uses *sitting?* when *shall we sit down?* would be appropriate. Later on in the study, however, the formula could be used productively in different contexts, e.g. *Shall we maybe go out coffee now, or you want later?* Wes also

seemed to have broken down formulas and created non-standard utterances. For example, he seemed to use the phrase *Can I X?* to produce 31-34 below.

> 31. can I getting some more coffee? (at home, intended as request)
> 32. can I have a light? ('turn on the light,' not 'give me a match')
> 33. if you back to room, can I bring cigarette? ('please bring me')
> 34. uh, can I? ('would you?')
>
> (Schmidt, 1983, p.183)

Overall, while Wes' progress in terms of sociolinguistic, discoursal and formulaic knowledge was extensive, Schmidt admits that he finds little evidence within the data for the 'evolution toward creative grammar from formulaics' (p.150). This is interesting because while others have pointed to social distancing as a cause of the (non)learning of a second language, e.g. Schumann (1976) and Shapira (1978), Wes was described as an 'extremely extroverted and outgoing person' (p.142), who was actively engaged in the local culture with a 'strong drive to communicate' (p. 143). It is not clear why then, with a stock of formulaic phrases, a need and want to communicate in the target culture, there was limited development in Wes' grammatical ability in terms of morpheme development or in his ability to readily extend his formulaic phrases to new contexts.

Yorio (1989) also provides data which questions the trajectory from formula to low-scope pattern to construction. The data comes from two sources: a case study of the written production of an 18-year old Korean learner called 'K', who had lived in the USA for 5 years and had little formal training in English, and the second from 25 ESL students who had been in the US for between 5 and 7 years. Both K and the ESL learners produced a large number of conventionalised expressions in their productions[7], although as can be seen from Table 2.12 many of the expressions they used were non-standard.

[7]Yorio (1989) uses the superordinate 'conventionalised expressions' to descibe 'idioms, formulas, collocations, etc.' (p.55)

**Table 2.12:** *Standard and non-standard forms used by 'K' and advanced ESL learners (Yorio 1989)*

| 'K' | Advanced Learners |
|---|---|
| I know what's going on | In conclusion |
| No matter what happens | According to |
| in a row | By way of illustration |
| right now | In summary |
| at this point | For example |
| Most of the time | *take advantages of [a] |
| *Sunday was lousiest day | *he had chance |
| *just two of us | *are to blamed for |
| *for short time | *at the present |
| *I have lots money | *those mention above |
| *at the morning | *being taking care of |
| | *A friend of her |

a) *Denotes an erroneous form

If one assumes that the errors were not evident in the linguistic environment, the origin of such non-standard formulas such as **for short time* is unclear. Yorio suggests the existance of such non-standard forms is evidence for two systems: a collection of unanalysed formulas and an apparently separate rule system which generates propositional speech. As K made the same errors in non-conventionalised propositional speech as he did in conventionalised expressions, Yorio argues that the separate rule system affects the grammaticality of the formulas.

> [conventiaonalised expressions] are not simply memorized or taken in as wholes, but that, in fact, they are subject to whatever interlanguage rules the learner is operating under' (Yorio, 1989, p.60).

In addition to arguing for the prevalence of interlanguage rules, Yorio finds no evidence that while this proposed rule system may affect the grammaticality of formulas, learners are able to use these formulas for grammatical development.

**2.1.2.1.4 Discussion** Although the data from these studies offer some support for the learning of formulas early on (Wes and Fatmah) and for formu-

las as a entry into creative construction, most notably Ziola's breakdown of *I don't know* and Wes' analysis of *Shall we*, overall, the grammatical development of each participant in these studies was minimal, with little improvement in morpheme development over considerable periods of time. As mentioned throughout this section, however, it could well be the case that defining acquisition as the suppliance of morphemes in obligatory contexts is incompatible with the view of development which emphasises the piecemeal acquisition of constructions. If a typical trajectory is from formula to low-scope pattern, then measuring the most systemic of rules, e.g. plural *-s*, may not be the best indicator of the development of local rules, which predict suppliance in some cotexts and not others.

Arguably, Schmidt's study of Wes, while chiefly focusing on systemic morpheme development, did track the emergence and contextual appropriacy of formulas, but even then, grammatical development was very limited, and the development of formulas hit and miss. Weinert (1995) argues that studies like these are more related to 'minimal or fossilized development' (p.190) than to the role of formulas in development, although even with this in mind any account of a trajectory from formula to low-scope pattern would still need to address why this trajectory seemed so limited.

Shapira offers a tentative explanation why this might be the case with Zoila. Zoila used a strategy which focused on the communicative words in the linguistic environment, and ignored words that added little to the message.

> Zoila: The sentence, sentence. In the sentence I never using this little, little words.
> Rina: Like what?
> Zoila: Ah, 'and,' and 'that,'[. . . ][=examples of 'little' words] as observed by Zoila)[. . . ] Well, maybe because I no study . . . never, and only hear the people and . . . and talking.
> Rina: Yeah, but people talk with these words.
> Zoila: Yeah, *pero*, I'm . . . hear and put more attention the big words. You know and . . . something 'house.' I know 'house' is the *casa* for me. And [. . . ] and little words is no too important

for me.

Abridged from Shapira 1978 (p.254, emphasis added)

Although Shapira explains this strategy within Schumann's theory that social and psychological distance was a major causal factor in SLA (e.g. Schumann, 1976), Schmidt's analysis of Wes' production clearly contradicts this. Wes was outgoing, motivated and 'acculturated', but still showed limited evidence of grammatical development. Alternatively, Zoila's focus on single words in the linguistic environment offers support for Wray's theory on the differences between analysis in FLA and SLA (Wray, 2002; Wray and Perkins, 2000). According to Wray, in FLA, children first go through a holistic learning phase (from birth to around 20 months), which involves the identification and storage of linguistic units in the environment that are deemed socially important. The second phase is analytical (from approximately 2 to 8 years), within which formulaic sequences are segmented and analysed with the aid of literacy into constituents.

Wray argues that due to their experience with analysis and literacy in the L1, the units in focus in an L2 are 'fundamentally different' (Wray, 2002, p.212) to the L1. In learning the L1, sequences are first stored and then analysed only if necessary. In contrast, for adults learning an L2, due to their experience with the L1, the focus is on smaller word-like units, not on formulaic sequences. As Wray puts it:

> Where the first language starts with large and complex strings, and never breaks them down any more than necessary, the post-childhood second language learner is starting with smaller units and trying to build them up. (p.206)

This difference can be seen in Figures 2.1 (FLA) and 2.2 (Adult SLA). In Figure 2.1, the emphasis is very much on the storage of multiword utterances (top left of the figure), which are consequently analysed into component parts (from multiword to formulaic word). This analysis is aided by the teaching of

literacy (bottom left), which encourages the analysis of multiword units into individual words. In adult SLA, multiword input (bottom left) is analysed into formulaic words, as are words in the input (bottom middle), with only a minimal number of formulaic word strings stored as such. Although Wray's emphasis is on taught input, it is possible in this model that adults also prefer analysis over storage of the formulaic sequences encountered in naturalistic input (top left of Figure 2.2).

Relating this to the data in Yorio (1989) and Hanania and Gradman (1977), if the emphasis is shifted in SLA in adults from formulaic sequences to word-like units, then formulaic sequences encountered in the environment are more likely to be analysed for the individual words and stored as such, without the necessary evidence for grammatical development. Consequently, the sequences, while adding to the learner's lexical repertoire, do not directly contribute to the learning of more abstract grammar. In a learner's production, then, 'whatever interlanguage rules the learning is operating under' (Yorio, 1989, p.62) are used to glue together lexical items, therefore allowing errors into formulaic sequences. If this is indeed the case, then the extraction of word-like units may go some way to explain productions such as *Mommy coming*, *Go wash the hand* and *The dog eating* in Hanania and Gradman (1977).

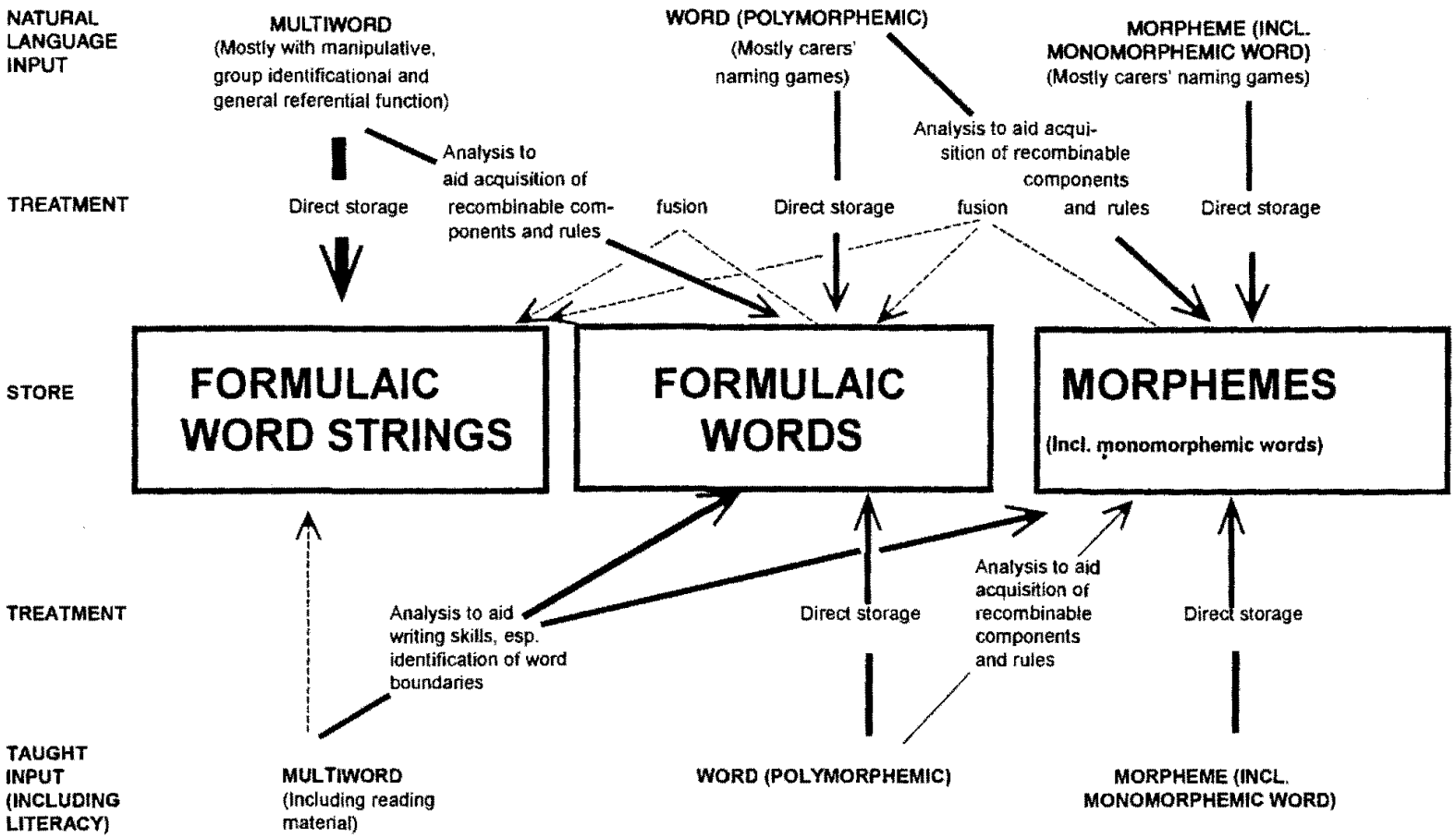



**Figure 2.1:** *Formulas and the creation of the lexicon in FLA (Wray 2002, p207).*

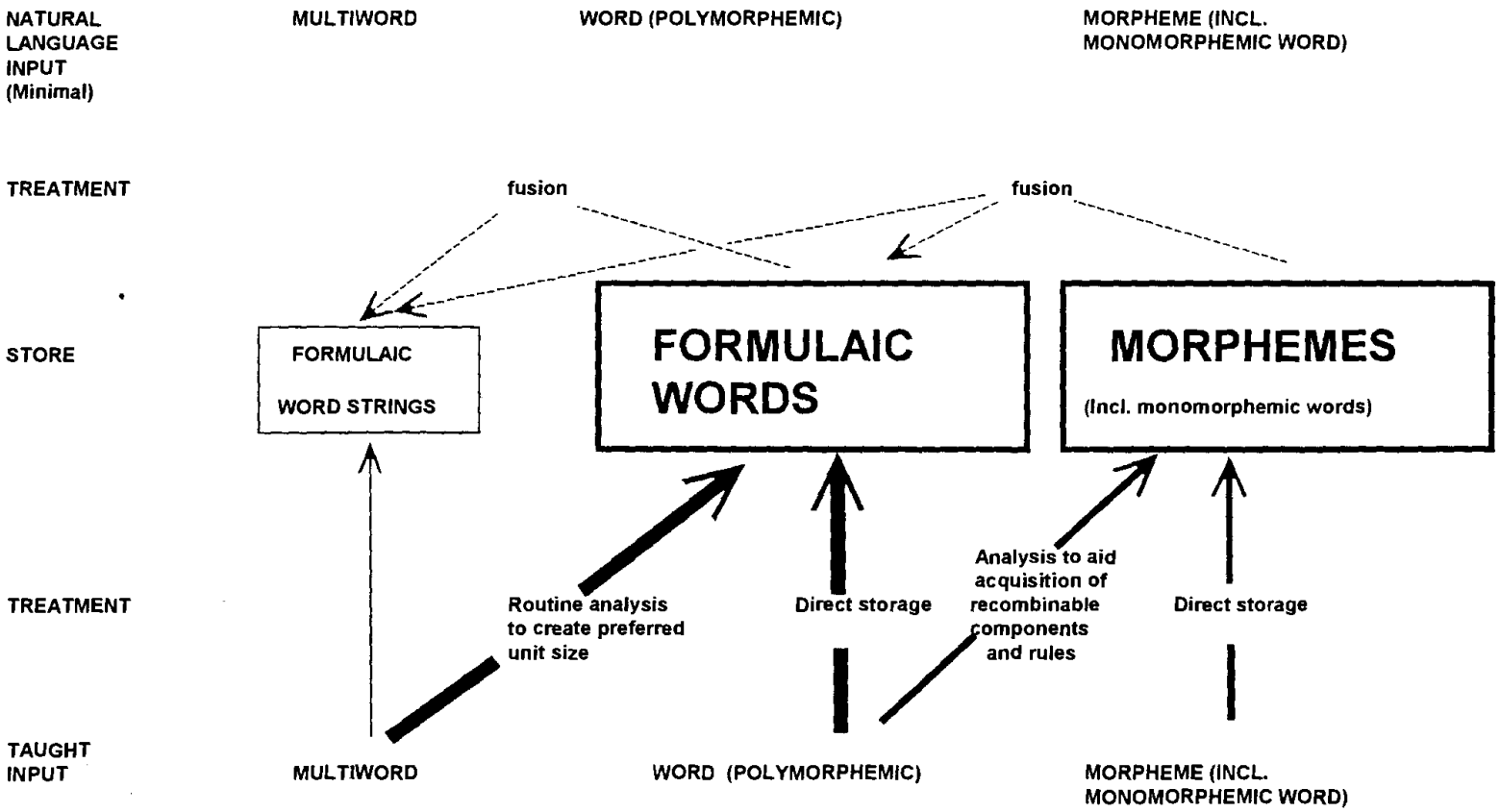



**Figure 2.2:** *Formulas and the creation of the lexicon in adult SLA (Wray 2002, p208).*

Wray's perspective can also be related to construction grammars, specifically Goldberg's (1995) analysis of verb argument constructions. If for example, tokens of the caused motion construction such as *He pushed it over* are not retained as formulas, but are analysed as individual lexical items, then it is likely that the constructional meaning 'X CAUSES Y TO MOVE $Z_{PATH/LOC}$' is lost, leaving an expression as the sum of its individual parts.

Overall, if Wray's hypothesis is accurate, then N.Ellis' trajectory from 'formula through low-scope pattern to construction' would not be the primary method for extracting grammatical information from the linguistic environment in naturalistic adult SLA, as the initial stage of formula extraction in adults could well be considerably different from that of FLA and child SLA.

#### 2.1.2.2 Evidence for a trajectory in instructed SLA

While the longitudinal adult SLA studies by Hanania and Gradman (1977); Schmidt (1983) and Shapira (1978) offered detailed production data, the emphasis was mainly on the acquisition of grammatical morphemes in naturalistic settings. The classroom-based studies detailed in this section focus specifically on the breakdown of formulas.

**2.1.2.2.1 R.Ellis (1984)** R.Ellis (1984) tracked the use of the formula *I don't know* of two Pakistani children (11 and 13) and an 11-year old Portuguese boy, studying beginner English in London. Drawing on trajectories suggested by Clark (1974) and Wong Fillmore (1976, 1979), the focus of the study was whether, or the degree to which, routines might be broken down into constituent parts over time, leading to rule-like behaviour. This breakdown was operationalised by tracking the development route of the constituents *I*, *don't* and *know* of the routine *I don't know* over a one-year period. Specifically, he traced:

1. when 'don't' was first used in similar but different expressions;
2. when an alternative subject to 'I' first occurred;

3. when 'know' was released for use without 'don't';
4. when an additional constituent first occurred.

(R. Ellis, 1984, p.60)

Findings from the study showed that when *don't* was first found in other expressions, its use was limited to specific verbs (e.g. *like* and *understand*), which could also be considered as unrelated formulas, not a breakdown of *I don't know*. Only one of the learners in the first year managed to replace the pronoun *I* with an alternative, and *know* was only used outside *I don't know* in the phrase *I know X*. Although the first three of R.Ellis' developmental indicators showed few signs of development, there was more evidence for the fourth - when an additional constituent first occurred. R.Ellis found that there was considerable use of the juxtapositioning of other phrases with *I don't know*, shown in Examples 35-37.

35. That one I don't know
36. I don't know what's this
37. I don't know 'holiday' spelling

(R. Ellis, 1984, p.60)

He also found the routine *I don't know* was blended with uptake from the teacher's previous utterance, as in Examples 38 and 39.

38. I don't know what is squirrel
39. I don't know making

(R. Ellis, 1984, p.60)

As R.Ellis acknowledges, while there is some evidence of active engagement with the formula, the data on whether a routine like this can form the basis for creative construction is far from conclusive. What can be gleaned from this study is that if the analysis of formulas does take place, which may then lead to rule-like behaviour, the trajectory in second language learners is a slow process. R.Ellis' data spans a year of immersion into a London-based school, and showed little grammatical development from initial formulas.

**2.1.2.2.2 Myles et al. (1998, 1999)** A more comprehensive picture emerges from two longitudinal studies on the breakdown of the formulaic phrases of secondary school pupils learning French (N=16) in the UK. In a three-year study, both Myles et al. (1998) and Myles et al. (1999) tracked the breakdown of formulas and their 'contribution to the development of a creative language capacity' (p.323). In the first study (1998) the chunks *j'aime* (I like), *j'adore* (I love) and *j'habite* (I live) were initially identified, based on their over-extended syntactic, pragmatic or semantic use (Examples 40 and 41).

40. **la Monique j'ai...j'ai adore...la tennis*
(Monique I have I have love tennis = Monique loves tennis)

41. **une famille...j'habite un maison*
(a family I live in a house = the family lives in a house)

Myles et al. (1998)

The pupils took part in six oral production tasks (termed 'rounds') with a proficient speaker at the end of each school term (years 7-9). Drawing on R.Ellis (1984), the three chunks were then tracked to determine:

(when) were *'j'/je'* used elsewhere?
(when) did alternative subjects to *'je'* occur?
(when) were *'aime/adore/habite'* released for use without *'je'*?

(Myles et al., 1998)

Similar to R.Ellis (1984), they found that for the majority of pupils *J'* was only used in the three chunks or in the chunk *J'ai* (I have), and for 14 of the 16 pupils it was not used creatively at all with other verbs. Only two pupils seemed to have segmented the subject and verb and therefore 'freed up' *J'* for use elsewhere. The uncontracted subject *Je* was used in a similar uncreative way for eight of the sixteen students. For the remaining eight students, *Je* was starting to be used with other verbs in a creative fashion.

In terms of alternative subjects, nine of the students only used third person subjects *elle* or *il* with the verbs *a*, *est* and *s'appelle (has, is, is called)* or juxtaposed *elle* or *il* with one of the chunks *j'aime j'adore* and *j'habite* (e.g. *elle j'aime (She I like))*. The other seven pupils showed greater creativity in their third person pronouns and verbs, suggesting that they had broken down the chunks and were starting to use them productively. Interestingly, this creativity and lack of creativity respectively seemed to correlate with the initial use of chunks, the more creative group initially having a greater knowledge of the chunks *j'aime j'adore* and *j'habite*, while the group that showed little progress showed less knowledge of chunks.

In relation to the verbs *aime*, *adore* and *habite* without *J'*, a similar pattern to the use of subject pronouns emerged: approximately two thirds of the pupils used the three verbs predominantly in the chunks *j'aime j'adore* and *j'habite* (*j'aime* showed greater signs of creative construction). The remaining third of the pupils showed significant signs of breaking the chunks down and using the subject and the verb independently in more creative construction. Again, these students were ones who had shown particular expertise in the use of chunks earlier on. Myles et al. (1998) conclude that the utility of formulaic utterances goes beyond solely enabling incipient communication, as the breaking down of the formulas seems to correlate with, in this case, the emergence of subject pronouns, and therefore creative construction of novel utterances (albeit with large between subjects variability).

The second study by Myles et al. (1999) used the same rounds, elicitation procedures, participants (N=16) and criteria for chunk selection, but the focus this time was the interrogative chunk *comment t'appelles-tu?*. The study investigated to what extent the chunk was broken down over time, whether the creative process and the use of chunks were independent processes, and what the relationship was if any between the use of chunks and creative construction.

To gauge the creative construction of interrogatives, Myles et al. (1999) pushed pupils into contexts in which they could not use their repertoire of

memorised interrogative chunks (e.g. Pushing learners into speaking about the name of a third person when the learner only has control of *comment t'appelles-tu?*). The productions from the oral tasks are detailed in Table 2.13. The first row shows the use of the formula *comment t'appelles-tu?* across the six rounds, with 18 occurrences in the first round of tasks. The second row details an increasing trend across the rounds in attempts by the learners to establish third person reference by juxtaposing the unanalysed chunk *comment t'appelles-tu?* with a NP, e.g. **Comment t'appelles-tu un garçon?*, a similar strategy to the juxtaposition of NPs with *I don't know* in R.Ellis (1984). The third and fourth row - *Comment t'appelle* (+NP) and *Comment s'appelle* (+NP) - seem to show some analysis of the internal structure of the chunk, as *tu* has been replaced with an NP referring to a third person. Only two participants in the study seemed to have gone further in the analysis and produced the token *(NP) il/elle s'appelle comment*, again a form emerging later in the rounds.

**Table 2.13:** *Frequencies of the comment t'appelles-tu? chunk with other uses of s'appeler, by round (Myles et al. (1999)*

| | Round | | | | | | |
|---|---|---|---|---|---|---|---|
| Chunk Type | 1 | 2 | 3 | 4 | 5 | 6 | Total |
| ***Comment t'appelles-tu?/ comment tu t'appelles?* (second person)** | 18 | 2 | 0 | 0 | 1 | 2 | 23 |
| *Comment t'appelles-tu?/comment tu t'appelles?* (+NP) (third person) | 0 | 9 | 1 | 7 | 14 | 1 | 32 |
| *Comment t'appelle?* (+NP) (third person) | 0 | 6 | 2 | 9 | 1 | 0 | 18 |
| *Comment s'appelle?* (+NP) (third person) | 2 | 7 | 2 | 12 | 13 | 0 | 36 |
| *il/elle s'appelle* (+NP) (third person) | 1 | 7 | 0 | 5 | 3 | 0 | 16 |
| (NP) *il/elle s'appelle comment* (third person) | 4 | 0 | 0 | 11 | 4 | 0 | 19 |
| Other *s'appeler* chunks | 2 | 1 | 3 | 4 | 7 | 0 | 17 |
| Total | 27 | 32 | 8 | 48 | 43 | 3 | 161 |

Although one could argue that the chunks were discarded as creative construction kicked in and that the use of chunks was solely a communicative crutch and not integral to creative construction (cf. Krashen and Scarcella,

1978), the study also traces the link between formulaic chunks and creative construction. By first eliminating from the data all unanalysed chunks, Myles et al. could trace the development of creative construction. They found that pupils fell back on either verbless utterance (e.g. *Où la plage?* (where the beach?)) or, at more advanced stages utterances with verb phrases (e.g. *je manger douze heures? (I eat-INF 12 o'clock?)*, and finally a finite verb stage (although the majority of learners did not reach this stage). Pupils who did not have a initial repertoire of chunks did not move beyond the v-less stage, suggesting that initial knowledge of chunks had an effect on further analysis. At the other end of the scale, learners who were able to memorise chunks early on, rather than abandon them in favour of an emerging creative system, continued to explore how they could be broken down and moved beyond the v-less stage. For example, one pupil who had memorised the *comment t'appelles-tu?* chunk went on to create **comment il t'appelles-tu?* and **il est comment t'appelles-tu?* in an attempt to fit the chunk to third person reference. In line with Myles et al. (1998), Myles et al. (1999) conclude that 'far from inhibiting progress in our learners, ready-made chunks or formulas provided by the classroom context were instrumental in the learning process' (p.76).

**2.1.2.2.3 Eskildsen (2012)** Further classroom-based studies that emphasise the importance of multiword phrases in grammatical development are those of Eskildsen (Eskildsen and Cadierno, 2007; Eskildsen, 2009, 2012). Using data from the Multimedia Adult English Language Corpus (MAELC), Eskildsen (2012) tracked the development of SUBJ AUX-dopresent/past tense NEG VERB constructions such as *I don't know* between 2001 and 2005 of an adult student called Carlos studying at an adult community college in the US. Figure 2.3 shows the Recording Periods (RPs 1-5) spread over four years. In the first year, Carlos' productions mainly consisted of the nontargetlike form *subject no verb*, such as *I no remember*, and Multi-Word Expressions (MWE) such as *I don't think so* and *I don't know*, with comparatively few targetlike utterances.

Over time, the targetlike SUBJ AUX-dopresent/past tense NEG VERB pattern increased, while the nontargetlike patterns declined to a point at which in the final RP there were no instances in the sample.

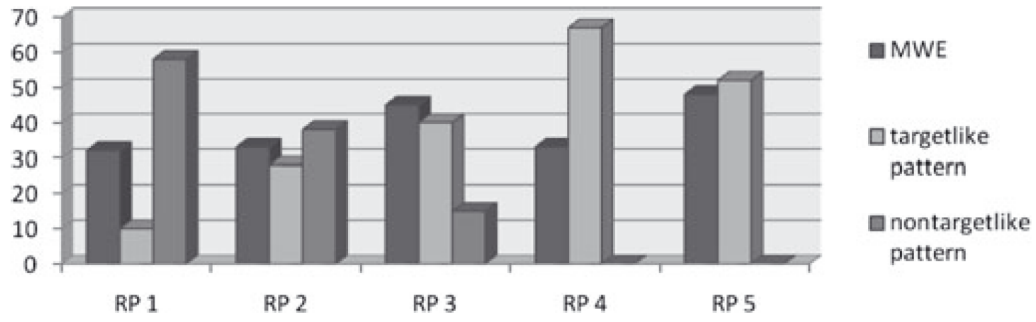



**Figure 2.3:** *Distribution of subject don't/no verb patterns over time (Eskildsen, 2012, p.342)*

**Table 2.14:** *Type-token rations for Carlos' SUBJ AUX-dopresent/past tense NEG VERB construction*

| | **Targetlike, subject don't verb** | | | | **Nontargetlike, subject no verb** | | |
|---|---|---|---|---|---|---|---|
| | Tokens | Types | Ratio | MWE[1] | Tokens | Types | Ratio |
| RP1[2] | 13 | 3 | 0.23 | 10 (.77) | 18 | 13 | .72 |
| RP2 | 24 | 7 | .29 | 13 (.56) | 15 | 11 | .73 |
| RP3 | 40 | 13 | .33 | 21 (.53) | 7 | 6 | .86 |
| RP4 | 18 | 9 | .50 | 6 (.33) | 0 | 0 | .00 |
| RP5 | 23 | 11 | .45 | 11 (.48) | 0 | 0 | .00 |

1. =Recording Period
2. The Multi-Word Expression *I don't know*, in tokens (and ratios)

Eskildsen (2012) adopts a usage-based model to explain the changes in targetlike and nontargetlike productions. In such approaches, token frequency is the number of instances of a pattern, while type frequency is the number of different instances of the same pattern, e.g. several tokens of *I don't know* fall under the targetlike pattern SUBJ AUX-dopresent/past tense NEG VERB, which results in a high token frequency, but as the tokens are identical there is

a low type frequency (if there were ten instances of *I don't know*, the type-token ration is .1). Conversely, several tokens such as *I don't know, She doesn't think, He doesn't want, I don't see, They don't go* results in a high type and token frequency with a ration of 1 (e.g. ten tokens and ten types). From a usage-based perspective, a high token frequency and low type frequency suggests a entrenched MWE, while a high type frequency suggests a schematised pattern than can produce a number of different instantiations.

As illustrated in table 2.14, in the initial stages the multiword expression *I don't know* accounted for over 75% of the total tokens of the construction with a low type-token ratio (0.23), which suggests an entrenched formula of limited productivity. At a later recording period, Carlos' targetlike production sees a increased variety of pronouns in the subject position and the conjugation of the verb, leading to a concomitant higher type-token ratio (.50 in RP4 and .45 in RP5), suggesting from a usage-based perspective that the pattern is schematic, and therefore productive. For the nontargetlike use of the *subject no verb* pattern, the initial type-token ratio is high (.72), suggesting a highly productive pattern, but diminishes over the four-year period and finally disappears, apparently giving way to the increasingly productive targetlike pattern. Eskildsen does not offer an explanation for the productive use of the *subject no verb* pattern in the initial stages, but it seems likely that transfer from the Spanish construction No VERB, such as *no voy (I don't go)*, *no me recuerdo (I don't remember)* or *No sé (I don't know)* may have been an influence.

Based on this analysis, Eskildsen (2012) argues that multi-word expressions such as *I don't know* are essential for the initial establishment of a type. Subsequent experience with similar tokens of the same constructional pattern allows then for the breakdown of the initial rigid multiword expression into a low-scope pattern, and finally into a SUBJ AUX-dopresent/past tense NEG VERB construction. *Section 2.2: Processes in a Usage-Based Trajectory* discusses in more detail the initial establishment of types and the subsequent development of schematic types within a usage-based framework.

### 2.1.3 Summary

This section set out to review the claim that there exists a trajectory in SLA from formula to low-scope pattern to construction (N. Ellis, 2002). In child SLA, there seems to be evidence that this is the case, at least with some linguistic structures. The in-depth study by Wong Fillmore (1976), for example, tracks the breakdown of *How do you do dese?* from an initially unanalysed formula through segmentation of its constituents to the freeing up of constituents to enable their more creative use. The extant literature on adult SLA paints a much bleaker picture, with subjects showing little grammatical progress over relatively long periods of time. As Weinert (1995) argues, however, this appears to show fossilised development, and does not necessarily show that formulas do not lead to more abstract constructions. A further point is that the emphasis of most of these studies was morpheme development such as *-ing*, not the development based on the breakdown of formulas. As morphemes such as *-ing* cut across many constructions, analyses of use in obligatory contexts may well have missed the breakdown of chunks into low-scope patterns.

A second possibility for such lack of development suggested by Wray (2002) is that adults have relatively fewer formulas at their disposal than children due to a more analytical approach to learning. This approach would theoretically increase the adults' repertoire of segmented individual items, while disabling their ability to retain whole sequences. This lack of sequences would prevent them from drawing on this type of data in the development of the grammar, alternatively relying on the L1 or consciously-learnt rules.

Partial support for this suggestion comes from Myles et al. (1998) and Myles et al. (1999), who show that learners who were 'force-fed a diet of [. . . ] chunks' (Myles et al., 1999, p.61) in the initial stages of learning (and also remembered those chunks) went on to outperform learners who did not initially learn the chunks. Myles et al. (1999) interpret these findings as support for a trajectory from formulaic chunks to v-less constructions to v-headed constructions.

This trajectory, they argue, is driven at least in part by the need to express functions that are beyond the scope of their repertoire of formulaic sequences (also Perdue, 1993; Wray, 2002). A trajectory from entrenched multi-word expressions to productive patterns is also supported by the type-token analysis of Eskildsen, who interprets such findings from a usage-based perspective in which initial tokens are entrenched in the mind of the learner and subsequently broken down as leaners 'encounter and learn a sufficient number of similar but different instantiations of a given pattern to be able to analogically form abstractions of what is generic to the pattern' (Eskildsen, 2012, p.338). The next section develops the ideas that learners can 'analogically form abstractions from similar but different tokens' and that a catalyst for the breakdown of tokens is the need to express functions beyond the ability of the learner.

## 2.2 Processes in a usage-based trajectory

While *Section 2.1.2: Evidence for a trajectory* discussed evidence for a trajectory from formula to low-scope pattern to construction in child and adult SLA, this section focuses on the motivation for moving along the trajectory, and the possible mechanisms involved. In terms of motivation, drawing on the work by Myles et al. (1998, 1999) and studies in Task-Based Learning (TBL) that use *task demands* as a variable (Robinson, 2007; Samuda, 2001; Schmidt, 1990)[8], *Section 2.2.1: Motivation for the breakdown of formulas* discusses the view that the breakdown of formulas is at least in part driven by the need to express functions which are beyond the capability of existing formulas. The section also reviews evidence that the motivation for the breakdown of formulas comes from the frequency distribution of forms in the linguistic environment (e.g. Casenhiser and Goldberg, 2005; Ellis and Ferreira-Junior, 2009a).

[8]While Schmidt's central concern is the *noticing* of forms in the linguistic environment not the empirical investigation of task demands, he cites task demands as 'a powerful determinant of what is noticed' (p.143).

In terms of the mechanisms involved in the development of patterns from formulas, *Section 2.2.2: Processes acting on the input: categorisation, analogy and similarity* will explore in more detail usage-based accounts of language acquisition. Within this broad framework, the main emphasis is on how tokens in the input can be categorised by processes of *analogy* and *similarity* (e.g. Gentner, 1983; Gentner and Markman, 1997; Gentner and Medina, 1997), and how the relative importance of these processes may vary depending on the stage of development between formula and construction.

## 2.2.1 Motivation for the breakdown of formulas

### 2.2.1.1 Motivation from the linguistic environment

*Section 2.1: Usage and a trajectory from formula to construction* reviewed evidence that suggests the target language can be described at points along a continuum of abstractness between inflexible idiomatic formulas and abstract highly-productive constructions. The question Wray (2002, 2007) asks is, given that some formulaic utterances are semantically or grammatically irregular (e.g. *by and large, kick the bucket*), and some are regular (e.g. the caused-motion construction), why would one analyse a priori only the regular ones, which would lead to productivity, but retain the idiomaticity of the irregular sequences? In an attempt to account for this, Wray offers what she terms *needs-only analysis*.

Needs-only analysis is a conservative learning strategy which predicts to what extent an unanalysed sequence will be analysed based on the occurrence of certain forms in the input. The default position for such a strategy, according to Wray, is not to analyse a reliable form-meaning mapping unless motivated by the linguistic environment. Sequences such as *by and large* can be reliably mapped on to a meaning, while the linguistic environment will not provide further examples such as *By and X* or *X and large*. The formula therefore remains intact and does not undergo further analysis. In contrast, when the linguistic environment provides evidence, formulas are analysed. To exemplify,

Wray (2007) uses the example *from now on.* Initially, the phrase is unanalysed and assigned a specific meaning. On experiencing related forms in the input e.g. *from then on, from Tuesday on, from that moment on* (Wray, 2007, p.57), the fixedness of the middle position will be 'loosened'. The first and third words, however, will not see motivation from the environment to loosen, and therefore remain fixed. This theory sees some support from Culicover (1999, Culicover and Nowak, 2003; Jackendoff, 2002), who describes 'the conservative attentive' (p28) learner as one who 'acquires [...] regularity only for the cases to which it applies in the learner's experience'. From this perspective, the default processing is not the analysis of tokens, but the storage of formulas. Further analysis requires motivation, either in terms of the distribution in the linguistic environment or the communicative needs of the learner.

Evidence that the frequency distribution of tokens in the linguistic environment plays an influential role in the breakdown of tokens comes from a series of empirical studies which focus on the frequency-based acquisition of verb-argument constructions (Ellis, 2010; Ellis and Ferreira-Junior, 2009a,b). Ellis and Ferreira-Junior (2009b) used longitudinal data from the European Science Foundation (ESF) project (Perdue, 1993) to investigate the relationship between the distribution of Verb-Argument Constructions (VACs) such as *He pushed it down the street* [Subj V Obj $\text{Obl}_{pathlloc}$] in native and non-native speakers learning English as a second language[9], and the association between this frequency distribution and the initial acquisition of the constructions. The VACs were separated into the following types:

Verb Object Locative constructions (VOL)
Meaning: X CAUSES Y TO MOVE $\text{Z}_{PATH/LOC}$
Form: Subj V Obj $\text{Obl}_{path/loc}$

Verb-Locative constructions (VL)
Meaning: X MOVES $\text{Z}_{PATH/LOC}$
Form: Subj V $\text{Obl}_{path/loc}$

[9]The corpus was established from NS/NNS dyads, with a total of 14,574 verb tokens.

Ditransitive constructions (VOO)
Meaning: X CAUSES Y TO RECEIVE Z
Form: Subj V $Obj_1$ $Ob_2$

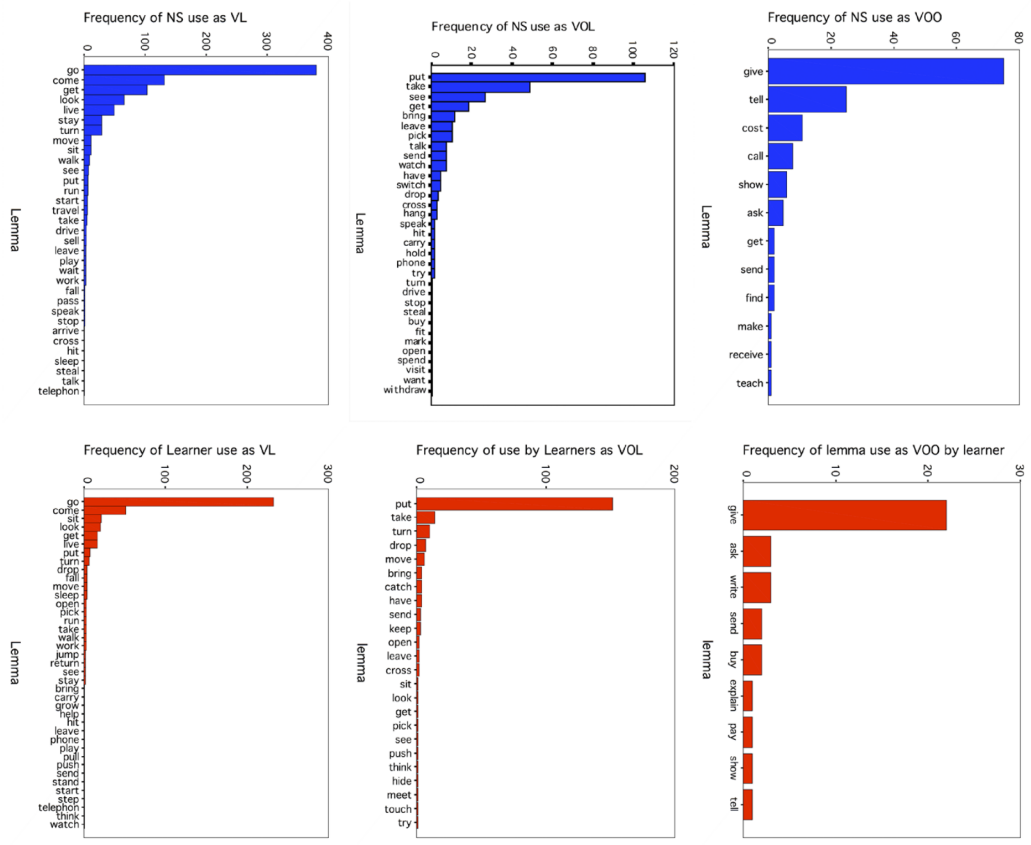



**Figure 2.4:** *Native and non-native use of lexical verbs in Verb-Locative (VL), Verb-Object Locative (VOL) and Verb-Object-Object (VOO) constructions (Ellis and Collins 2009*b*).*

As can be seen from the upper row of graphs in Figure 2.4, Ellis and Ferreira-Junior (2009b) found that, for the native speakers, one verb accounted for the lion's share of tokens in each construction (*go* for the VL construction, *put* for the VOL construction and *give* for the VOO construction). The second most frequent verb (*come* for the VL construction, *take* for the VOL construction and *tell* for the VOO construction) was considerably less frequent than the most frequent verb, and the third verb considerably less frequent than the second

verb, and so on. They argue that the pattern between the frequency of the verbs is not linear but follows a Zipfian profile. A Zipfian profile is derived from Zipf's law (Zipf, 1949), which states that the occurrence of a word (for example *give* in the VOO construction) will be related to its rank in a frequency table by a power function. Ellis and Ferreira-Junior (2009b) found that the non-native use of verbs in the three VACs followed a similar profile in their frequency distribution, e.g. *put* as the most frequent and *take* as the second most in the VOL construction (See lower row of graphs in Figure 2.4). They also found that the verbs that were acquired first correlated with the most frequent verbs in the constructions, and that these verbs were also prototypical in meaning. In a later study, Ellis and Ferreira-Junior (2009a) found that similar Zipfian distributions were evident in other 'islands' in the construction, e.g. in subject, object and preposition positions. Ellis and Ferreira-Junior (2009b) conclude by suggesting that:

> ... learners acquire the most frequent, prototypical, and generic exemplars. Learning is driven by the frequency and frequency distribution of exemplars within constructions and by the match of their meaning to the construction prototype. (p.384)

That the frequency distribution of elements within tokens (or exemplars in Ellis and Ferreira Junior's terms) plays a key role is similar to the stance taken by Wray in her 'Needs-only analysis' (Wray, 2007). In a needs-only analysis, while there is no motivation in the linguistic environment to breakdown *by and large* into smaller constituents, there *is* motivation when alternative tokens exist in the environment. In the case of VACs above, the motivation to break down the initially prototypical constructions (e.g. *go* for the VL construction, *put* for the VOL construction and *give* for the VOO construction) appeared to come from the distribution of linguistic forms in the linguistic environment, i.e. VAC constructions with alternative verbs to the entrenched prototypical construction. In Wray's terms this distribution 'loosened' the VP slot within the formulaic and prototypical tokens.

#### 2.2.1.2 Motivation from communicative need

In addition to the properties of the linguistic environment, there is also the probability that *communication need* is a causative factor in the breakdown of formulas. Myles et al. (1998, Myles, 2004) suggest that the breakdown of formulas that occurred in their 1998 study was due to a combination of the exigencies of the tasks and the inadequacy of the pupils' current interlanguage to deal with these exigencies. This tension between task demands and current ability can be seen in the following example, taken from round two of the study:

> [. . . ]oh she déteste it (. . . ) et la Monique . . . umm . . . j'ai adore la géographie . . . la Monique . . . non she adore j'ai adore le Monique does . . . oh it doesn't matter (. . . ) j'ai no hang on a minute elle aime . . . la reg-non umm . . . umm . . . oh how do you say it? umm elle aime . . . umm . . . regarder le musique or something like that (. . . ) she likes musique . . . j'ai adore /la, le/ musique (. . . ) j'ai déte- umm elle no . . . elle a déteste la piscine (. . . ) euh j'ai adore . . . oh no Monique j'ai adore . . . no Monique ELLE adore la . . . regarder la television [. . . ] (p. 350)

The pupil here first uses a first person reference (*J'ai adore*) instead of third person reference, and consequently seems to understand that this is not adequate for the task (*j'ai no hang on a minute elle aime*). She then attempts on numerous occasions to replace *J'ai* with the third person *elle*. This drive for adequate reference leads Myles et al. (1998) to conclude that:

> [t]he need to establish reference correctly is a strong driving force behind their analysis of the pronominal system within chunks. Because the chunks they learn in the early stages typically contain verbs and their pronominal subjects [. . . ] and because reference is crucial in the kind of information-gap classroom activities they typically engage in, subject clitic pronouns seem to be particularly salient and are the first items to be dropped or added to when they are inappropriate. (Myles et al. 1999, p.76)

The notion that the need to establish reference is a strong driving force can also be seen in the investigation of *task demands* in the area of Task-Based Learning (TBL). Although not directly related to the breakdown of

formulas, studies in TBL point to the importance of the conceptualisation of task demands in the selection of sequences from the linguistic environment. Samuda (2001), for example, argues that an initial focus on meaning in a task environment creates a 'need to mean' (p.122), which learners can only partially fulfil with their current resources. The gap created by this need drives learners to mine the linguistic environment for resources that can meet the need. That task demands create a need is similar to the proposal by Myles et al. (1998) that the exigencies of the task drive the breakdown of chunks as both see communicative pressure as resulting in a need for further resources. The difference, however, is that in Samuda's work, the linguistic environment is seen as the resource that can fill the gap between current ability and language useful to the task, while Myles et al. see communicative pressure as the driving force behind the breakdown of already-acquired formulas.

Further evidence for a link between the conceptualisation of task demands and the selection of task-relevant tokens from the input comes from Robinson (2005), who investigated the effect of varying the conceptual complexity of intentional reasoning tasks on the uptake of task relevant tokens in a picture retelling task. During completion of the oral task, participants were provided with nine task-relevant phrases of the construction type verb-continuous-object, e.g *he's carrying a plank*, and nine phrases of the construction type verb-$3^{rd}$ps-adverb, e.g. *he walks tiredly.* Uptake of the task-relevant phrases was divided into two subcategories: exact uptake, which was the selection and use of the full verb-continuous-object construction, for example *he walks tiredly*, and partial uptake, which was the selection and use of elements from the input, including lexical elements (e.g. *walk, tiredly*), partial phrases (e.g. *walks tiredly*), or altered phrases from the input (e.g. *walks very tiredly*). Results showed that in terms of partial uptake, learners who took part in more complex tasks selected and used significantly more partial phrases than in the less complex tasks (although not exact phrases).

Robinson's study (2005) is interesting, as it links task demands, input and performance. For example, learners in the study selected phrases such as *he's carrying a plank* that were helpful in the current communicative context, demonstrating that communicative need is related to the selection of forms in the input. This, hypothetically, may have led to the retention of a token such as *he's carrying a plank* which encoded the relevant element of the task conceptualisation. However, the study does not explore how the exigencies of the task may have led to the breaking down of input tokens, and how this breaking down may have led to low-scope patterns. The following section picks up the idea that task demands create a motivation for the selection of phrases, and discusses how this may factor in the breakdown of tokens.

### 2.2.2 Processes acting on tokens in the input: categorisation, analogy and similarity

The focus of the previous section was the extent to which the linguistic environment and communicative need could motivate the selection and/or breakdown of formulas. This next section builds on these motivations by reviewing more fine-grained processes that may be involved at the point of breakdown. The principal process considered is that of the categorisation of tokens and how this categorisation may be influenced by drawing analogies between tokens.

#### 2.2.2.1 Categorisation of tokens

Langacker (2000) sees the ability to compare two forms as a fundamental function of cognition. This comparison, he suggests, is a form of categorisation, in which a unit such as the unanalysed formula *he pushed it off* in the linguistic system of a language learner is compared to an instance such as *he pushed it off* in the linguistic environment. In Langacker's terms, the conventionalised unit in the linguistic system is the *standard*, and the item in the linguistic environment, or usage event, is the *target*, although more common terminology is *type*, for standard and *token* for target. Hudson (2008) uses the example *cat* in

a isa network (e.g. a cat *isa* animal): in the utterance *Where's my cat?*, *cat* is a token. To understand the token, it is necessary to match it to a stored word, which is a type. The stored type *cat* is also a noun, which is a more abstract type higher in the hierarchy.

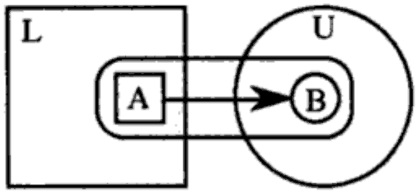



**Figure 2.5:** *Categorising usage events (Langacker, 2000, p.10).*

In Figure 2.5, L denotes the linguistic system (of, for example, an L2 learner), and A a conventionalised unit within that system, e.g. the unanalysed formulaic type *he pushed it off*. The circle is the usage event (U) (for example, a conversation) with a unit B as part of the event, e.g. the token *he pushed it off*. The arrow shows the relationship between the conventionalised unit and the usage event, the latter being an instantiation of the former. With the recursive categorisation of a token in the linguistic environment such as *he pushed it off* and a nearly identical formulaic type in the linguistic system, the formula is gradually *entrenched* further in the linguistic system.

Langacker also suggests that when there exist differences between a type such as *he pushed it off* and a token such as *he pulled it off* the categorization of the token as an instantiation of the type has the effect of producing a higher-level schema of the commonalities between both units. For example, if there exists *give it me* as an established linguistic unit in the system (a type) and a usage event *give it me* is encountered (token), there are very few differences between the system and the event (although there will be differences in the context and pronunciation), and therefore, as outlined above, this leads to the categorisation and further entrenchment of *give it me* in the system. However, if a slightly different token such as *give it her* is encountered, only *give it* is further

entrenched and the *her* and *me* lead to a schematic slot in the construction (cf. Wray's 'loosening' of fixed formulaic sequences based on evidence from the linguistic environment). Diagrammatically, this can be seen in Figure 2.6. Here, the linguistic system is represented as A in a box (e.g. *give it me*), a usage event by B in a circle (e.g. *give it her*), and the 'residue' of the unit and event by $A'$ (e.g. *give it her/me*). The resultant unit *give it* $X_{her/me}$ can be considered a very low-level schema with one constrained variable. If this schema becomes entrenched, it can be utilized in the interpretation of, for example, *give it* $X_{her/them/that\ man}$, which extends the abstractness of the schema (e.g. *give it* $X$). This process, given suitable and recursive usage events, would ultimately produce the schema Subj V $\text{Obj}_1$ $\text{Obj}_2$.

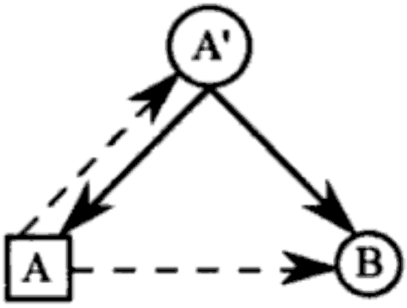



**Figure 2.6:** *Categorisation of and Schematisation from two tokens. Full arrows denote an instantiation of the type. Dotted arrows denote an extension of the type. (Langacker, 2000, p.13).*

Although presented differently, this process of gradual schematisation is essentially the same as that put forward by Ellis (2010, see also Tomasello 2003). In Figure 2.7, *put it on the table* is initially entrenched in the system. Subsequent tokens *put it on the desk/bag/bed* are categorised by the unit *put it on the table*, with a slightly schematised unit as the result (*put it on the X*). This schematised unit can then be marshalled to categorise other more diverse usage events, e.g *set it on the table*, which again extends the schematised unit. This process of extension and schematisation eventually results in the entrenchment of an abstract unit 'V Obj $\text{Obl}_{pathlloc}$'

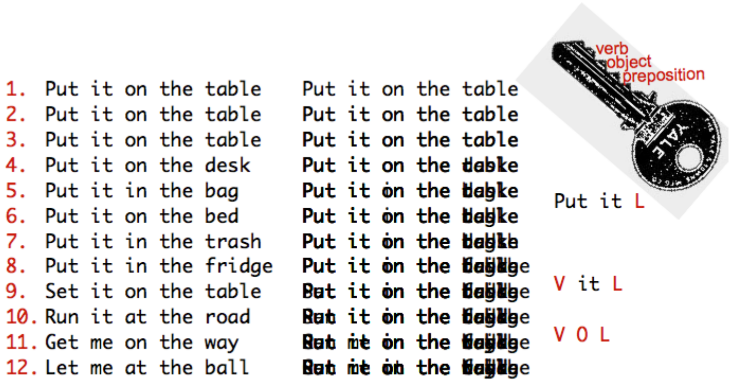



**Figure 2.7:** *Extension and schematisation of usage events (N.Ellis 2010, p.32).*

A key question is what motivates an entrenched type such as *put it on the table* to categorise tokens in the linguistic environment. As discussed previously, with high token frequencies, the hypothetical outcome is an entrenched formula, as the two forms are by and large identical. However, at the other extreme, if we consider *put it in your pocket*, *pull him along the top* and *threw them on the floor* as tokens of the type V Obj $\text{Obl}_{path/loc}$, without a priori knowledge of V Obj $\text{Obl}_{path/loc}$, it is not clear how a learner can determine that such tokens with different surface elements are the same type. She could in fact accept every token as a separate token and not a type. The next section addresses this by considering the role of analogy and specifically the degree of similarity in the categorisation of tokens.

#### 2.2.2.2 The role of analogy and similarity in the categorisation of tokens

At least three factors may determine whether a token in the linguistic environment is categorised by a stored linguistic unit. The first is surface similarity between tokens, e.g. *He put it* in *He put it there* and *He put it here*. The second and third factors are more structural. Items may be categorised if they share the same or similar meaning, as in the caused-motion construction, X CAUSES Y TO MOVE $\text{Z}_{PATH/LOC}$ or the same form, such as Subj V Obj $\text{Obl}_{path/loc}$. A theory which accounts for both the abstract and the concrete is

Gentner (1983)'s *structure mapping theory*. Within structure mapping theory, the central processes that facilitate comparison and categorisation are those of *analogy* and *similarity*. After first reviewing this theory outside the area of language acquisition, this section looks at how it can apply to the surface and structural properties of language.

In Gentner's structure mapping theory (1983) a distinction is made between *literal similarity* and analogy[10], in the mapping of a *base* domain and a *target* domain. In the case of literal similarity, a base domain, such as the universe, is mapped on to a target domain, such as another universe. Here, the base and the target share *objects* (e.g. suns, planets) and *object attributes*, such as the stars in both the base and the target are yellow and medium-sized. They also share *relational structure*, i.e. they share relations between objects, such as the planets in both the target and base domains revolve around the suns. To create a comparison, the objects, object attributes and relational structure are mapped from base to target, as illustrated in 2.8. In the case of analogy, the objects and object attributes are different but the base and target share relational structure. Gentner (1983) uses the example of our solar system and an atom: attributes of the sun, such as medium-sized and yellow, are not shared between the base and target, but the relational structures are, e.g. the planets revolve around the sun and electrons revolve around a nucleus, and a nucleus attracts electrons and the sun attracts the planets. To create a comparison here the relational structure needs to be mapped from the base to the target domain.

Gentner and Markman (1997) argue that the mapping of relational structure dominates over the mapping of literal similarity. For example, the two ratios 1:3 and 3:9 are related both in terms of attributes (the shared *3*), and relational structure (they are proportional). Although both mappings are possible,

[10]Gentner adds the possibility of *mere similarity*, which would include tokens which appear to be similar but do not share similar structure. Mere similarity will not be taken any further here.

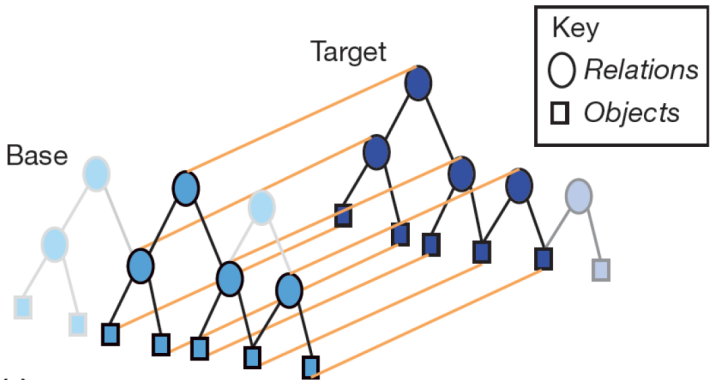



**Figure 2.8:** *Mapping objects and relations from a base domain to a target domain (Gentner and Smith, 2012)*

they argue that the dominant comparison is based on the relational structure of proportion. Markman and Gentner (1993) tested this hypothesis by asking participants which of the objects in the picture at the top of Figure 2.9 'goes with' an object in the bottom picture. The participants responded by choosing the identical cars. However, when the participants first compared the similarities between two pictures and were then asked which objects in the first picture went with which objects in the second picture, they mapped the car on to the boat, suggesting that they were mapping relational structure (A is towing B) and ignoring object similarity.

#### 2.2.2.3 Structure mapping and schema abstraction

While the previous section focused on mapping a base and a target domain based on similarity and analogy, this section focuses on the extent to which this process of comparing two tokens leads to the the *development* of structure. Gentner et al. (2004, Gentner et al., 2003; Loewenstein et al., 2003) use the term *analogical encoding* to describe the process whereby the explicit comparison of two analogical cases, i.e. two tokens, 'can foster the discovery of common principles and allow transfer to new structurally similar situations' (para 1).

Gentner et al. (2004, Experiment 1) investigated the extent to which analogical encoding leads to this abstraction of common principles. The partici-

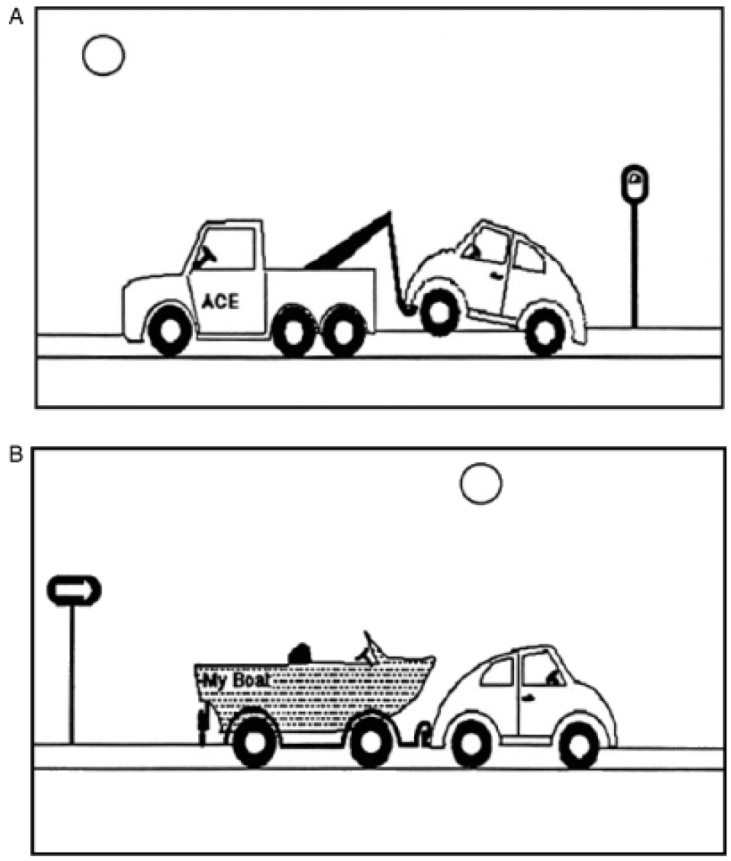



**Figure 2.9:** *Structural alignment and similarity between scenes (Markman and Gentner 1993)*

pants, all management consultants, were presented with a negotiation role-play on the agreement of contingent contracts (contracts based on a future performance[11]). Beforehand, half of the participants (n=124) were given two analogous exemplars of contingent contracts and were asked to compare the two cases (the Comparison Group) by finding similarities and parallels between the cases, while the other half studied the same exemplars separately (the Separate Group). In the subsequent role-play, participants were graded on how well

[11]Gentner et al. (2003)) give the following example of a contingent contract: 'a book author was negotiating with her book publisher about royalty rates. The publisher offered a 12% royalty rate; the author wanted an 18% royalty rate. The publisher believed that book sales would be relatively modest, but the author believed that book sales would be high. Whereas a compromise solution would involve settling for a 15% royalty rate (halfway between the parties' stated demands), a contingent contract agreement would tie royalty rates to sales. For example, the royalty rate would be 12% if book sales were less than or equal to some agreed-on level and would rise linearly as book sales increased' (p.393)

their performances matched the prototypical schema for the negotiation of contingent contracts and how well they retrieved specific past examples. Gentner and colleagues found that the Comparison Group significantly outperformed the Separate Group on both measures, and that the Separate Group did not seem to be able to treat the two cases as exemplars of the same principles, i.e. they did not analogically encode the cases although they were presented one after the other.

The previous study highlighted the extent to which analogical encoding resulted in target gains in experts (management consultants). In a similar study, Gentner et al. (2003, Experiment 1) used the same procedures and the same distinction between groups as Gentner et al. (2004), i.e, one group was given two analogous cases and asked to draw parallels between the two, while the other group studied both cases sequentially. The participants in this study, however, were novices selected from undergraduate programmes (n=128). Participants were asked to write the solution to an analogous case and were scored on the extent to which they had an understanding of the schema that the cases were instances of. Similar to Gentner et al. (2004), they found that the participants who initially drew parallels between the two cases were twice as likely to transfer principles to the test case than the participants who studied the analogical cases separately. Furthermore, the participants who studied the cases separately showed 'little evidence of schema abstraction' (p.400), with less that 20% making a comparison between the two analogical cases, and overall not significantly outperforming the baseline.

In both studies, Gentner et al. conclude that the participants who studied the cases separately encoded information specific to that case, while the 'analogical' participants drew parallels between the exemplars. This comparison, they argue, results in the abstraction of a schema based on the principles that the two analogical cases have in common. This abstract schema can then be drawn upon in the categorisation of future related tokens.

The findings from both these studies can be related to Langacker's view of schematisation (Figure 2.6 on page 60). Langacker describes how a token in a usage event is categorised by a established linguistic unit (be that a entrenched formula or a more abstract construction), the result of which is either an entrenched formula if the tokens are approximately equal, or a schema if the tokens had different properties but were categorised as exemplars of a type. Gentner et al.'s findings demonstrate that that one major factor in this categorisation is the ability to draw analogies across exemplars based on the identification of shared structure.

#### 2.2.2.4 A developmental perspective on the role of similarity and analogy

In the two studies outlined above, Gentner et al. (2003, 2004) argue that the encoding of analogous cases results in the abstraction of commonalities between the cases, and that this abstracted schema can then be employed for further analogous cases and so on. Gentner and Rattermann (1991, Gentner 2003) refer to this progression as the *career of similarity*. Early on in the career (e.g. with novices), categorisation is made possible by literal similarity between exemplars, as no shared structure exists. The result of this categorisation is the schematisation of shared properties between the tokens, which subsequently enable more abstract categorisation based on both literal similarities and the newly-formed relational structure. Later on in the career and following numerous categorisations of tokens based on similarity and relational structure, categorisation is possible by drawing on solely relational structure. Gentner et al. (2007), for example, found that when attributes were very similar (and relational structure was shared), younger children found it easier to make a comparison between tokens. This initial comparison then made it subsequently easier to categorise objects that had different object attributes. This suggests that the initial categorisation resulted in a more abstract understanding of the relational structure, which then facilitated the categorisation of objects with

similar relational structure but fewer object attributes than in the initial comparison. The idea of varying degrees of utility for similarity and analogy is important to SLA, as it is this process, arguably, that may enable learners to bootstrap into the processing of more abstract structures. This is picked up in more detail in *Section 2.2.2.6: A career of similarity in language learning.*

#### 2.2.2.5 Analogy and similarity in language learning

Analogy and similarity in language acquisition would seem to be particularly complex, as any theory would need to address the similarity between surface forms as well as structurally align formal elements (e.g. Subj V Obj $\text{Obl}_{path/loc}$) and constructional meaning (e.g. X CAUSES Y TO MOVE $\text{Z}_{PATH/LOC}$).

In FLA, Tomasello (2003) proposes a distinction between *schematisation* and *analogy*. He sees schematisation as the outcome of categorisation based on attribute similarity, e.g. the shared *give* in *give it me* and *give her them*. Analogy, he argues, is reserved for the categorisation of two elements based on the structural alignment of both the form and function of language without attribute similarity, e.g. both *give it him* and *send them the fish* share the same ditransitive meaning of physical or metaphorical transfer and the same categorial form V $\text{Obj}_1$ $\text{Obj}_2$, but do not share physical characteristics. This is illustrated in Figure 2.10, in which the surface forms do not share physical attributes but the two forms are structurally aligned in terms of both formal categories and meaning. Figure 2.11, on the other hand, shows surface forms with similar attributes (*pushed, it*), which are also structurally aligned in terms of meaning (X CAUSES Y TO MOVE $\text{Z}_{PATH/LOC}$). The formal categories are omitted in Figure 2.11 as Tomasello sees schematisation as an earlier stage of language acquisition when, it is argued, formal categories such as Subj V Obj $\text{Obl}_{path/loc}$ are not available. In line with the notion of schema abstraction as a byproduct of analogical processing (Gentner et al., 2003, 2004; Langacker, 2000; Loewenstein et al., 2003), the processes of schematisation and analogy in

Figures 2.10 and 2.11 have as a byproduct of such processing an abstraction based on the commonalities in the base and target domain.

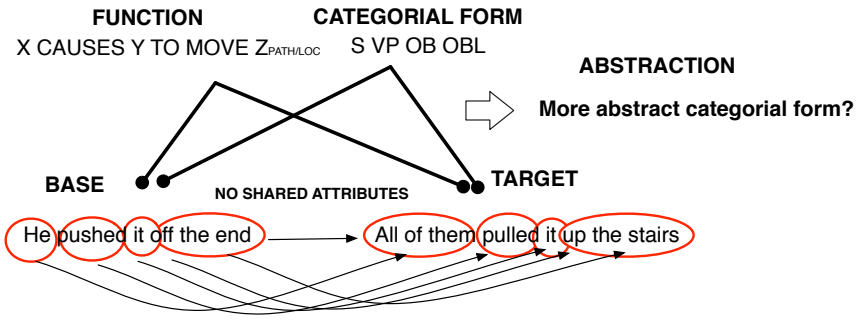



**Figure 2.10:** *Structural alignment in the analogical processing of tokens (based on Tomasello 2003).*

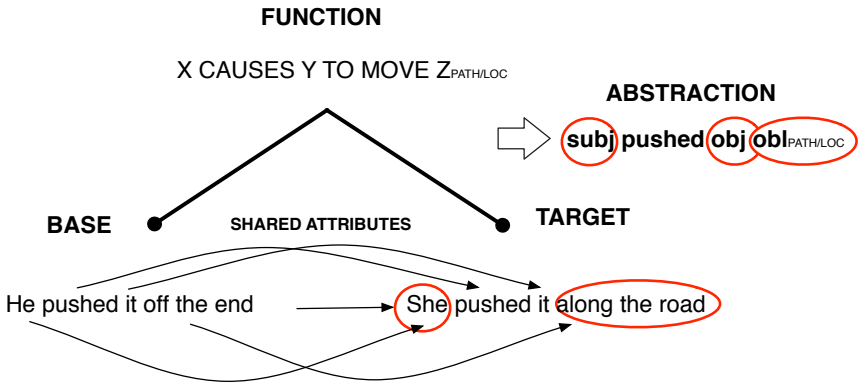



**Figure 2.11:** *Schematisation based on attribute similarity (based on Tomasello 2003).*

Bybee (2010) takes a different view on analogy and similarity by conflating the two into one mechanism with considerable emphasis on shared attributes (in addition to the structural alignment of meanings, such as X CAUSES Y TO MOVE $Z_{PATH/LOC}$, shared by two instances). This would seem to be a consequence of Bybee's view of the cognitive representation of language in the form of exemplars, whereby 'every token of experience is classified and placed in a vast organizational network as a part of the decoding process' (Bybee, 2006, p.716). By emphasising the storage of exemplars and deemphasising the role

of abstract structure, the processing of exemplars relies more on the similarity between surface forms and less on the structural alignment of categorial form.

Both Bybee's and Tomosello's views on analogy and similarity have assumed that the end product of learning is the ability to be productive, either via abstract constructions (Tomasello, 2003), or via some form of critical mass of exemplars (Bybee, 2010). However, as discussed in *Section 2.1: Usage and a trajectory from formula to construction*, the synchronic view of language taken in the present study is that there exists a continuum from formulas (e.g. *by and large*) to low-scope patterns (*Why do-TENSE n't NPi pick on someone PROi-gen own size!*) to abstract constructions such as the passive construction Subj aux VPpp (PPby). At the more abstract end of the continuum, either structural alignment between tokens in the input based on analogy is pervasive (Tomasello, 2003), or structural alignment in addition to object similarity is pervasive (Bybee, 2012). At lower levels of abstraction, however, such as in the case of formulas and low-scope patterns, literal similarity is key, as the end point of learning a construction such as *The Xer the Yer* retains morphological material from the input (*the, er/more*), which can be employed in categorisation based on similarity.

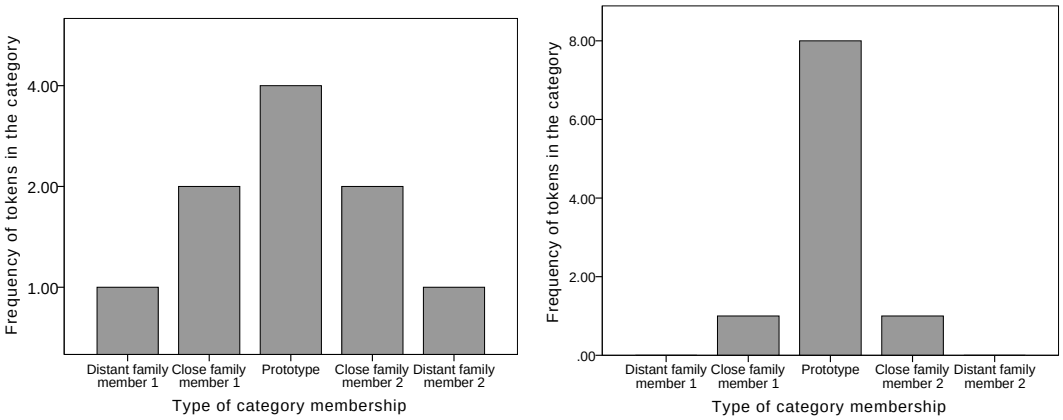
#### 2.2.2.6 A career of similarity in language learning

The relative prominence of similarity and analogy at different stages of language development can be related to Genter et al.'s *career of similarity* (Gentner, 2003; Gentner and Rattermann, 1991, discussed in *Section 2.2.2.4: A developmental perspective on the role of similarity and analogy*). As Gentner et al.'s framework outlines, in the initial stages of categorisation it is assumed that little or no structure can be utilised to categorise analogically related tokens, and therefore categorisation here is based on similarity between tokens. If, however, initial analogical encoding results in the abstraction of a low-scope schema, this schema can then be drawn upon in the analogical encoding of further analogically-related tokens that share fewer object attributes. This

updated, more abstract schema, can subsequently be drawn upon in future categorisations, and so on.

Parallels can be drawn between a career of similarity (e.g. Gentner and Rattermann, 1991) and Tomasello's schematisation-analogy distinction . Although Gentner and Rattermann (1991) see the varying importance of analogy or similarity in the development of structure as a cline between similarity and analogy, not a binary distinction, both Gentner and Rattermann (1991) and Tomasello (2003) see similarity between two tokens as prominent in the early stages of acquisition when no categorial structure is available. At more competent levels when the learner has some grasp of structure, the prominence shifts to analogical processing. This view of analogy and similarity is also comparable to N.Ellis' trajectory (2002) from formula to low-scope pattern to construction. In the initial stages, categorising tokens such as *give it me* and *give it her* based on literal similarity theoretically leads to a more abstract low-scope patterns (*give it x*), which can then be employed in the categorisation of tokens with fewer shared attributes, e.g. *give it them/him/John.*

That similarity between tokens can affect the initial acquisition of a construction sees support from both Elio and Anderson (1984) and Casenhiser and Goldberg (2005). Outside the area of language acquisition, Elio and Anderson (1984, Experiment 1) presented three groups of participants (n=45) with tokens derived from a category that showed typical prototype effects, i.e. the central member of the category was more likely to occur than peripheral members (Figure 2.12a). Group one were presented with a representative sample of the category, as in Figure 2.12a; group two were presented first with a skewed frequency of tokens which focused on the central prototype and close family members with a high number of overlapping features between tokens (Figure 2.12b), and were subsequently presented with the full category membership; and group three were first presented with full category membership and then with a skewed frequency. A transfer test was then used where participants rated seen and not-seen-before category members on a cline between typical of




**(a)** *Presented tokens representative of category (high variance)*

**(b)** *Presented tokens skewed towards prototype (low variance)*

**Figure 2.12:** *High and low variance presentation of category members (Elio and Anderson, 1984)*

the category and not typical of the category. They found that the group who were presented with a skewed frequency first significantly out-performed the other groups, although all groups at the end of the study phase had encountered the same number and type of tokens of the category. They argue that the initial low variance allows learners to get a 'fix' on central membership of the category based on similarity between tokens, which will enable later tokens to be categorised by an emerging schema.

In a similar study in terms of high levels of similarity across tokens in the initial stages, but this time in FLA, Casenhiser and Goldberg (2005, Experiment One) presented novel expressions, e.g. *The rabbit the hat moopoed*, with a novel meaning, e.g. a rabbit appearing on a hat, to English-speaking children (n=51). The children were divided into a control group which did not receive any input, a balanced frequency group which was presented with five nonsense verbs each occurring once or twice, and a skewed frequency group which was presented with again five nonsense verbs, one of which occurred four times, and the others once each. The groups were tested using a forced choice comprehension test with novel verbs, in which participants pointed to a scene which corresponded to a description. An ANOVA showed a significant main effect for

group, with both the balanced and skewed frequency groups outperforming the control, and the skewed frequency group performing significantly better than the balanced frequency group.

Boyd and Goldberg (2009) explain these results by suggesting that a *skewed* frequency allows learners to 'anchor' (p.422) both the meaning and the form of the construction. Once anchored (or 'fixed' in Elio and Anderson (1984)'s terms), further evidence in the linguistic environment can allow for the breakdown of these prototypical formulas into more productive constructions. Relating this explicitly to the breakdown of formulas, these 'anchors' can be seen as high frequency tokens which have become entrenched as formulas with limited or no variability. Their breakdown is then motivated by evidence for variability in the linguistic environment (e.g. alternative verbs in the VAC constructions in Ellis and Ferreira-Junior, 2009b, as discussed in *Section 2.2.1.1: Motivation from the linguistic environment*).

Although there is little research in SLA on 'anchoring' a token by presenting low-variance, high similarity tokens, McDonough and Kim (2009) provide evidence that once a token is anchored, a high type frequency leads to productivity within a *wh*-question syntactic frame. For example, if there initially exists the unanalysed frame *What's he doing?*, subsequent contact with different tokens of the same type e.g. *What's the dog doing?*, *What's he eating?, When's he eating?, What was he doing?* (ibid., p.388) leads to the initial unanalysed token becoming productive. McDonough and Kim tested this theory of high type frequency with Thai learners with an incipient knowledge of *wh*-questions studying EFL (n=85), by exposing the learners to instances of *wh*-questions (primes) either with a high type frequency (36 questions with 36 different verbs), or low type frequency (36 questions with 6 different verbs), and then prompting the learners to produce *wh*-questions (prompts) by providing them with either a low prompt type frequency (six verbs) or high prompt type frequency (36 verbs). Based on the accurate production of *wh*-questions in pre and post-tests, they found a significant main effect with a medium effect size for

high prompt type frequency over low prompt type frequency ($p$=.021, $\eta_p^2$=.064), i.e. the production of *wh*-questions based on a high number of prompts led to gains in accurate production between pre- and post-tests. However, there was no significant main effect for prime type frequency, i.e. exposure to instances of *wh*-questions did not seem to have any effect on gain scores. As McDonough and Kim note, although such a short intervention (2x20-minute sessions) may not have allowed for the processing of the primes, the high prompt type frequency may have forced the learners into the grammatical processing of the *wh*-question syntactic frame. More longitudinal studies could presumably establish the extent to which the reception of instances (primes) can lead to greater productivity of *wh*-constructions specifically, and of other construction more generally.

### 2.2.3 Summary

While *Section 2.1: Usage and a trajectory from formula to construction* outlined evidence for a trajectory from formula to construction in SLA, this section was concerned with the motivation and processes that would allow for such a trajectory. Empirical evidence from Myles et al. (1998) and Myles et al. (1999) suggest that a key motivation for the breakdown of formulas is the need to establish reference that is beyond the capability of existing formulas. As task demands, by definition, create the need to establish reference, arguably the investigation of task demands as a causative factor in the breakdown of formulas would be an interesting direction for future research, although to date there is little work in this area.

The initial breakdown of formulas also seems to be at least in part dependent on the shape of the linguistic environment, specifically the 'skewed' frequency distribution within constructions. Ellis and Ferreira-Junior (2009a), for example, correlate native-speaker productions with the acquisition and use of non-native productions and put particular emphasis on the prototype which initially seeds the formula, and the skewed 'Zipfian' profile of the input. The

link between frequency distribution and the role of categorisation can be seen in the work of Elio and Anderson (1984) and Casenhiser and Goldberg (2005), who both suggest that the skewed frequency of the input facilitates an initial 'fix' on a formula and its meanings, which in turn facilitates the categorisation of similar tokens.

Gentner's *career of similarity* offers a model within which there is a shift from categorisation based on literal similarity to categorisation based on analogy. Initial categorisation of similar tokens leads to the abstraction of common principles that both tokens share. This abstraction then leads to the categorisation of a base and a target based on more abstract structural properties. In terms of language development, very similar formulas such as *give it me* and, with different context and pronunciation *give it me* lead to entrenchment, or strengthening, of the formula. This established formula can then categorise similar formulas such as *give it her*, the result of which is the creation of a low-scope schema. Following Gentner, this low-scope schema can then be drawn upon to categorise even less similar tokens, until the point at which there exists no shared morphological material between the base and the target, at which point categorisation is based on pure analogy. In SLA, although McDonough and Kim (2009) provide some evidence that when learner's have an incipient knowledge of a category (e.g. *wh*-questions), a high type frequency can further the productivity of the type, there is little research in this area.

## 2.3 The current research

### 2.3.1 Summary and research space

The starting point of this literature review was the hypothesis that there exists a trajectory from the detection of tokens in the linguistic environment and their entrenchment as rigid formulas, through the breaking down of such formulas into more productive low-scope patterns, to the emergence of more abstract constructions. Although earlier longitudinal studies in SLA such as Hanania

and Gradman (1977) and Shapira (1978) found little evidence to suggest such a trajectory from token to formula to pattern, these studies tended to operationalise acquisition as the use of morphemes in obligatory contexts rather than the detection, use and development of formulas. More recent empirical research suggests that when SLA is operationalised as the development of formulas from tokens rather than the development of abstract morphemes, there is evidence that formulas are initially rote learnt, and that they are subsequently broken down into more productive patterns (Eskildsen, 2009, 2012; Myles et al., 1998, 1999).

Evidence that there exists a trajectory from token to formula to construction raises the question of what motivates both the initial selection of tokens and the subsequent development along the proposed trajectory. In terms of the initial selection, Samuda (2001) found that an initial need to mean within a task- based environment motivates the selection of relevant tokens from the input that convey the intended meaning. The notion that need to convey a meaning motivates selection is shared by Robinson (2005), who found that increased task complexity leads to a greater number of selections of task relevant tokens from the input. While both Robinson and Samuda cite task demands as motivation for the selection of tokens, neither examine the role that task demands may have in the subsequent development of selected tokens. In terms of the development of structure, Myles et al. (1998, 1999) suggest a correlation between a repertoire of entrenched formulas and the development of a linguistic system (the pronominal system in this case), and further suggest that a catalyst for development is a gap between a learner's current repertoire of formulas and the exigencies of the task to fit a current communicative need.

However, while evidence from studies such as Eskildsen (2009) and Myles et al. (1998, 1999) suggests that there does exist a trajectory for at least some learners with at least some constructions, the relatively sparse sampling of these studies, and their inability to control for input to learners does not allow for a fine-grained investigation of the relationship between the selection of tokens

from the input and the use of such selections. This study therefore builds on such research by investigating both the degree to which past-counterfactual instances were selected from the input, and the degree to which the selections were broken down to meet communicative task demands. By tracking the selection and breakdown of instances, the study provides much needed research into how communicative need may encourage the initial selection of instances, their subsequent repurposing (seen here as changes to selections in order to fit the task demands), and whether or to what extent this repurposing may lead to the development of a past-counterfactual construction.

For a more fine-grained analysis of what may be involved when there exists a gap between current knowledge and task demands, this study employs analogy and similarity in the categorisation of selected tokens. In a non-linguistic domain, Gentner and colleagues (e.g. Gentner, 1983; Gentner and Markman, 1997) outline the mapping of the structure and objects of a *base*, i.e. the solar system, onto the structure and objects of a *target*, i.e. an atom, with similarities between the objects in the base token and objects in the target token facilitating the analogy. To relate this to task demands and the selection of task-relevant tokens, it is hypothesised that the conceptualisation of task demands can drive the selection of base tokens from the linguistic environment, and that such tokens once selected can subsequently be deployed in production (targets) to satisfy the demands of the same conceptualisation. Drawing on Tomasello's interpretation of analogy and schematisation in a linguistic domain, Figure 2.13 illustrates the proposed mapping of a base token *She would've come, if you had come* on to target token *We would've been happier, if they hadn't said anything.* The illustration presumes no prior structural knowledge of past counterfactual constructions.

In terms of development from such a process, Gentner and colleagues (Gentner et al., 2003, 2004; Loewenstein et al., 2003, see also Gick and Holyoak 1983) argue that the analogical process of mapping bases on to tokens is the abstraction of a schema based on the commonalities between bases and targets.

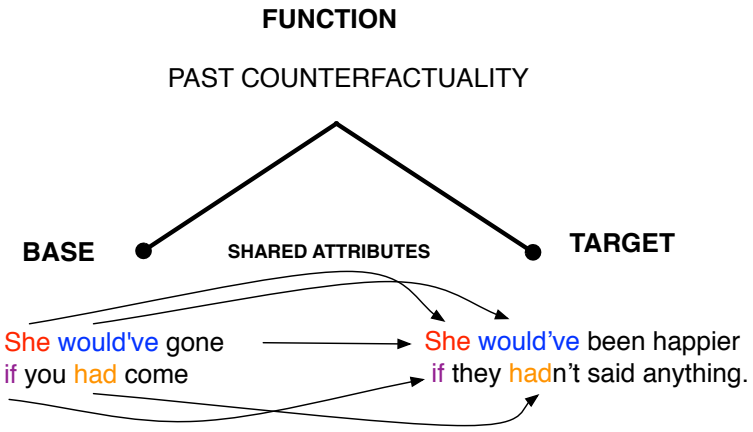



**Figure 2.13:** *Mapping of form and function in the analogical processing of past counterfactual tokens with shared attributes (based on Tomasello 2003)*

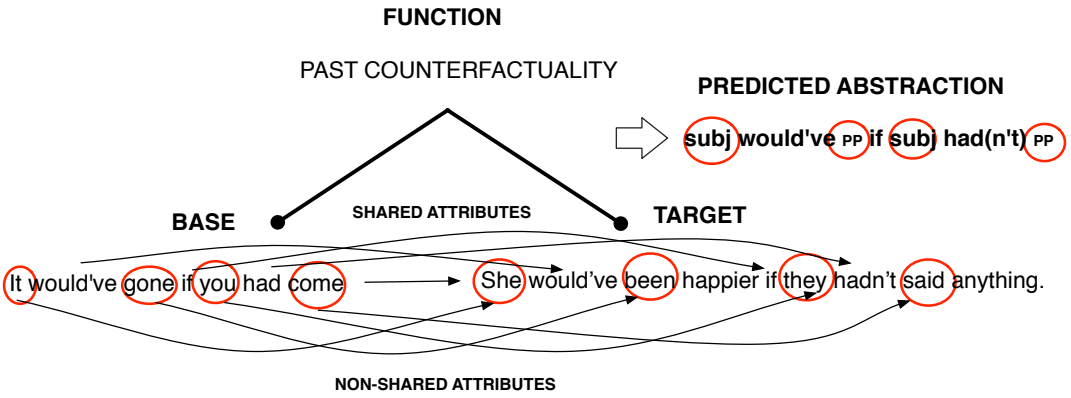



**Figure 2.14:** *Predicted schematisation of low-scope patterns based on the mapping of form and function in the analogical processing of past counterfactual tokens with shared attributes*

This abstraction is illustrated in Figure 2.14, in which a past counterfactual base token is mapped on to a target token. The predicted result of such a process is the entrenchment of almost identical items (*would've, if, hadn't*), and the initial loosening of constituents in for example the subject position (*it* → *it/she*), resulting in the establishment of a low-scope pattern with paradigmatic slots, in this case *it/she would've gone/been, if you/they hadn't come/said anything*. If such a result is possible, the low-scope pattern may be employed to categorise further not-seen-before past counterfactual tokens in the input.

Studies such as Loewenstein et al. (2003) suggest that with the application of recursive mapping of base and target the end product is an increasingly more abstract construction, in this case, hypothetically, the partially abstract and partially concrete construction *subj would've pp if subj hadn't pp*.

While the evidence from non-linguistic domains suggests that analogies are formed between tokens, and also that the by-product of such analogical processing leads to the development of structure, to my knowledge there exists no evidence in SLA that analogical processing of tokens in the input can be driven by task demands, or that analogical processing can lead to the development of constructional patterns. The current study, therefore, brings these lines of research together to investigate the role both task demands and analogy may play in the selection of tokens from the linguistic environment and the subsequent development of constructions.

In more precise terms, the current study hypothesises that task demands can be employed to drive the selection of tokens from the input (a base token) and their subsequent *exact* use in task performance (a target token), both under the same constructional meaning. This *literal processing*, i.e. the exact use of forms from the input, is illustrated in Figure 2.15. As the base and target tokens are identical, it is predicted that the by-product of such processing is the entrenchment of a token, but not a schematised pattern.

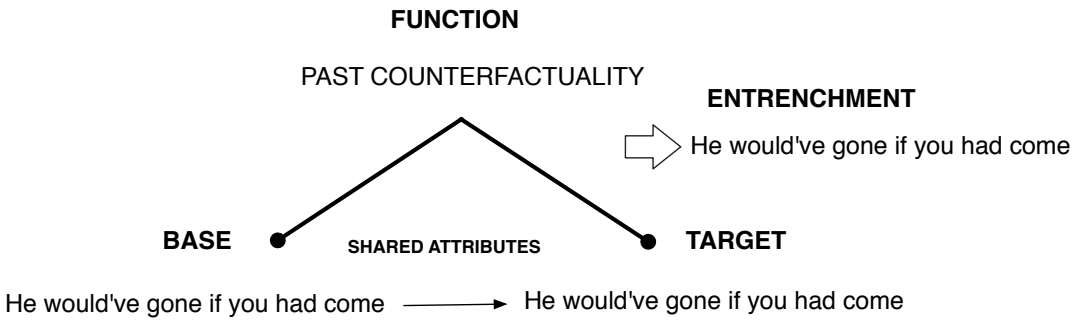



**Figure 2.15:** *Literal processing of selected input tokens (base) and output tokens (target)*

In contrast to the literal processing of tokens, it is further predicted that

task demands can drive the repurposing of the tokens selected from the input, in an attempt to produce output to meet these demands. *Repurposing* is used here to describe the alteration of a base token to fit a different purpose in production, e.g. the base token *It would've gone* is drawn upon to produce the target token *She would've been.* This *analogical processing* is illustrated in Figure 2.14, and is evidenced by the number of differences between tokens in the input and tokens in the output. This type of process will, as a byproduct of task completion, hypothetically lead to the abstraction of a schema. That is, the instances both in the input and the repurposed output will result in the formulation of a partially schematic construction that is similar to the input (*If, had, have*) and more abstract (the modal auxiliary slot and past participles in both result and condition clauses). This will render the resulting construction partially productive, and therefore able to 'sanction[...]novel expressions' (Langacker, 2000, p.26) in the subsequent categorisation of tokens. These effects of both literal and analogical processing are articulated in the following research questions.

### 2.3.2 Research questions and hypotheses

The first two research questions relate to task demands and the selection of tokens from input.

*RQ1: When learners are aware that task demands will require output that is identical to instances of counterfactuals in the input, does this encourage the selection of task-relevant instances of past counterfactuals?*

*RQ2: When learners are aware that task demands will require output that is analogous but not identical to instances of counterfactuals in the input, does this encourage the selection of task-relevant instances of past counterfactuals?*

Research question 3 and 4 relate to task demands and the repurposing of tokens

selected from the input.

*RQ3: When learners are aware that task demands will require output that is identical to instances of counterfactuals in the input, does this lead to the repurposing of instances of task-relevant counterfactuals?*

*RQ4: When learners are aware that task demands will require output that is analogous but not identical to instances of counterfactuals in the input, does this lead to the repurposing on instances of task-relevant counterfactuals?*

Research questions 5, 6 and 7 relate to the development of past counterfactual knowledge from the processing of the input.

*RQ5: Does the selection and use of exact tokens correlate with gains in past counterfactual knowledge?*

*RQ6: Does the selection and repurposing of tokens correlate with gains in past counterfactual knowledge?*

*RQ7: Do repurposing operations result in an abstract past counterfactual construction?*

Research question 8 relates to the effect that cross-linguistic differences in the encoding of past counterfactuality in L1s may have on the development of past counterfactual constructions in the L2.

*RQ8: Do cross-linguistic differences in the encoding of past counterfactuality in the L1 affect the development of counterfactual constructions in the L2?*

These research questions are supplemented by three directional hypotheses:

*H1: There will be no difference between a) the selection of tokens in the input when task demands require output that is identical to instances of counterfactuals in the input and b) the selection of tokens in the input when task demands require output that is analogous but not identical to instances of counterfactuals in the input.*

*H2: The repurposing of tokens to fit task demands will be greater when task demands require output that is analogous but not identical to instances of counterfactuals in the input.*

*H3: The selection and repurposing of tokens will lead to superior performance on post-tests than selection and exact use of tokens.*

# Chapter 3

# Methodology

This chapter outlines the methodology employed to explore the research questions and hypotheses developed in the previous section. The chapter is subdivided into two sections; the first presents the research design, including a description of the groups in the intervention, the linguistic target, the tasks used to elicit data relevant to the hypotheses and the testing instruments. The second section presents the procedures employed to score the data.

## 3.1 Research design

To explore the hypotheses, this intervention employed a mixed-method design, with a quasi-experimental pre-/post-test phase, supplemented by qualitatively-analysed stimulated recall interviews. Intact classes were randomly assigned to one of two treatment groups, and to a control (Analogous, Literal and Control). As can be seen from Figure 3.1, all groups were required to complete three task cycles, each comprised of their own planning and production phases.

In each cycle, the participants in all three groups wrote a story on the same topic, while the treatment groups also received pre-production input seeded with the target construction. The two treatment groups differed in the relationship between the task demands and the input to which they were exposed (the independent variable). The *Analogous Group* were provided with input that was analogous to their task demands, i.e. the group were required to con-

ceptualise a past counterfactual situation before being exposed to input that was functionally similar to their task demands, but which comprised lexical items that differed to ones needed in their own production, e.g. input (base): *If it hadn't been for the weather, it would've been good.*, target production: *If I hadn't failed my exams, I would've got a good job.* The *Literal Group* were given the same treatment as the Analogous Group, except that the base input matched the target output in terms of function, grammatical frame *and* lexical items, e.g. input (base): *If it hadn't been for the weather, it would've been good*, production to meet task demands (target) *If it hadn't been for the weather, it would've been good.* The *Control Group* completed the same tasks as the Analogous and Literal Groups, but received no input.

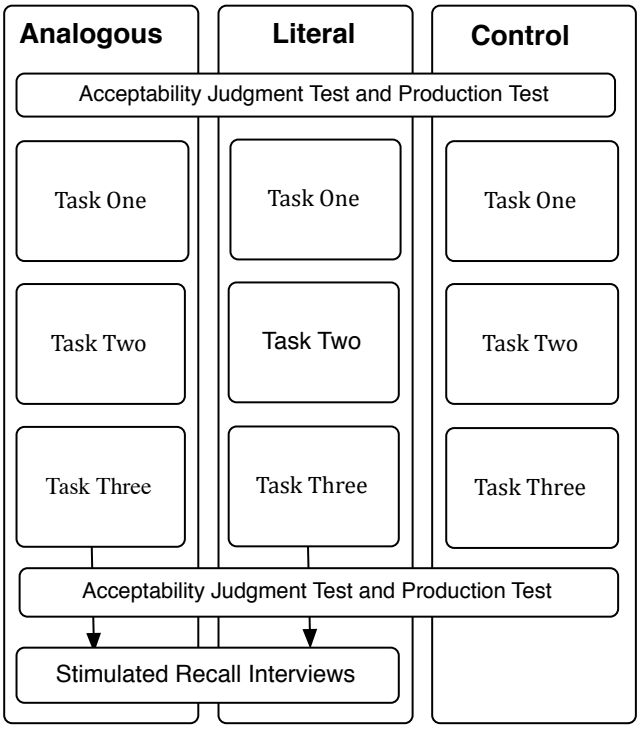



**Figure 3.1:** *Summary of Research Design*

There were three dependent variables (see Table 3.1): the selection of tokens from the input, the number of repurposed tokens in production, and scores

on the use of past counterfactuals in two pre- and post-intervention tests. All three dependent variables are described in more detail in sections 3.2.3: *Testing Instruments and Tasks* and 3.3: *Data Scoring*.

**Table 3.1:** *Dependent and independent variables*

| | |
|---|---|
| Independent Variable | Input at Three Levels:<br>1. Analogical Input<br>2. Literal Input<br>3. No Input (Control) |
| Dependent Variables | 1. Scores on acceptability judgment and production tests<br>2. The number of partial and full selections of condition and result tokens from the input<br>3. The number of repurposed past counterfactual tokens generated from selected tokens from the input |

### 3.1.1 Participants

The participants were a convenience sample of 92 university students between the ages of 18 and 40, studying on pre-sessional English for Academic Purposes (EAP) programmes at The University of East London. Participants were from a large number of linguistic backgrounds and were studying on intensive (15-20 hrs/week) programmes in order to reach the English language entry requirements of the University. Participants were to progress onto undergraduate (UG, n=34) or postgraduate (PG, n=58) programmes in September 2010, January 2011 or September 2011. The criteria for inclusion was that the students had achieved an score on the International English Language Testing System (IELTS), or equivalent, of between bands 5 and 6.5, and were taking part in pre-sessional courses. This particular sample population was chosen due to the similar proficiency of the students, their similar goals and course structure, as well the convenience in terms of the researcher's prior knowledge of both pre-sessional students in general and these students in particular .

**Table 3.2:** *Biographical data of participants in the study*

| | |
|---|---|
| L1 | Arabic (28), French (18), Vietnamese (9), Persian (6), Mandarin (4), Japanese (4), Bangla (3), Russian (3), Thai (3), Hindi (2), Italian (2), Bangladeshi (3), Spanish (2), Bulgarian (1), Gujarati (1), Kurdish (1), Lithuanian (1) and Urdu (1). |
| Gender | Female (36), Male (56) |
| Pathway (Under or postgraduate) | Postgraduate (58), Undergraduate (34) |
| Months of Classroom Study | Mean: 6 years, SD=5.7 |
| Months in UK | Mean: 8.4 Months, SD=12.2. |
| Months on Pre-sessional Course | Between 2 weeks and 6 weeks |
| Proficiency at start of course | All participants were B2 level on the Common European Framework (Council of Europe, 2001). 53 participants self-reported IELTS scores: average: 5.5[a] with no higher than 6.5, and no lower than 5. 7 gave TOEFL (Computer or iBT): IELTS equivalent of between 5 and 6. 32 participants did not state a test score, as they were tested internally (and achieved the equivalent of between IELTS 5 and 6) |

a) IELTS and TOEFL test scores were mapped on to the CEF using submissions to the UKBA

After an explanation of the study in the second week of the course, all students were provided with a written description of the study (Appendix C) and asked if they wished to participate. Those who wished to participate were given a consent form (Appendix A), which they were asked to sign only if they understood and agreed to what was proposed. Participants were reminded that they could leave the study at any point and for any reason, and that their personal details would not be disclosed. They were also reminded that the data would be destroyed after the study had been completed (See Appendix B for the related ethics approval form). As stimulated recall interviews did not directly focus on language learning, participants were offered £5 for their participation. Participation in the interviews was voluntary, and, after being

made aware that the same consent form and related practices applied, all those who volunteered were interviewed (n=27).

In addition to the consent form, participants were asked to complete a brief demographic questionnaire (Appendix A), the results of which are summarised in Table 3.2.

## 3.2 Materials

This section presents the materials used in the intervention, the linguistic target in focus and the testing instruments employed to measure acquisition.

### 3.2.1 Initial piloting of intervention materials and testing instuments

The tasks in the interventions and the testing instruments were piloted prior to the main study. In the first round of piloting, various task types (without testing instruments) were piloted with intact classes of pre-sessional students to determine the task type that was most appropriate. The second round of piloting (again without testing instruments) was used to assess whether participants interpreted the rubrics of the chosen tasks as they were intended and could negotiate the content of the input. This round of piloting was also used to determine the optimal timing needed to complete each stage of the tasks.

Both the tasks and testing instruments were then piloted with students at a different university (n=30). The main aim of this round was to determine the validity of the tests by assessing both performances on the tests and the students' opinions of the level of difficulty. This round also allowed for the whole design from pre-test through the three tasks to post-test to be piloted. The specific influences that the pilot study had on tasks, testing instruments and procedures are noted in the context of the main study below.

### 3.2.2 Intervention materials

#### 3.2.2.1 The linguistic target: past counterfactual constructions

The construction selected for this study is the past counterfactual, exemplified in 42, 43 and 44.

42. *If it hadn't been for the kids, I'd have gone*
43. *If I hadn't done it, it would've been a lot better*
44. *I would've enjoyed myself more if you'd been there*

A number of reasons suggested that the past counterfactual would be a suitable construction for the study. Firstly, past counterfactuals are often acquired late in formal instruction settings (Celce-Murcia et al., 1999), and therefore it is a structure which students at this level were likely to either find problematic or have little knowledge of. Secondly, there exist a number of studies which have focused on the acquisition of past counterfactuals by adult learners (e.g. Izumi and Bigelow, 2000; Song and Suh, 2008), and therefore this study can be situated and compared to previous work. Thirdly, past counterfactuals are semantically opaque, and grammatically complex. This means they are are an ideal construction to to investigate a trajectory from incipient knowledge to more generative knowledge.

**3.2.2.1.1 Past counterfactuals: abstract and concrete** As specified in Section 2.1.1: *A property theory as a continuum of abstractness*, the degree of abstractness of a given construction determines the extent of its trajectory (*by and large* has a minimal trajectory and S V $Obj_1$ $Obj_2$ a maximal trajectory). Dancygier and Sweetser (2005), working within a construction grammar framework in their analysis of counterfactuals, suggest that although counterfactuals can be described as an abstract grammatical construction, e.g. Example 45 below, past counterfactuals often have very limited productivity, e.g. *If it hadn't been for NP* in Example 42 above.

45. *If NP VP*-en *(COMP), NP* $_{AUX[might,could,would]}$ *VP*-en *(COMP)*

Even the most productive of past counterfactuals share attributes between tokens, sharing *had* and *have*, and often *if* and *would/could/might*. This suggests they are an appropriate construction for probing analogical processing, as their shared surface attributes are more easily identified as types due to the perceptual similarity across tokens (Casenhiser and Goldberg, 2005; Gentner and Medina, 1998; Gentner and Namy, 2006; Gentner and Rattermann, 1991). This is particularly relevant in relatively short interventions, in which learners do not have prolonged exposure to the forms and therefore have limited opportunities to categorise tokens.

**3.2.2.1.2 A hypothetical trajectory for past counterfactuals** From the perspective of the L2 user, any description of knowledge is dependent on their stage of development. An incipient knowledge of counterfactuals could be described as formulaic, and then at later stages partially lexically-open and finally fully lexically open. For example, 46 may be a single unanalysed formulaic chunk, not generated by an underlying system.

46. *[If it hadn't been for the weather, it would've been good]*$_s$

Alternatively, it could be described as partially productive, as in Example 47. In this particular case, the VPs could be limited in scope to a small number of verbs and nouns.

47. *If I'd* $_{[VP-\text{en }(COMP)]}$[1]*, I would've* $_{[VP-\text{en }(COMP)]}$
e.g. *If I'd* $_{gone/been\ there/come}$*, Iwould've* $_{had\ a\ good\ time/been\ happy}$

A later stage would relinquish constraints on both the VPs and complements (48).

48. *[If NP VP*-en *(COMP), NP* $_{AUX[might,could,would]}$ *VP*-en *(COMP)]*

[1]The convention here follows that of Jackendoff (2002). The VP is subscripted to show that it is not a fully-productive VP, but one that is limited in scope to, for example, semantically related verbs, or a number of semantically unrelated memorized verbs. A non-subscripted VP is free within semantic and pragmatic constraints.

**3.2.2.1.3 Meaning of past counterfactuals** Past counterfactuals encode a speaker's negative epistemic stance on the truth of the situation. This stance is expressed in the distancing of the verb forms in both the condition and result clauses (Dancygier and Sweetser, 2005; Quirk et al., 1985). In the condition clause, this counterfactuality can be seen if we compare factuals and counterfactuals in the present[2].

49. *If I go (factual, present time)*
50. *If I went (counterfactual, present time)*

Here it can be seen that present counterfactuality is indicated by a back-shifting of the form of the verb (from *go* to *went*) to a 'fake' past (Iatridou, 2000). To encode past time, and therefore *past* counterfactuality, the form of the verb is back-shifted further, resulting in Examples 51 and 52:

51. *If I went (counterfactual, present time)*
52. *If I had gone (counterfactual, past time)*

The conditional clause in the past counterfactual is encoded in the same way as the past perfect, although semantically it maintains a simple past meaning (Quirk et al., 1985, p.1010). In the main result clause, the past time is encoded in the perfect infinitive - *have gone* - which follows the modal verb. Again, this can be illustrated by comparing counterfactuals in both the present and the past:

53. *If we went, it would be good. (present counterfactual)*
54. *If we had gone, it would have been good. (past counterfactual)*

[2]A fuller description of the use of the simple past in conditionals without a past time reference would make a distinction beteen hypothetical and counterfactual meanings of simple past verb forms (Celce-Murcia et al., 1999), or introduce a cline of hypothetical meanings from real to counterfactual (Comrie, 1986). However, for clarity, only conditionals with counterfactual meaning without reference to past time will be considered here.

**3.2.2.1.4 Cross-linguistic variation** In English conditionality is most commonly marked by the conjunction *if*, and in the Tense, Aspect and Mood (TAM) system of the verb phrases, as in the examples *if I were, I would be...* (present) and *if I had been, I would have been...* (past). Using such condition and counterfactual markers as a baseline in order to shed light on the possible interference between L1s and English, the following section details how the main L1s in this study typically encode conditionality and counterfactuality by grammatical and lexical means.

The main language, or language families, considered here are Arabic (n=28), Romance, including French, Italian and Spanish, (n=22), Indo-Aryan, including Gujarati, Bengali, Urdu and Hindi, (n=11), Vietnamese (n=9), Mandarin Chinese (n=4), Persian (n=6), Japanese (n=4), Russian (n=3) and Thai (n=3). The languages or language families fall into three broad groups, influenced although not exclusively by Comrie's cross-linguistic taxonomy of conditional constructions (Comrie, 1986). The first group is comprised of languages such as English and French that have a distinction in the TAM systems between factual and counterfactuals, and a distinction between a past and present reading. The second group mark counterfactuality and tense partly through the TAM system, and partly through additional particles. The third group rely on context and additional particles to retrieve counterfactual intentions, not predominantly on their TAM systems.

Considering the first grouping, there are similarities between romance languages and English in the use of TAM to signal counterfactuality and pastness. In French, for example, present counterfactuality is signalled with the simple imperfective past, and past counterfactuality with the use of a past imperfective auxiliary and a past participle (Examples 55 and 56), while the result clauses take the conditional mood in the present and the conditional mood plus the auxiliary in the past.

55. *Si les blancs jouaient*$^{[ipf,past]}$ *simplement au premier coup* $<$ $Cc3 >$ *suivi de* $< Cb1 >$, *les blancs se retrouveraient*$^{[conditionnel]}$

*alors dans la position initiale.*
If white in the first move simply played Nc3 followed by Nb1, then white would be back in the initial position.

56. *Si les noirs avaient*$^{[aux,ipf,past]}$ *joué 288, les blancs auraient*$^{[aux,conditionnel]}$ *gagné en jouant 29. Ce7.*
If black had played 288, white would have won by playing 29. Ne7.

(Examples from Grønn, 2008, p.9)

Languages in the second grouping mark conditionality and past counterfactuality within the TAM system and with the use of particles. In Arabic, the result and condition clauses of counterfactual conditionals are formed by the perfective aspect, with the TAM system making no temporal distinction between a past and present interpretation. Unlike Romance languages and English, however, Arabic employs conditionality markers depending on whether a counterfactual or non-counterfactual reading is intended. The markers *in (if)* or *idhā (whenever)* are used for for factual conditionals, and the conjunction *lau (if)* for counterfactuals (Lardiere, 1992). Lardiere (1992) notes that counterfactuality can be more explicitly marked by employing additional lexical markers such as *as if* or *as though.*

57. *in kataba waṣala al-maktūibu al-yauma*
If he writes, the letter will arrive today.
58. *idha waṣala fa-adkhil-hu*
When (whenever, if) he arrives, show him in!
59. *lau lam yaktub la-katabtu ana nafsī*
If he hadn't written, I would have written myself.

(Examples from Lardiere, 1992, p.241)

A similar distinction between the marking of non-counterfactual conditionals and counterfactual conditions is made in Persian, with the conjunctions *har vaqt (ke) (if ever), har mouqe (ke) (if and when)* and *agar* (if) used to signal factual conditionals, while the signalling of counterfactual conditionality is limited to *agar* (Mace, 2002, p.162). In both condition and result clauses the verb takes the perfective aspect, using context to infer the time frame (Example 61

60. *har vaqt (ke) u rā dідid xabar bedehid.*
    If ever/If and when you see him, tell (me).
61. *agar ān rā mi dānestam.*
    If I knew that I would tell you. OR
    If I had known that I would have told you.

(Examples from Mace, 2002, p.162)

As can be seen in the indeterminate readings in Example 63, in Russian there is also no tense distinction between present and past counterfactuality (Comrie, 1986). Counterfactuality in this case is marked by adding the particle *by* to the conjunction *esli*.

62. *Esli oni najdut (Future) étu vodu, oni budut (Future) spaseny.*
    If they find this water, they will be saved.
63. *Esli by oni našli (Past) étu vodu, oni byli (Past) by spaseny.*
    If they found that water, they would be saved.
    OR
    If they had found (find) that water, they would have been saved.

(Examples from Wierzbicka, 1997, p.38)

Japanese is another example of counterfactuality being marked not within the TAM system but with additional particles, albeit in a more complex way to Russian. Wierzbicka (1997) describes the optional use of the conjunction *moshi (if)* to signal conditionality and the suffix *ba* to signal counterfactuality, as in Example 64.

64. *(Moshi) karera ga ano mizu o mitsukete (karera wa) ima-goro shinde iru daroo. inakattara (ba).*
    If they hadn't found that water, they would have died.

(Example from Wierzbicka, 1997, p.45)

Hasada (1997), however, argues that the conjunction *moshi (if)* and the suffix *-ba* may lead to a indeterminate reading in terms of factuality. For a counterfactual reading, the particle *ni* is used as 'it invariably implies that the presupposed statement in the conditional clause is contrary to fact' (1997,

p.283). The present hypothetical and past counterfactual readings can be seen in Examples 65 and 66.

65. *(Moshi[mo]) John ga sono paatii ni ika-nakat-ta no deare-ba, Mary wa kare ni semaru chansu ga nakat-ta daroo (-ke(re)do(mo)/-ga).*
If John did not go (had not gone) to the party, Mary would not have had a chance to approach him.

66. *(Moshi[mo]) John ga sono paatii ni ika-nakat-ta no deare-ba Mary. wa kare ni semaru chansu ga nakat-ta noni/daroo-ni*
If John had not gone to the party, Mary would not have had a chance to approach him.

(Examples from Hasada, 1997, p.283-284)

In his treatment of counterfactual conditionals, Bhatt (1997) argues that generalisations can be made across Indo-Aryan languages such as Hindi, Gujarati, Bengali and Urdu. In Hindi, counterfactuals are signalled with the conjunction *agar (if)*, in addition to the atypical use of perfective, habitual or progressive aspects (in the sense that the past tense/pluperfect in English counterfactuals is a 'fake' past). In Example 67 the perfective aspect allows for an interpretation of both counterfactuality and pastness, a system parallel to the pluperfect in English past counterfactuals (Bhatt, 1997, p.11). Habitual and progressive aspects, seen in Examples 68 and 69 respectively, also signal a counterfactual interpretation, although the time reference can be interpreted as present or past, dependent on contextual factors[3].

67. *agar Ram-ne phal khaa-yaa ho-taa*
*if Ram-Erg fruit ate-Pfv be-hab*
If Ram had eaten the fruit,... (Perfective Contrafactive PSTCF?)[4]

68. *agar Ram phal khaa-taa ho-taa*
*if Ram fruit ate-hab be-hab PRS/PST*
If Ram had been eating fruit habitually,... (Habitual Contrafactive)

[3]Bhatt (1997) offers no alternative reading in Examples 68 and 69, and they are not added here.

[4]Bhatt notes that the clauses in counterfactual conditionals in Indo-Aryan languages are parallel, in the sense that the same morphology occurs in the result and condition clauses.

69. *agar Ram phal khaa rahaa ho-taa*
*if Ram fruit ate PROG be-hab PRS/PST*
If Ram had been eating the fruit,... (Progressive Contrafactive)

The third group consists of isolating syllabic languages with very little morphology to mark grammatical meanings (Bystrov and Stankevič, 2005). Mandarin Chinese, for example, relies heavily on pragmatic interpretation to infer past counterfactuality. Conditionality may be expressed with the particles *rúguǒ, jiǎrú, jiǎshǐ* or *yàoshi*, but as can be seen in Examples 70 and 71 it is often optional, relying on contextual factors for interpretation (Li and Thompson, 1989).

70. *jiǎrú xià yǔ wǒmen jiu zài wūlǐ chī fàn*
*if descend rain we then at outdoors eat food*
If it rains, we'll eat outdoors

71. *bàba qù wǒ gēn tā qù*
*father go I with 3sg go*
If father goes, I'll go with him

(Examples from Li and Thompson, 1989, p.632-3)

In Mandarin Chinese there is also no formal distinction between real, hypothetical and counterfactual conditionals, leaving any distinction to pragmatics. For instance, the conditional *Zhāngsān hē-le jǐu, wǒ jìu mà tā* can be interpreted as *If Zhangsan has drunk wine, I'll scold him, If Zhangsan drank wine, I would scold him* and *If Zhangsan had drunk wine, I would have scolded him* (Comrie, 1986, p.91), depending on the context.

Vietnamese and Thai also rely heavily on context to infer counterfactual intentions, as this is not encoded by the TAM systems. However, in Vietnamese the markers *nếu (nếu như, nếu mà)*, *if, if.would* or *gía (gía như, gía mà, gía sử)*, *if.would* (Bystrov and Stankevič, 2005, p.499) are employed to distinguish between factual and counterfactual/hypothetical, while in Thai the markers *thâa* and *hàak* signal a hypothetical/counterfactual interpretation (Nakagawa, 2005).

72. Vietnamese: *nếu không có Phát giúp, tôi sẽ làm con số bị loại*
if.would not there.is NAME help I FUT be number undergo expel
If Phat had not helped, I would have been among those expelled

(Example from Bystrov and Stankevič, 2005, p.506)

73. Thai: *thâa/hàak chăn kɔ̀ɔt nay sà-măy ee-dôo chăn àat cá mii oo-kàat phóp kàp too-kù-ŋaa-wâa ii-ee-yaa-súi kɔ̂ɔ dây.*
If I were born [*sic*] in the Edo era, I might have been able to see Ieyasu Tokugawa.

(Example from Nakagawa, 2005, p.47)

While this discussion on cross-linguistic variation has sacrificed depth of analysis for a broad view across the relevant L1s, it does allow for a discussion on the possible issues that may face learners who encode counterfactuality differently in their L1s in terms of particles that signal conditionality and the TAM system that signals counterfactuality and tense. The first issue, and arguably the easier of the two, is the use of particles to signal conditionality; while English commonly signals conditionality with *if*, with the optional particle *then*, languages such as Arabic, Persian and Russian make a bipartite distinction between factual and counterfactual particles. Japanese would seem a extreme case of this with the particles *moshi, -ba* and *-ni* signalling both conditionality and counterfactuality. At the other extreme, an explicit particle is optional in Mandarin Chinese, leaving pragmatics to infer the intended meaning. As English requires *if* and makes no distinction between *if* used in counterfactuals and *if* used in factuals, at the very least L2 learners of English need to determine that firstly English signals counterfactuality and tense in the verbal system, not in a distinction between particles, and secondly that omission of the *if* particle can lead to a non-conditional reading, even in a situation when pragmatics would suggest that a counterfactual situation is likely.

A related issue is that of cross-linguistic variation in the TAM system. Learners who have as their L1 languages which 'neutralise' (Comrie, 1986, p.94)

tense distinctions in past counterfactuals, such as Russian, will need to make the distinction between the bipartite tense distinction in the English TAM system that allows for a past or present counterfactual reading. For languages such as Mandarin Chinese, no distinction is made in the TAM system between present and past or factual and counterfactual, leaving the interpretation mainly to context. A Chinese learner learning English therefore would need to make a two-way distinction in the English TAM system between morphology that signal counterfactuality and morphology that signals tense.

The grouping of language and language families based on the method they use to encode past counterfactuality, and the possible issues that learners may have on the development of past counterfactuals in English based on their L1s is summarised in Table 3.3.

**Table 3.3:** *Summary of cross-linguistic differences in past-couterfactuals with issues for SLA*

| Group | L1s | Counterfactuality and Time | Possible issues for SLA |
|---|---|---|---|
| 1 | English<br>Romance | TAM system determines counterfactuality and time | Limited differences between English and Romance in terms of signalling conditionality and the encoding of time and counterfactuality in the TAM system |
| 2 | Arabic<br>Persian<br>Indo-Aryan<br>Japanese<br>Russian | Both TAM system and particles combine to achieve counterfactuality and time frame | Distinction between present and past time frames within TAM system; no distinction between hypothetical/counterfactual particles in English |
| 3 | Vietnamese<br>Mandarin<br>Thai | Particles and context determine counterfactuality and time frame | Distinction between present and past time frames within TAM system to signal both past time and counterfactuality; no distinction between hypothetical/counterfactual particles in English |

#### 3.2.2.2 Task materials

The three rounds of tasks used in the intervention are illustrated in Figure 3.2. All tasks were identical apart from the content of the input that learners were exposed to. In the first round of tasks, the Analogous Group were required to plan the 'Money and the bank manager task', and were subsequently provided with a performance of the 'Last night' task. This input was seeded with past counterfactuals, with lexical items that were unrelated to the 'Money and the bank manager task'. The Literal Group were also required to plan the 'Money and the bank manager task', but this time were provided with a performance of the same task. The control group planned the same task but received no input. All groups subsequently wrote a performance of the task.

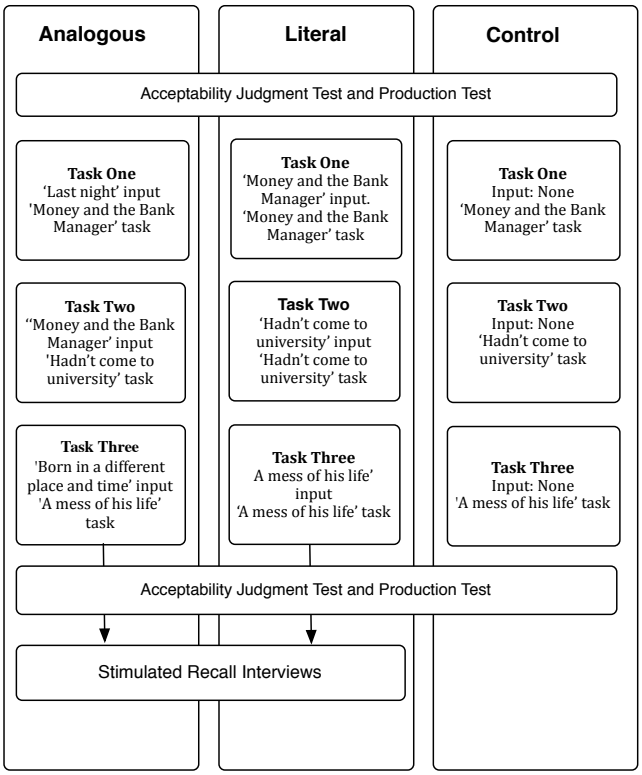
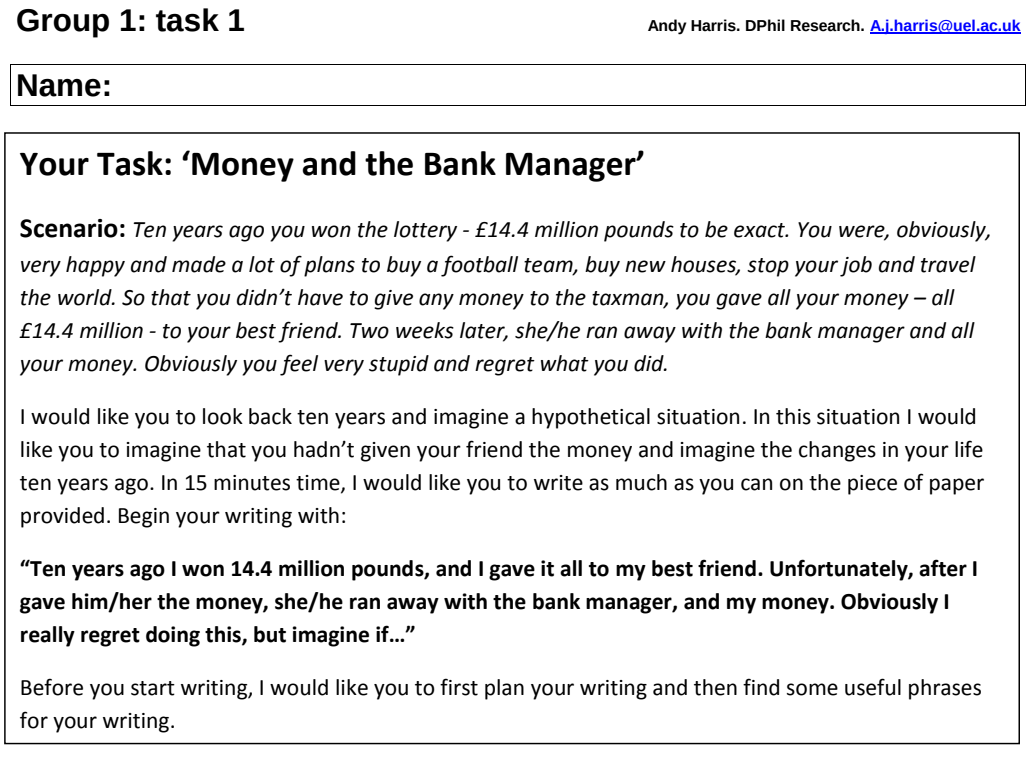



**Figure 3.2:** *Summary of research design with task types*

The sections of the tasks are described in more detail below, while Appendices F, G and H illustrate in full materials for Task One for the Analogous,

Literal and Control groups.

**3.2.2.2.1 Establishing the task demands** The first section of the intervention materials provided the rubric for the task (Figure 3.3). It was designed in such a way as to establish both a *base* (the factual situation) and a counterfactual space (the counterfactual situation) by employing space builders such as *imagine, hypothetical and regret* (Fauconnier, 1997). After initial piloting, it was decided that participants needed more guidance in the production task, and therefore they were also provided with the first sentence of the narrative, which reinforced the base and guided them further towards a counterfactual conceptualisation with the repetition of *regret, imagine, and if*. This assistance was hypothesised to be enough to guide the participants, but not so much as to render the task a form-focused language exercise (Ellis et al., 2009).



**Group 1: task 1**

Andy Harris. DPhil Research. A.j.harris@uel.ac.uk

**Name:**

**Your Task: 'Money and the Bank Manager'**

**Scenario:** *Ten years ago you won the lottery - £14.4 million pounds to be exact. You were, obviously, very happy and made a lot of plans to buy a football team, buy new houses, stop your job and travel the world. So that you didn't have to give any money to the taxman, you gave all your money – all £14.4 million - to your best friend. Two weeks later, she/he ran away with the bank manager and all your money. Obviously you feel very stupid and regret what you did.*

I would like you to look back ten years and imagine a hypothetical situation. In this situation I would like you to imagine that you hadn't given your friend the money and imagine the changes in your life ten years ago. In 15 minutes time, I would like you to write as much as you can on the piece of paper provided. Begin your writing with:

**"Ten years ago I won 14.4 million pounds, and I gave it all to my best friend. Unfortunately, after I gave him/her the money, she/he ran away with the bank manager, and my money. Obviously I really regret doing this, but imagine if…"**

Before you start writing, I would like you to first plan your writing and then find some useful phrases for your writing.

**Figure 3.3:** *Task rubric for Analogous, Literal and Control groups (Task One)*

**3.2.2.2.2 Analogous and literal input** This next section of the materials provided task-relevant input for the Analogous and Literal groups. For the Analogous Group, the input was based on regrets in the past and was therefore functionally and grammatically similar to the learners' task, but consisted of lexical items unrelated to their task demands. For example, in Figure 3.4, the input relates to a man who regrets going out the previous night, while the learners' task focuses on regrets related to giving money away (Figure 3.3). The input for The Literal Group was matched to the demands of the task in terms of function, grammatical frame *and* lexical items, i.e. the learners' task focuses on regrets about giving money away (Figure 3.3), and the input also focused on regrets about giving money away (Figure 3.5). The Control Group received no input.

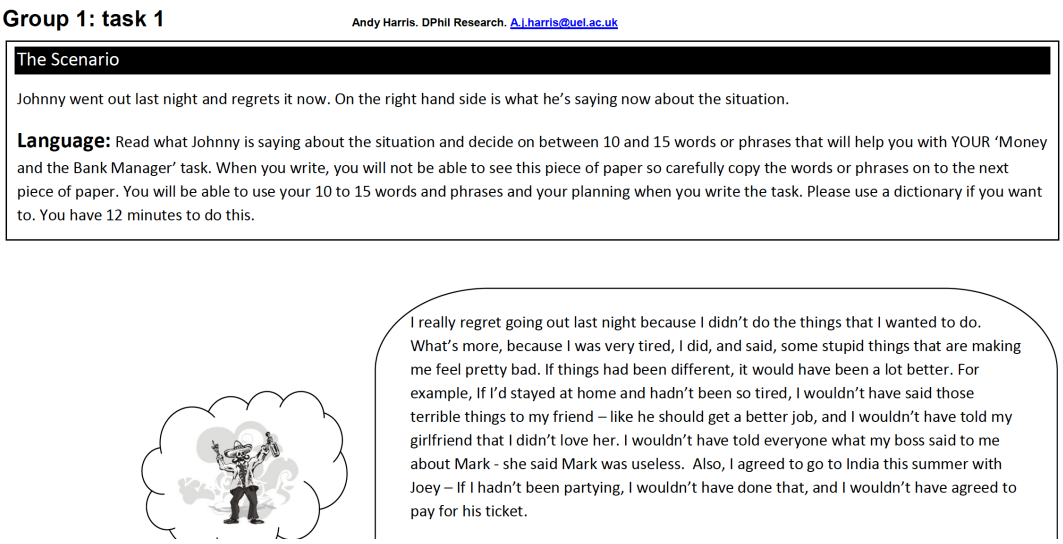



**Figure 3.4:** *Task rubric with a sample of analogous pre-task input*

As summarised for all tasks in Table 3.4, the input for both the Analogous and Literal groups was flooded with counterfactuals to allow participants ample opportunity to map function to form. The total percentage of conditional words (operationalised as *if+had+pp +would /could /might+have +pp*) was between twenty-two and twenty-nine per cent of the total input. There were approximately twice as many result clauses as condition clauses since context

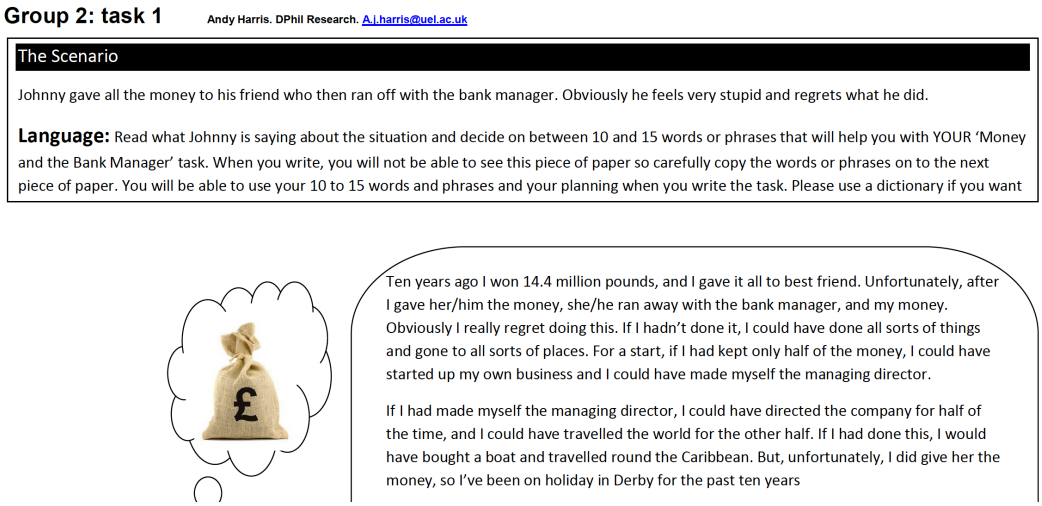



**Figure 3.5:** *Task rubric for Group Two with a sample of literal input.*

allowed the condition to be assumed.

**Table 3.4:** *Number and percentage of seeded past counterfactuals in the input.*

| Input | Number of words | Condition clauses | Result clauses | Conditional words (%) |
|---|---|---|---|---|
| Last Night | 289 | 7(21 words) | 14(42 words) | 22 |
| Money and the Bank Manager | 297 | 7(21 words) | 15(45 words) | 22 |
| No University | 322 | 8(24 words) | 21(63 words) | 27 |
| Different World | 262 | 9(27 words) | 16(48 words) | 29 |
| Mess of my Life | 278 | 7(21 words) | 17(51 words) | 26 |

**3.2.2.2.3 Task production** The final section of the materials for all groups consisted of a blank page to complete the task. The first line of the narrative was repeated to remind and assist the participants.

### 3.2.3 Testing instruments

#### 3.2.3.1 Operationalisation of acquisition

The operationalisation of acquisition in this study is dependent on the theoretical underpinnings outlined in the literature review. To illustrate this dependence, Norris and Ortega (2003) cite a study by Trahey and White (1993), which was carried out within a UG theoretical framework. In the study, French learners were exposed to input flooded with preverbal adverbs (SAVO), and in post-treatment tests, the learners accepted SAVO, but did not reject post-verbal adverbs (*SVAO). Trahey and White interpret this as an absence of acquisition, indicating that the relevant parameter, which controls adverb placement, had not been reset to allow SAVO and reject *SVAO. However, within a constructionist framework the acceptance of SAVO would be seen as evidence of embryonic or 'incipient' acquisition (Norris and Ortega, 2003, p.728). The Literature Review suggested that the trajectory for the acquisition of symbolic units of variable abstraction was from formula to low-scope patterns to full productivity. With this in mind, evidence of acquisition here is considered as piecemeal, non-categorical changes in interlanguage in the direction of the linguistic target, without necessarily being target-like. This can be seen in the manufactured examples in 74, 75 and 76.

74. Target-Like Utterance
    *If it hadn't been for the weather, it would've been good.*
75. Utterance showing signs of pastness and counterfactuality
    *If she hadn't took the position, she could earned so much money.*
76. Utterance showing no signs of pastness
    *They would look better if they dress better.*

#### 3.2.3.2 Test materials

Two tests were administered pre- and post-intervention to measure acquisition. The first test was an acceptability judgment test (Appendix E), which was designed to test for implicit knowledge of the structure. The second was a

production test (Appendix D) designed in order to measure use of the structure in obligatory contexts. A delayed post-test was not possible due to students leaving the university at the end of their pre-sessional course. The rationales and designs of the tests are detailed below.

**3.2.3.2.1 Acceptability judgment test** Acceptability judgment tests probe the extent to which structures are acceptable to learners, in terms of their grammaticality. These tests are extensively used both in linguistics and in applied linguistics, but they have been challenged in terms of their construct validity (Loewen, 2009; Sorace, 1996). Sorace, for example, comments on the indeterminacy of L2 grammars, describing them as 'unstable, transitional states of knowledge' (p.385). This unstable indeterminacy, Sorace argues, is more acute in the periphery than the core (from a UG perspective). As previous chapters argue for a constructivist and usage-based approach, it would seem valid to assume that there is considerable instability in a learner's interlanguage. For example, a learner may have stable control over Example 77, but reject Example 78, and Example 79, as inaccurate, as Example 77 could operate as a very limited type, and whose scope does not extend to Examples 78 and 79. Sorace suggests that to capture this indeterminacy more accurately, degrees of acceptability, instead of categorical decisions, may be more effective.

77. *If it hadn't been for the weather, it would've been good.*
78. *If she hadn't taken the position, she could not have earned so much money.*
79. *They might have looked better if they had dressed up.*

A second potential danger with acceptability judgment tests is their ability to tap both explicit and implicit knowledge (Ellis et al., 2009, *passim*, and specifically Loewen (2009)). In the case of an untimed acceptability judgment test, it is assumed that participants are free to tap both types of knowledge. If the timing is limited, however, it is suggested that there is insufficient time to tap explicit knowledge.

With both these issues in mind, the acceptability judgment test was designed with a six-point cline to gauge level of acceptability of the sentence (Figure 3.6, and in full in Appendix E). To reduce the possibility of learners drawing on explicit knowledge, the test was timed, allowing nine seconds for each sentence. Loewen (2009), in a review of the literature, gives a range of between 3 and 10 seconds. The test here opted for 9 seconds due to the multiple clauses and the length of sentence, compared with other studies, as well as the difficulty in completing the test in piloting when allocated five seconds per sentence. Nine seconds is also in line with other tests on past counterfactuals (Izumi et al., 1999; Song and Suh, 2008).

| | How good does the sentence feel? | | | | | |
|---|---|---|---|---|---|---|
| Sentence Number | Feels very wrong | Feels wrong | Feel a little wrong | Feels OK | Feels right | Feels very right |
| Sentence 1. | | | | | | |
| Sentence 2. | | | | | | |
| Sentence 3. | | | | | | |
| Sentence 4. | | | | | | |

**Figure 3.6:** *A six-point cline of acceptability in the acceptability judgment test*

The test consisted of forty sentences, twenty of which were distracters, and twenty of which were past counterfactuals (See Appendix E for the full list of sentences). Out of the twenty counterfactuals, ten started with the result clause and ten with the subordinate condition clause. Half of both the distracters and the counterfactuals were produced by a standard grammar (e.g. *I would have gone*), while the other half were not (e.g. *I would have go*). Of the non-standard sentences, the subordinate clause contained approximately half of the total number of errors, and the main clause approximately half.

The order of the sentences was randomly generated using an Excel-based random number generating function. All sentences were written in PowerPoint 2008 for Mac, with one sentence per slide, and a nine-second pause between sentences. The PowerPoint file was then converted to Adobe Captivate 4 for

audio encoding and conversion to a browser-compatible .swf file. When the file was played, participants saw the sentences on screen and also heard them.

**3.2.3.2.2 The production test** Although a number of mechanisms were put in place in an attempt to ensure the acceptability judgment test tapped implicit knowledge, one test alone may not have fully tapped the learners' knowledge of the target (Ellis and Barkhuizen, 2005). For this reason, an additional test of production was utilised.

Ellis and Barkhuizen (2005) subdivide methods for the collection of production data from L2 learners into three sampling methods: naturally-occurring, clinically-elicited and experimentally-elicited (Figure 3.7). Naturally occurring data is described as data collected in a real-life situation with no control over a participant's production. Both clinically elicited and experimentally elicited data, on the other hand, result from some form of control over the elicitation. The difference between the two is whether the main emphasis is on the form of the language, or on its meaning.

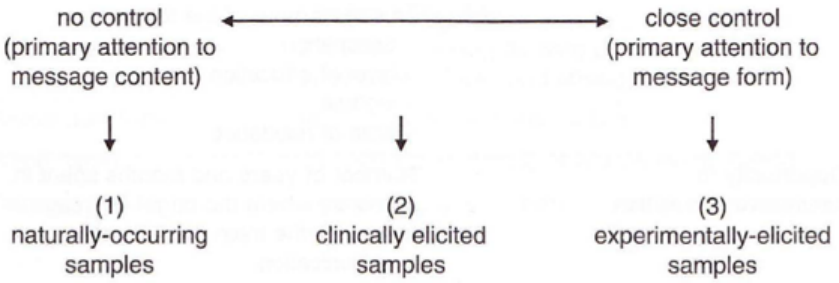
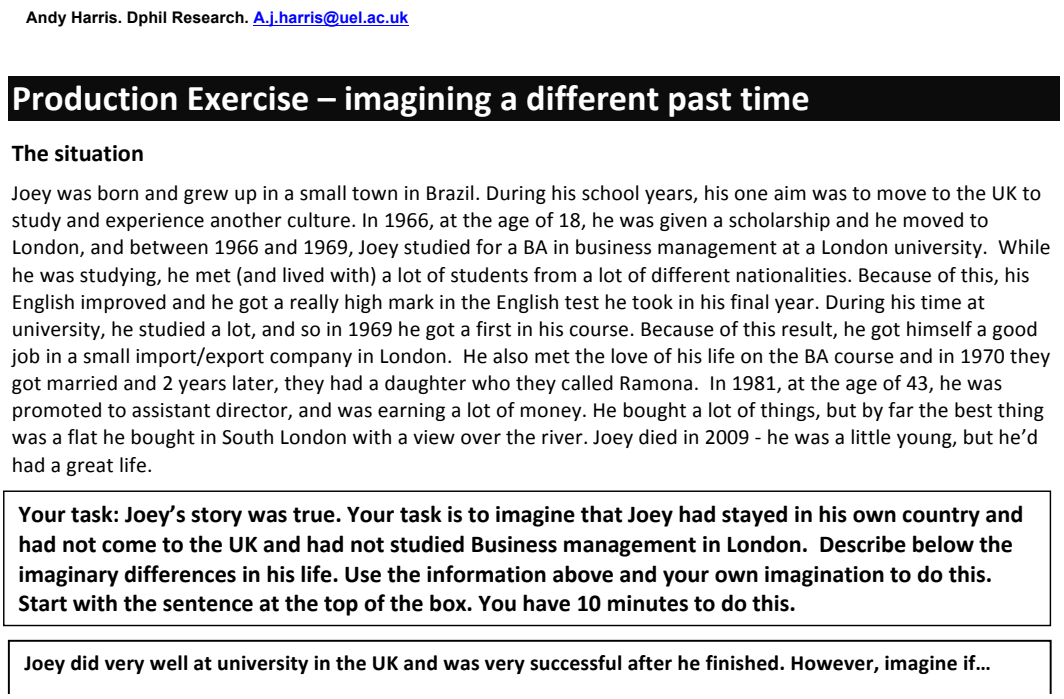



**Figure 3.7:** *A cline between naturally-occurring and experimentally-elicited samples (Ellis and Barkhuizen, 2005)*

As the main orientation of the current thesis is on the conceptualisation of past counterfactuality and the selection and use of tokens, a suitable data sampling method would need to emphasise meaning, which rules out a heavily form-focused method. As naturally occurring data collection is impractical in this study, a focused clinically elicited method was employed. As illustrated

in Figure 3.8, the rubric of the production test established a context in the form of a situation, thus providing a base on which to interpret the task. The task itself provided a communicative purpose, which necessitated the use of past counterfactuals for successful task completion. This was achieved by employing as many *space builders* (Fauconnier, 1997) as possible, such as *imagine, if*, as well as past tense references.



Andy Harris. Dphil Research. A.j.harris@uel.ac.uk

**Production Exercise – imagining a different past time**

**The situation**

Joey was born and grew up in a small town in Brazil. During his school years, his one aim was to move to the UK to study and experience another culture. In 1966, at the age of 18, he was given a scholarship and he moved to London, and between 1966 and 1969, Joey studied for a BA in business management at a London university. While he was studying, he met (and lived with) a lot of students from a lot of different nationalities. Because of this, his English improved and he got a really high mark in the English test he took in his final year. During his time at university, he studied a lot, and so in 1969 he got a first in his course. Because of this result, he got himself a good job in a small import/export company in London. He also met the love of his life on the BA course and in 1970 they got married and 2 years later, they had a daughter who they called Ramona. In 1981, at the age of 43, he was promoted to assistant director, and was earning a lot of money. He bought a lot of things, but by far the best thing was a flat he bought in South London with a view over the river. Joey died in 2009 - he was a little young, but he'd had a great life.

**Your task: Joey's story was true. Your task is to imagine that Joey had stayed in his own country and had not come to the UK and had not studied Business management in London. Describe below the imaginary differences in his life. Use the information above and your own imagination to do this. Start with the sentence at the top of the box. You have 10 minutes to do this.**

**Joey did very well at university in the UK and was very successful after he finished. However, imagine if…**

**Figure 3.8:** *Production test rubric*

### 3.2.4 Procedure

The intervention, including pre- and post-tests, was spread over five days. All participants, in class groups, first took part in the two pre-tests (Appendix D and Appendix E). Participants were tested by the researcher in groups using a pencil and paper test. The first test was the production test, as it was considered that this would not pollute the acceptability judgment test, as the production test, unlike the acceptability judgment test, did not provide tokens

of the past counterfactual. Key phrases in the rubric of the production test were highlighted by the researcher and subsequently participants were given two minutes to read the text in more detail and ask questions. Ten minutes was allowed for the completion of the production test, as this amount of time proved optimum in piloting. Following the production test, the procedure for the acceptability judgment test was explained and demonstrated, and four unrelated test sentences were used to practice the procedure. The total amount of time for participants to complete both tests was forty minutes. The post-tests used the same task types followed by exactly the same procedure as the pre-test, apart from a random reordering of the 40 sentences on the acceptability judgment test to lessen any practice effects from the pre-test.

The tasks themselves were spread over three days. Each task lasted approximately forty minutes and was administered in the participants' class groups by their regular teacher. All teachers in the study were trained beforehand on how to set up the tasks, told what was permissible in the session, and how to answer possible questions. A summary of the stages of the tasks is given in Figure 3.9. After reading the rubric to their task, participants were given ten minutes planning time to reread the rubric, conceptualise the task and make notes relevant to their production. Ten minutes is often considered the optimal time for planning in task performance (Crookes, 1989; Foster and Skehan, 1996; Skehan and Foster, 1997, see Mehnert (1998) for the effects of varying planning time). In line with Crookes (1989) and Foster and Skehan (1999), the participants were allowed to take notes and consult dictionaries during the planning phase, and were allowed to retain their notes for the task performance. Although they were given significant context, they were not given direction in how to plan for the task.

After the 10 minutes planning time, the Analogous and Literal Groups were asked to underline areas in the input that would be useful to them in their own production and were then asked to transfer their choices to their production sheet, so that they could use the choices in their task production (see

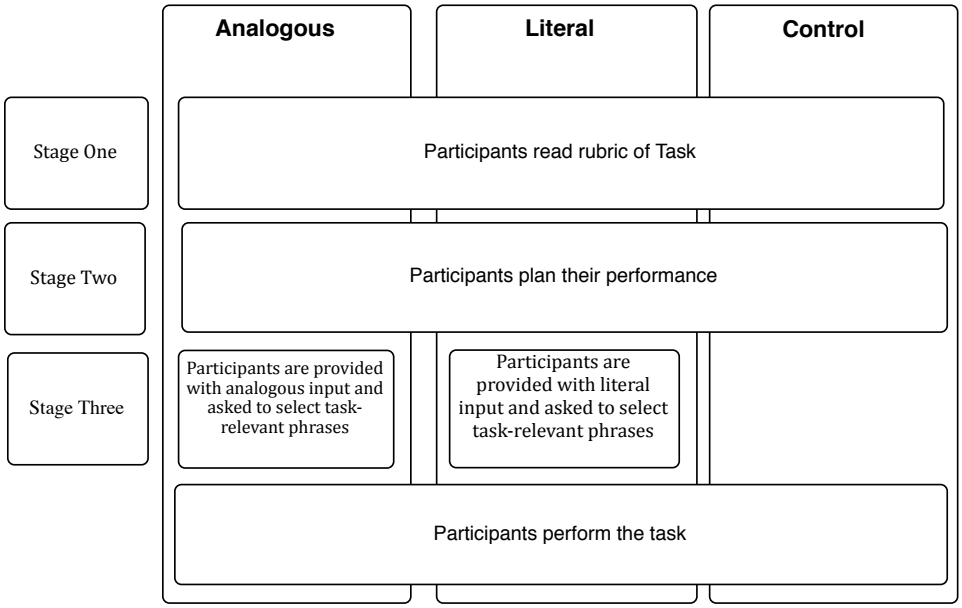



**Figure 3.9:** *Task procedure*

Appendices H and G for examples). Both groups were given the same time to complete the exercise (7 minutes), again based on the piloting of the materials. The Control Group did not receive input, so moved directly to the production stage. Often when a control is used an alternative task is set in order to reduce the number of intervening variables, i.e. here the Control Group would have completed an unrelated exercise while the Literal and Analogous Groups selected task-relevant language. However, when an additional task was added in piloting, it appeared to distract learners from the main task by confusing the purpose of the final task production. For this reason, an additional task was not employed.

All groups were given ten minutes to complete their written tasks. They were told that they could consult dictionaries, and ask questions, although no explicit information was given on past counterfactuals. At the end of the ten minutes, all participants handed in all tasks and notes.

Twenty-four hours after the final task, those students who had volunteered took part in the stimulated recall interviews. These interviews took place in the researcher's office on a one-to-one basis. Each interview lasted approximately

five minutes. See 4.2.2: *Stimulated Recall Interviews* for a description and coding scheme for the interviews.

## 3.3 Data scoring

This section first details how the performances on both the acceptability judgment test and the production test were coded and scored. The section then considers the scoring of the participants' selections of past counterfactual tokens from the input and the repurposing of selections for use in task performances. The final part of this section presents the coding of the stimulated recall interviews.

### 3.3.1 Scoring the tests

#### 3.3.1.1 The acceptability judgment test

The first step in scoring the acceptability judgment test was to remove the distractors from further analysis. The remaining twenty sentences were then subjected to analysis using a variable scale based on how confident participants were of their decision (see Figure 3.6 on page 102, along with the scoring in Table 3.5). The scoring took into account the accuracy of the judgment and the confidence with which it was reported. This method was designed to be more sensitive to variation in the learners' interlanguage than a binary score based on either targetlike or nontargetlike production.

**Table 3.5:** *Scoring of acceptability judgment test*

| Score | Description (unacceptable Sentences) | Description (acceptable Sentences) |
|---|---|---|
| 5 | Feels very wrong when unacceptable. | Feels very right when acceptable |
| 4 | Feels wrong when unacceptable. | Feels right when accurate |
| 3 | Feels a little wrong when unacceptable. | Feels a little right when acceptable |
| 2 | Feels a little wrong when acceptable. | Feels a little right when unacceptable |
| 1 | Feels wrong when acceptable. | Feels right when unacceptable |
| 0 | Feels very wrong when acceptable. | Feels very right when unacceptable |

#### 3.3.1.2 The production test

The analysis of the production test was more complex than that of the acceptability judgment test due to the difficulties of defining and operationalising an obligatory context (Bley-Vroman 1983; Lakshmanan and Selinker 2001; Lardiere 2003; Shirai 2007).

One issue with operationalising obligatory contexts, as Selinker and Lakshmann point out, is attempting to determine the intention of the learner when producing certain utterances, something they refer to as *interlanguage intention*. Identifying the intention is not an easy task, as a learner's interlanguage is often variable, and by definition nontargetlike. This means that non-standard forms could be mapped in a non-standard manner to functions. The present analysis utilised a number of tools to increase the likelihood that a certain utterance was produced to satisfy a certain underlying intention. The first of these was the space builders (Fauconnier, 1997) *Ten years ago*, and *Look back ten years*, which were used in the rubric to set a past time frame, as well as the space builders *imagine a hypothetical situation* and *Obviously I really regret doing this, but imagine if...* to set a counterfactual frame. Additionally, the tasks themselves, unless misinterpreted, did not allow for a conceptualisation in present or future time, as the protagonist in the task was dead, or a specific time in the past was being considered. Clearly, however, there existed ways to interpret the task which circumvented the set rubric, and therefore additional tools were necessary to attempt to establish intention. One of these was lexical reference to time frames. For example, in Example 80 below, the student uses *in my earlier life*, *in past* and *could have achieved* to locate the utterances in the past. This context can then be used to code *if there were...* and *If I had...* as obligatory contexts for past counterfactuals.

> 80. I could have achieved success if there were no mess in my early life... My life could be better if i had no criminal activities in past.

In a similar way to lexical past time reference, the immediate factual discourse often allowed for an interpretation of counterfactuality. In example 81, the phrase *contrary to London* set the frame of factuality (the student completed the task in London), and therefore, *would not have meet* can be coded as an obligatory context for past counterfactuality.

> 81. I would have travelled around the world, perhaps, but I would not have meet a lot of friend, contrary to London.

In cases when there did not exist lexical devices to signal counterfactuality or past time reference, and there was a possibility that past time reference was not obligatory, the sample was omitted from analysis due to its ambiguity. This ambiguity can be seen in Example 82, in which it is possible that the time reference is not the past, and that the past tense forms refer to solely counterfactuality.

> 82. If I won 14.4 million pounds I will be happy. I could buy new village in France. I could buy a new car. I will open my own company and travel the world.

Learner production was also coded as an obligatory context if the initial utterances framed the subsequent discourse, an important function of past counterfactual condition clauses (Biber et al., 1999, p837). For example, in 83, the first two clauses suggest a past time frame (with a non-standard form), setting an obligatory context for a past counterfactual in the second sentence. Similarly, the underlined sections in Example 84 were considered obligatory contexts as the first clauses again have scope over the subsequent discourse (note that the rubric states that the antagonist of the story is dead).

> 83. However, imagine if...
> ...he did not come to london, he would stay in his own town. <u>There was a posibility to intended to football clubs and trained very well with his team</u>...likely this man would be a profissional football player and also might be become famouse in that time...

84. However, imagine if...
    ...If Joey had not been given a scholarship, his life would be continue in Brazil and be successful in his profession. In 1966, after failure of scholarship, Joey decided study BA in business management in the capital of Brazil. By his effort, he got the first class honour then many big company invited him working for...

The final case was that of complete avoidance. In a number of learner productions, the past counterfactual space builders, as well as the situation did not lead the learner to show any signs of past counterfactual condition or result clauses (e.g. in Example 85).

85. However, imagine if...
    ...Joey didn't go to the UK. What's his life? I try to explain it...After the school in his small village he went to Brazilia to study at University. During the university years...

In these cases, if there was no possibility of a reference to present or future time (e.g. because the antagonist was dead), it was assumed that the learner had no knowledge of past counterfactuals, and therefore to complete the task, a past or present tense was used. These were counted as obligatory contexts. Verbless clauses and elided structures were omitted due to their ambiguity.

The second stage of the analysis was to divide the obligatory contexts into condition and result contexts, and determine whether the form used to encode the intention was a Target-Like-Utterance (TLU). The analysis of the obligatory contexts and condition and result clauses was based on Pica's (1983) formula (Figure 3.10). The main difference between Pica's formula for TLUs and the one used here was the incorporation of '$n$ suppliance in non-obligatory contexts'. Although this was essential in Pica's work on morpheme acquisition, the use of past counterfactual conditions out of context was considered unlikely. Analysis of the pre- and post-test scripts confirmed this.

While the analysis of TLUs gives one indication of performance, in second language acquisition, its use has been criticised as it does not take into account

$$TLU = \frac{(n\ suppliance\ in\ oligatory\ contexts)}{(n\ obligatory\ contexts)+(n\ suppliance\ in\ non-obligatory\ contexts)}$$

**Figure 3.10:** *Determining target-like utterances (Pica, 1983)*

the independent status of the interlanguage system (cf. Bley-Vroman's comparative fallacy, 1983). That is, if a learner's production is solely measured against the target system (e.g. TLU), interlanguage is framed not as a system in its own right with measurable systemic generalisations, but as an erroneous version of the target language. However, as Shirai (2007) points out, there is not a problem with the use of TLUs per se; the issue is that additional analysis is needed to determine not *solely* TLUs but also the status of a learner's interlanguage.

Two studies are of particular interest in terms of the scoring of interlanguage in the production test. The first is a task-based study that focused on pre-task explicit input in the form of relative clauses (Mochizuki and Ortega, 2008), while the second focused on past counterfactuals and the Output Hypothesis (Izumi et al., 1999). In Mochizuki and Ortega, a six-point interlanguage scoring was used that was based on a bottom-up analysis of errors. Participants scored the maximum for correctly supplying the form, less for erroneous forms, and finally zero for non-suppliance in an obligatory context. This, Mochizuki and Ortega argue, allows for an 'interlanguage-sensitive measurement of performance of relativization rather than an all-or-nothing binary accuracy scoring' (2008, p.21). Izumi et al. (1999) used a scoring system that focused on suppliance or non-suppliance of counterfactual words that form past counterfactuals (e.g. the suppliance of *would, have, had*). These studies provided a tentative a priori coding scheme for the coding of both targetlike utterances and Interlanguage, with the final coding system emerging from the investigation of the data.

Table 3.6 shows how the analysis was divided into two superordinate categories: the first was TLU in an obligatory context; the second broad category

**Table 3.6:** *Coding of production test for condition and result clauses (targetlike and interlanguage measures).*

| | Condition | Result | Score |
|---|---|---|---|
| TLU | **Cond.TLU**: condition clause, target-like utterance (e.g. if I had gone) | **Res.TLU**: result clause, target-like utterance (e.g. I would/could have gone) | 3 |
| IL | **Cond.signs of counter and past**: shows signs of both counterfactuality and pastness but is not target-like (e.g. if I had went, if I would have gone) | **Res.counterfactual modal with signs of pastness**: shows signs of both counterfactuality and pastness but is not target-like (e.g. I would went, I would had gone, I would gone) | 2 |
| IL | **Cond.signs of counter or past**: shows signs of either counterfactuality or pastness and is not target-like (e.g. if I was born, if I went) | **Res.counterfactual modal with no signs of pastness**: shows signs of counterfactuality but not pastness and is not target-like (e.g. I would go, I could have go) | 1 |
| IL | **Cond.no sign**: no sign of either pastness or counterfactuality (e.g. if I go) | **Res.no sign**: no sign of either pastness or counterfactuality (e.g. I go, I will go, I can go) | 0 |

was an interlanguage measure, which, based on the data, was broken down into three categories. In the result clauses, the first and second interlanguage measures focused on the use of a counterfactual modal verb, with or without signs of pastness. The final banding for the result clause covered utterances which showed no sign of past or a counterfactual modal. The interlanguage measures for the condition clauses were coded slightly differently due to the ambiguity of past time and counterfactual reference. This can be seen in the 'signs of counter OR past' category, which covered such examples as *If I were born*, for which the past tense form could be interpreted as either reference to past time or reference to counterfactuality without pastness. Following Mochizuki and Ortega (2008), TLUs were allocated three points, and interlanguage measures from zero to two. Spelling was ignored, unless it rendered the meaning grammatically ambiguous. In a similar way, subject-verb agreement was ignored,

unless it impacted on the analysis of the target structure.

### 3.3.2 Scoring selections from the input

In the pre-task input phase, participants were asked to select instances from the input vis-à-vis the task demands and copy them to their production sheet (Appendices F and G). The selections relating to past counterfactuals were coded in two ways: selection of the entire main clause or selection of the entire verb phrase (*Full Selection*) and selection of elements of the verb phrase (*Partial Selection*). Examples of the Full Selection of result clauses are shown in 86, 87 and 88.

86. I wouldn't have agreed to pay for his tickets
87. I wouldn't have spent it all
88. So I could have spent it on other things

Selection of the entire verb phrase from a result clause, e.g. with a missing subject or object or both, can be seen in Example 89, which corresponds to the phrase *I could have done many things* in the input to the task.

89. I could have done

The second category, Partial Selection, included the selection of words that formed past counterfactuals result clauses in the input, but not the full VP or the entire clause. This category covered items such as single words, e.g. *would*, *could* and past participles, two-word combinations within the VP, e.g. *would have*, *have had*, and multi-word chunks that omitted part of the VP, e.g. *I would have*. Selections of instances from the condition clauses in the input followed the same two-level categorisation of full and partial selection, albeit with differences in the exact details. In line with other studies that focused on the selection of past counterfactual instances from input (Izumi and Bigelow, 2000; Izumi et al., 1999; Song and Suh, 2008), the final score was obtained by

dividing the number of partial and full past counterfactual instances by the total number of selected items (Figure 3.11).

$$Score_{partial} = \frac{Number\ of\ partial\ selections_{condition\ or\ result}}{Total\ number\ of\ selections}$$

$$Score_{full} = \frac{Number\ of\ full\ selections_{condition\ or\ result}}{Total\ number\ of\ selections}$$

**Figure 3.11:** *Scoring the selections of past counterfactual condition and result clauses*

### 3.3.3 Scoring repurposing operations

This section considers the repurposing of the input to meet the task demands. Repurposing was operationalized a priori as the number of lexical changes that were made to instances of past counterfactuals selected from the input texts. On coding the data, a more fine-grained coding schema emerged, resulting in five categories for both result and condition clauses (see Figure 3.7). The first of these categories for the result clauses was the verbatim use of whole instances, including the subject, verb phrase and object (Example 90).

90. **Text from the Input**
What's more, because I was very tired, I did, and said, some stupid things that are making me feel pretty bad. If things had been different, it would have been a lot better.

**Output of Task Performance**
*. . . I really regret giving money all to my best friend. Because I have trusted her absolutely. Moreover, because I was very lazy to do everything. Therefore, I have done stupid things. If things had been restored, it would have been a lot better.*

The learner production was only coded in this way if it matched the input to that particular task; input in other tasks on different days was not taken into account. It could, of course, be argued that the learner already knew the phrase *It would have been a lot better*, and had either constructed it creatively

from scratch or had acquired it previously as a formulaic phrase. To guard against this, if the learner had showed targetlike use in the pre-test, their task production was not coded. This was the case for the coding of all result and condition clauses in task productions.

**Table 3.7:** *Coding for the use of selections from input in task performance*

| Category | Condition | Result |
|---|---|---|
| 1 | Use of input-COND-SVO* | Use of input-RES-SVO* |
| 2 | Use of input-COND-changed subject | Use of input-RES-changed subject |
| 3 | Use of input-COND-changed object | Use of input-RES-changed object |
| 4 | Use of input-COND-aux+changedPP (but maintains the subject and object from the input) | Use of input-RES-modal + aux + changedPP |
| 5 | - | Use of input-RES-change to modal + aux + InputPP |
| 6 | Use of input-COND-change to negative/positive aux + InputPP | Use of input-RES-modal + change to negative/positive aux + InputPP |

*COND=conditional clause, RES=result clause, COND-SVO=use of subject, verb (and object) from a selected condition clause, RES-SVO=use of the subject, verb (and object) from a selected result clause

The second and third categories described the verbatim use of the verb phrase with a change of subject or object respectively. Example 91 shows a change in the subject of the sentence (*it* is changed for *My life*). If a quantifier had been inserted or omitted as in the omission of *a lot* in Example 91, it was decided that the object had not been changed.

91. **Text from the Input**
What's more, because I was very tired, I did, and said, some stupid things that are making me feel pretty bad. If things had been different, it would have been a lot better.

**Output of Task Performance**
*. . . I had not made a real mess of my life, my life would have been better. If I had not stolen cars, I would not have been caught and I wouldn't have had a criminal record. . .*

The next three categories of changes in the result clauses relate to changes to the verb phrase, and are therefore either a change to the type of modal

(e.g. *could* to *would*), the polarity of the modal verb or a change to the past participle. Examples 92 and 93 show, in relation to the input, changes to the modal, and changes to the past participle respectively.

92. **Text from the Input**
... and so I could've spent it on other things, like some CDs, a new tent, or I could have gone out for a meal instead. Or, better, I could have gone ...

**Output of Task Performance**
*I would have also made my husband feel happy, we would have gone out for a meal*

93. **Text from the Input**
... I would have bought a boat and travelled round the Caribbean ... I would have bought a large house in Scotland ... I also would have invested some money ...

**Output of Task Performance**
*... didn't go to the UK, my english wouldn't improved, But I would have worked to earn money to go in a chinee's university where I could learn chinese and Business.*

### 3.3.4 Stimulated recall interviews

In addition to the tests and tasks, stimulated recall interviews were conducted twenty-four hours after the final task to tap into how participants processed the task demands and the selected instances of past counterfactuals.

To frame the stimulated recall procedure, Gass and Mackey's classification scheme was used (2000). This scheme, based on a similar scheme by Faerch and Kasper (1987), puts forwards six dimensions with which to classify stimulated recall sessions. The first of these - *relationship to specific action* - refers to how concrete the original event was. As Gass and Mackey note, if the process is non-specific/abstract, the participant may well be reaching for conclusions based on their introspection of specific events. This would, to a certain extent, distance the introspection from the specific event and maybe bring into question the validity of the conclusions (unless the target of the recall session

was to tap into learner beliefs). The study therefore used the materials that the participants had used in their tasks, including their planning, underlining of task-relevant input, task production and related rubrics, i.e. completed versions of the worksheets in Appendices F, G and H. Questions about their tasks were made as concrete as possible by referring to their individual conceptualisation of the task and subsequent selection and processing of instances in the input.

The second dimension - *temporal relationship to action* - reflects the ability to recall specific processes over different time periods. It seems reasonable to argue here that the longer the time period between the recall session and the event, the less accurate the recall of that event. It was decided that although immediately after each task would have been ideal, this would have biased results on subsequent tasks and the post-tests. Therefore the interviews took place twenty-four hours after the final task.

*Participant training* is the third dimension in the classification scheme. No training session was given to the learners. This decision was based on the lack of evidence that training sessions have a positive effect on the quality of recall (Gass and Mackey, 2000).

The fourth dimension - *procedural structure* - refers to the structure of the interview itself. A highly-structured interview would employ highly structured a priori categories, which would give little flexibility to the interview. It would, however, offer a specific direction for the study. This study falls towards the unstructured end of the cline, with only a small number of guiding questions that approximately followed the stages in the study (planning - selection - repurposing - meaning of selections - explicit knowledge). These questions can be seen in Table 3.8.

The fifth dimension refers to the stimuli(us) used in the sessions. As stated earlier, the stimuli were the participant's planning notes, their underlining of task-relevant input, task production and related rubrics (Appendices F, G and H).

**Table 3.8:** *A priori questions for stimulated recall interviews*

| | |
|---|---|
| **Planning** | What were you thinking about when you were planning your task?<br>What were you thinking about when you read the reading that I gave you after the planning? |
| **Selection and Repurposing** | Why did you choose this or that?<br>Did you change the language or use it exactly? Why?<br>How did you change it? |
| **Meaning** | Did you have the same technique in each task?<br>Do you know what the language means?<br>Do you know when to use it? |
| **Explicit Understanding** | Do you know the grammar?<br>Can you explain it to me?<br>Did you know the language before? Did you use it? What did you learn from the exercises? |

The final dimension - *the initiation of the recall questions* - refers to whether the learner or the researcher initiates the questions. Here, an attempt was made throughout the interviews to echo responses in order to elicit more comment, and lower the risk of suggesting ways to answer questions.

### 3.3.5 Coding the stimulated recall interviews

Participants who took part in the stimulated recall interviews came from both the Analogous and Literal Groups (n=27, 14 in the Analogous Group and 13 in the Literal Group). All interviews were transcribed and coded in Nvivo. During coding, three broad categories emerged:

94. Task demands and the selection of constructions from the input
95. Explicit reference to patterns in the input and learner production
96. Knowledge of the meaning of counterfactual constructions

These three themes are detailed below.

#### 3.3.5.1 Task demands and the selection of tokens from the input

This category refers to the effect that task demands had on the subsequent selection of past counterfactuals tokens from the input. For example, in Excerpt 97 below, the selection of past counterfactuals fulfilled a need that was created by the task demands.

97. Task demands and the selection of constructions from the input. AH=researcher.

| | |
|---|---|
| AH: | OK, OK, fine. And then I gave you a, some reading, on a piece of paper, and you chose this language down here (researcher points to the selections the student had made). |
| ST14: | Yeah. |
| AH: | Why [did you choose this]? Can you tell me a bit about that? |
| ST14: | Because I don't know, how can I describe? Of course I know in Japanese, but I don't know English how can I describe the past. Especially if I imagine something. |

This category also refers to how pre-task input affected the interpretation of the task demands. In Excerpt 98, ST4 changed her interpretation of the task demands because of her interpretation of the input.

98. Task demands and the selection of constructions from the input.

| | |
|---|---|
| AH: | OK, and if I ask you some questions about it, is it about the past, or the present, this grammar? |
| ST4: | About the past... |
| AH: | About the past. But your planning was all about the present. |
| ST4: | The present, yes. |
| AH: | So what happened there? What did you..? |
| ST4: | Um... just I look how to write correctly and in the story was past, and I decide to use past. |
| AH: | So you changed your plan to the past. |
| ST4: | Yeah. |

#### 3.3.5.2 Explicit reference to patterns in the input and learner production

This category refers to the explicit recognition that a pattern exists in the input or during production. To be included in this category, references had to be made to the abstract properties of the input or a reference to 'going beyond' the selection. In Excerpt 99, it can be seen that the learner makes direct reference to the pattern underlying the phrases and the use of her own words within that pattern.

99. Explicit reference to patterns in the input and learner production

| | |
|---|---|
| AH: | So to fit your plan. You had your plan, and you had the reading, and you put these together. |
| ST4: | Yeah, yeah. I... uh... mostly I use the pattern, the sentence, the pattern, after, the pattern of sentence, but my word. |
| AH: | But your words. |
| ST4: | Yeah, put my words in the patterns. |

#### 3.3.5.3 Knowledge of the meaning of counterfactual constructions

The third category related to whether and to what extent participants knew the meaning of the constructions that they had selected and were working with. For example, in Excerpt 100 below, ST5 seems to have grasped both the counterfactual and past nature of the tokens. However, in Excerpt 101, ST10 seems to have taken *have been* as an example of the present perfect tense, not as a past time reference.

100. Knowledge of the meaning of counterfactual constructions

| | |
|---|---|
| AH: | OK, and is this language (AH points to past counterfactual selections) about the past, or the future, or the..? |
| ST5: | It's about the past. |
| AH: | And is it about a real situation, or an unreal situation? |
| ST5: | It's unrealistic but... |
| AH: | Did it happen? That's what I want to say. |
| ST5: | No it didn't, it did not happen. |

101. Knowledge of the meaning of counterfactual constructions

ST10: And I used the past.
AH: And you used the past.
ST10: And present perfect.
AH: Present perfect.
ST10: Yeah.

### 3.3.6 Inter-rater reliability

An experienced EAP lecturer was employed to code 20% of the transcribed data for obligatory past counterfactual contexts and use of past counterfactuals in the production tests, repurposing operations on the input and the stimulated recall interviews.

**Table 3.9:** *Inter-rater reliability of coding schemes*

| Coding Scheme | Reliability (%) | Cohen's Kappa | .95 CI | |
|---|---|---|---|---|
| | | | lower | upper |
| Production Test Coding | | | | |
| Use of Counterfactuals | 82 | .79 | .71 | .87 |
| Obligatory contexts | 92 | .71 | .61 | .89 |
| Repurposing operations | 92 | .90 | .79 | .99 |
| Stimulated Recall Interviews | 76 | .72 | .59 | .85 |

The codes were initially developed by the author, and then discussed with the second rater. This led to mutually agreed changes to the initial codes. Following this, there was a period of training on a small section of data within each of the coding schemes. Finally, 20% of the data were coded. Table 3.9 shows the inter-rater reliability and Cohen's Kappa for the coding schemes. Landis and Koch (1977) provide benchmarks for interpreting Cohen's Kappa: above 0.81 shows 'almost perfect agreement', above 0.61, 'substantial agreement' and above 0.41, 'moderate agreement'. All values for Cohen's Kappa in this study fall in the 'substantial agreement' bracket, 0.61-0.8.

# Chapter 4

# Findings

This chapter is divided into three sections. The first section is quantitative and provides descriptive and inferential statistics related to the eight research questions and the three hypotheses. To illustrate and extend these general statistics, the second section is qualitative in nature and describes and analyses five case studies of individual task performances, while the third section documents the stimulated recall interviews.

## 4.1 Quantitative Analyses

### 4.1.1 Selection of tokens from the input

The research questions and the hypothesis related to partial and full selection of condition and result clauses by the Analogous and Literal groups are listed in Table 4.1. Tables 4.2 and 4.3 provide descriptive statistics for the selection from the input of partial and full condition and result clauses for the Analogous and the Literal groups. Figure 4.1 illustrates the mean percentage of selections[1] for both Groups (See Appendix I for raw counts).

Due to significantly non-normal Kolmogorov-Smirnov tests, Kruskal-Wallis tests were run on the data to determine if there was a significant difference in

[1]The selections were divided into two types: partial selection was the selection of conditional elements (e.g. *had, if, gone*), and full selection was the selection of the VP, e.g. *would have gone* or full selection of the complete result or condition clause, with corresponding subject and complements/objects.

**Table 4.1:** *Research questions and hypothesis related to the selection of partial and full condition and result clauses from the input*

*RQ1: When learners are aware that task demands will require output that is identical to instances of counterfactuals in the input, does this encourage the selection of task-relevant instances of past counterfactuals?*

*RQ2: When learners are aware that task demands will require output that is analogous but not identical to instances of counterfactuals in the input, does this encourage the selection of task-relevant instances of past counterfactuals?*

*H1: There will be no difference between a) the selection of tokens in the input when task demands require output that is identical to instances of counterfactuals in the input and b) the selection of tokens in the input when task demands require output that is analogous but not identical to instances of counterfactuals in the input.*

**Table 4.2:** *Descriptive statistics for partial and full selection of result clauses.*

| | | **Partial Selection** | | | | **Full Selection** | | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | n | Mean[ab] | SD | 95% CI (lower) | 95% CI (upper) | Mean | SD | 95% CI (lower) | 95% CI (upper) |
| Analogous | 33 | 10.31 | 12.74 | 5.79 | 14.83 | 23.77 | 20.45 | 16.51 | 31.02 |
| Literal | 31 | 15.40 | 12.16 | 10.95 | 19.86 | 17.74 | 17.87 | 11.18 | 24.29 |

a) Mean percentage of partial and full selections was calculated by taking the number of selected result or condition clauses and dividing this figure by the total number of selections
b) Raw counts and percentages are provided in Appendix I.

**Table 4.3:** *Descriptive statistics for partial and full selection of condition clauses.*

| | | **Partial Selection** | | | | **Full Selection** | | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | n | Mean[ab] | SD | 95% CI (lower) | 95% CI (upper) | Mean | SD | 95% CI (lower) | 95% CI (upper) |
| Analogous | 33 | 0.67 | 1.96 | -0.3 | 1.36 | 14.61 | 14.69 | 9.41 | 19.82 |
| Literal | 31 | 0.94 | 1.41 | 0.42 | 1.45 | 9.05 | 10.34 | 5.26 | 12.85 |

a) Mean percentage of partial and full selections was calculated by taking the number of selected result or condition clauses and dividing this figure by the total number of selections
b) Raw counts and percentages are provided in Appendix I.

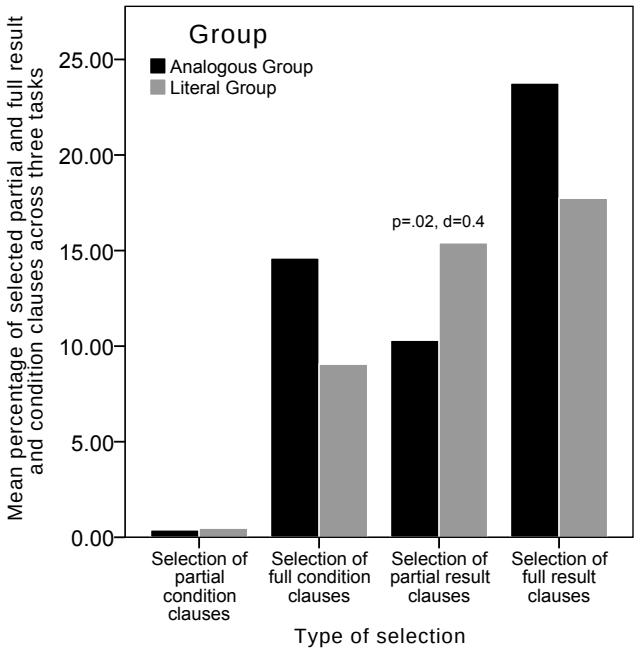



**Figure 4.1:** *Full and partial selection of result and condition clauses from the input*

the mean percentages between the groups[2]. As recommended by Norris and Ortega (2000), Cohen's $d$ was used to report effect sizes, interpreted as $d<.2$ for small effects, $.2<d<.8$ for medium effect sizes and $d>.8$ for large effect sizes[3]. Results showed no significant between-group difference[4] in the selection of full result clauses, H(1)=1.1, $p$=.295, $d$=.31, full selection of condition clauses H(1)=2.3, $p$=.13, $d$=.44, or the partial selection of condition clauses H(1)=1.3, $p$=.24, $d$=.07, but the Literal Group selected significantly more partial result

[2]Exact tests were performed to obtain significance levels on all non-parametric tests used in this study, unless computation was an issue. In this case, the Monte Carlo method was used.

[3]Effect sizes are reported throughout for both significant and non-significant results. Although for non-significant results any apparent differences may well be due to chance, they could also be due to low power and small sample sizes, in which case a higher sample size could well lead to a statistically significant result. Therefore, reporting effect sizes offers an additional interpretation of the raw scores which is of *practical significance*, if not statistical significance (Kirk, 1996).

[4]Alpha is set to .05 in all analyses, unless a correction for multiple tests was made. Exact $p$ is reported throughout, as recommended by Larson-Hall (2010) and Wilkinson and the Task Force on Statistical Inference (1999).

clauses, H(1)=4, $p$=.02, $d$=.40.

### 4.1.2 Exact and repurposed use of selected tokens

The three task performances of all participants were transcribed and analysed to identify the exact use of selected past counterfactual result and condition tokens in production, e.g. a selected base token such as *She would've come* redeployed as the target token *She would've come* in task production, and the repurposing of result and condition tokens, such as the same base token *She would've come* repurposed as a target token such as *He would've done it.* Given that the Analogous Group were provided with input that matched the necessary output in terms of function and grammatical frame, *but* not in terms of lexical items, it was predicted that this group would repurpose the input to meet the task demands. In contrast, the Literal Group were provided with input that matched the necessary output in both function, grammatical frame *and* lexical items, and therefore it was predicted that this group would use more exact phrases in their output. The relevant research questions and hypothesis are listed in Table 4.4.

**Table 4.4:** *Research questions and hypothesis related to the exact and repurposed use of result and condition clauses in task production*

| |
|---|
| *RQ3: When learners are aware that task demands will require output that is identical to instances of counterfactuals in the input, does this lead to the repurposing of instances of task-relevant counterfactuals?* |
| *RQ4: When learners are aware that task demands will require output that is analogous but not identical to instances of counterfactuals in the input, does this lead to the repurposing on instances of task-relevant counterfactuals?* |
| *H2: The repurposing of tokens to fit task demands will be greater when task demands require output that is analogous but not identical to instances of counterfactuals in the input.* |

In terms of repurposed use first, Figure 4.2 and the descriptive statistics in Table 4.5 report the repurposed use of result clauses in task production,

while Figure 4.3 and Table 4.6 report the repurposed use of condition clauses. Both a Levene's test of equality of variance and a Kolmogorov-Smirnov test of normality were conducted on the raw scores and were both found to be significant, resulting in the use of non-parametric tests. A Mann-Whitney Test on repurposed use showed a difference approaching significance on the number of repurposed result clauses, U(62)=409, z=-1.387, $p$=.083, with a medium effect size, $d$=.57, and a significant difference between the Analogous Group and the Literal Group for condition clauses, U(62)=370, z=-1.92, $p$=.028, again with a medium effect size, $d$=.49.

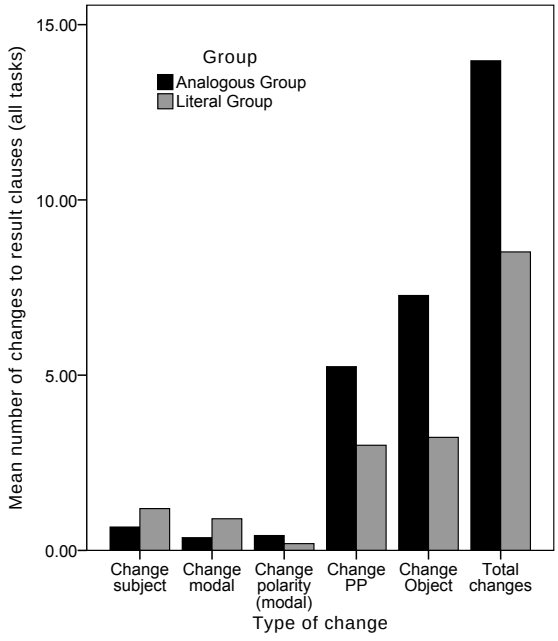



**Figure 4.2:** *Mean number of repurposed result clauses in task production*

In terms of the use of exact selections from the input in task performance, Figure 4.4 and Tables 4.5 and 4.6 illustrate the use of condition (Table 4.6) and result (Table 4.5) clauses exactly identical to the input for both the Analogous and Literal groups. A Mann-Whitney test showed a significant difference between the two groups for the exact use of condition clauses, U(62)=249.5, z=-3.8, $p$=.000, with a large effect size, $d$=1.03, and the exact use of result

**Table 4.5:** *Descriptive statistics for the exact and repurposed use of result clauses.*

| | | Exact Use | | | | Repurposed Use | | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | n | Mean | SD | 95% CI (lower) | 95% CI (upper) | Mean | SD | 95% CI (lower) | 95% CI (upper) |
| Analogous | 33 | 0.45 | 0.75 | 0.19 | 0.72 | 13.97 | 13.50 | 9.18 | 18.76 |
| Literal | 31 | 2.48 | 2.87 | 1.43 | 3.54 | 8.52 | 9.99 | 4.85 | 12.18 |

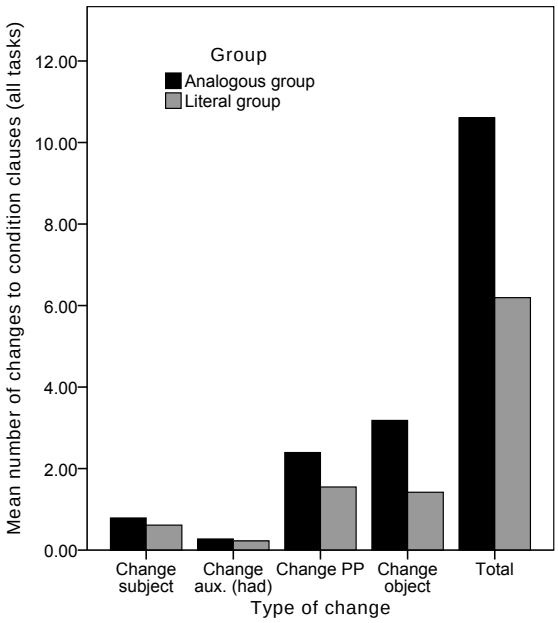



**Figure 4.3:** *Mean number of repurposed condition clauses in task production*

**Table 4.6:** *Descriptive statistics for the exact and repurposed use of condition clauses.*

| | | Exact Use | | | | Repurposed Use | | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | n | Mean | SD | 95% CI (lower) | 95% CI (upper) | Mean | SD | 95% CI (lower) | 95% CI (upper) |
| Analogous | 33 | 0.58 | 1.15 | 0.17 | 0.98 | 10.61 | 9.71 | 7.16 | 14.05 |
| Literal | 31 | 2.52 | 2.42 | 1.63 | 3.40 | 6.19 | 8.38 | 3.12 | 9.27 |

clauses, U(62)=269.5, z=-3.5, $p$=.000, also with a large effect size, $d$=.97.

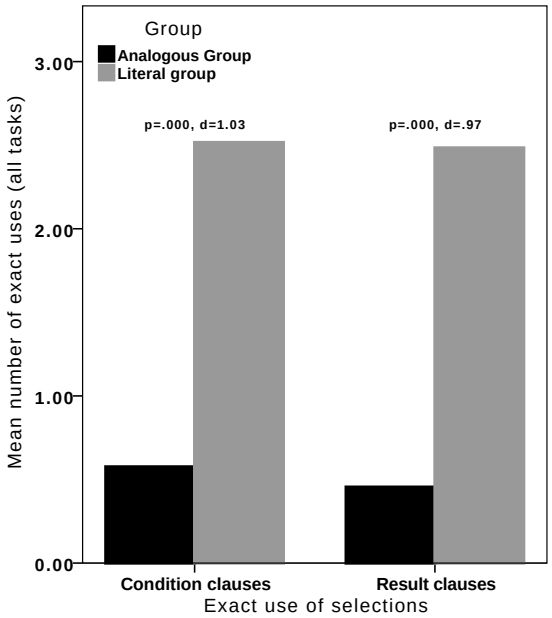



**Figure 4.4:** *Exact use of result and condition clauses in task production*

### 4.1.3 Treatment effects on acceptability judgment and production tests

This section reports between and within-group differences on the acceptability judgment test (Section 4.1.3.1), and the performance test (Section 4.1.3.2). The related research question and hypothesis are provided in Table 4.7.

**Table 4.7:** *Research questions and hypothesis related to the development of past-counterfactuals*

| |
|---|
| *RQ5: Does the selection and use of exact tokens correlate with gains in past counterfactual knowledge?* |
| *RQ6: Does the selection and repurposing of tokens correlate with gains in past counterfactual knowledge?* |
| *H3: The selection and repurposing of tokens will lead to superior performance on post-tests than selection and exact use of tokens.* |

#### 4.1.3.1 Acceptability judgment test

Table 4.8 and Figure 4.5 report the descriptive statistics for the acceptability judgment test at pre- and post-test for the three groups. These results provide an initial indication that the Analogous Group outperformed the Literal and Control groups. In-depth profiles of individual performances on the acceptability judgment tests are displayed in Figure 4.6. These profile plots, in addition to the large confidence intervals for the mean scores on the acceptability judgment test (Table 4.8), indicate considerable variability across performances.

**Table 4.8:** *Descriptive statistics for pre/post-test scores on the acceptability judgment test.*

| | | **Pre-test** | | | | **Post-test** | | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | n | Mean[a] | SD | 95% CI (lower) | 95% CI (upper) | Mean | SD | 95% CI (lower) | 95% CI (upper) |
| Analogous | 30 | 46.62 | 9.16 | 43.20 | 50.04 | 54.35 | 8.71 | 51.10 | 57.60 |
| Literal | 30 | 47.12 | 5.78 | 44.95 | 49.26 | 49.95 | 7.19 | 47.27 | 52.64 |
| Control | 26 | 45.51 | 6.41 | 42.92 | 48.10 | 49.07 | 7.83 | 45.90 | 52.23 |

a) The maximum score for the acceptability judgment test was 100

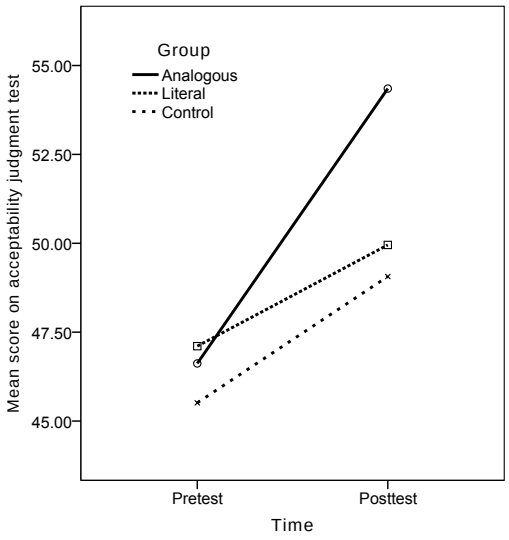
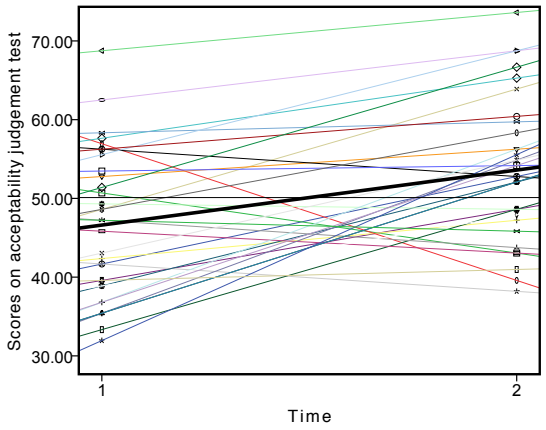
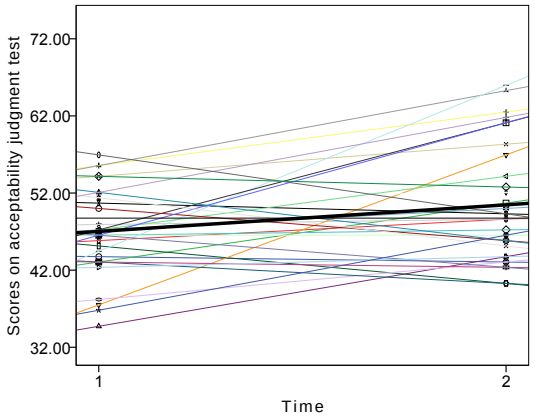



**Figure 4.5:** *Pre and post-test scores on the acceptability judgment test for the Analogous, Literal and Control groups*

**(a)** *Analogous Group*




**(b)** *Literal Group*




**(c)** *Control Group*

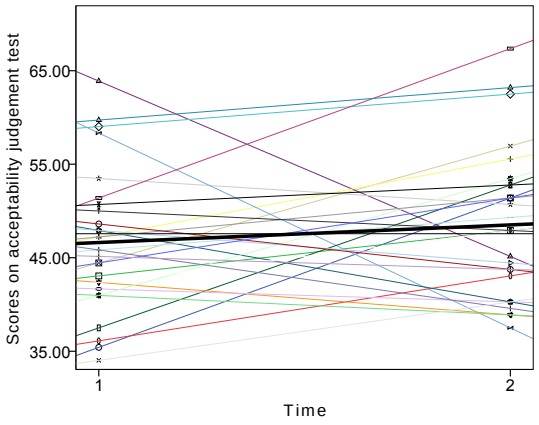



**Figure 4.6:** *Profile plots for Analogous, Literal and Control groups*

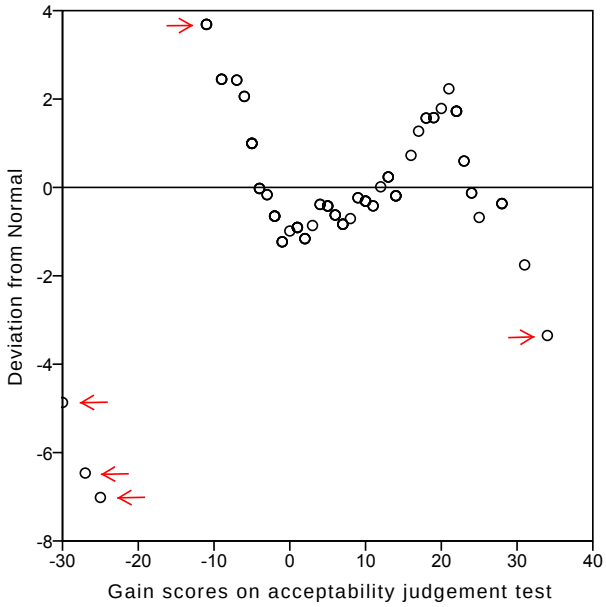



**Figure 4.7:** *Q-Q plot of acceptability judgment test scores, showing outliers three or more standard deviations from the mean.*

Prior to inferential tests, the data was pruned of outliners that were three or more standard deviations away from the mean (illustrated in Figure 4.7). Initial Kolmogorov-Smirnov tests were also run, and showed a normal distribution of the data. Subsequently, a parametric repeated measures ANOVA was performed with one between-group variable at three levels (Analogous, Literal and Control groups), and one within group variable at two levels (time 1: pre-test, and time 2: post-test). Levene's test of homogeneity of variance was violated on the pre-test, $p$ <.05), and therefore ANOVA results not assuming homogeneity of variance are reported.

Results from the ANOVA showed a significant main effect for Time, F(2, 83)=37.97, $p$=.000, $\eta_p^2$=.314, and a significant Time x Group interaction, F(2, 83)=4.12, $p$=.02, $\eta_p^2$=.09. Effect sizes (Partial eta squared, $\eta_p^2$) were interpreted following Cohen (1988) as small>.01, medium>.06 and large>.14. Table: 4.9 shows post hoc Scheffé between-group comparisons. These revealed no significant difference between the Literal Group and the Control Group, $p$=.927, with

a small effect size, $d$=.1, and a significant differences between the Analogous Group and the Literal Group, $p$=.024 with a medium effect size, $d$=0.69. The difference between the Analogous and Control groups approached significance, $p$=.09, with a medium effect size between the two groups, $d$=0.58.

It is worth pointing out at this point that although significance was reached with a medium effect size between the Analogous and the Literal group, the 95% confidence interval range shows a large amount of imprecision in the results (.48-13.59 for the Analogous and Literal groups, and -.8 and 12.82 for the Analogous and Control groups).

**Table 4.9:** *Between-group differences in gain scores on the acceptability judgment test*

| Group Difference | Significance | 95% CI | 95% CI | Effect Size |
|---|---|---|---|---|
| Analogous > Literal | $p$=.024[a] | .48 | 13.59 | .69 |
| Analogous > Control | $p$=.09*(n.s.)* | -.8 | 12.82 | .58 |
| Control > Literal | $p$=.927*(n.s.)* | -7.8 | 5.8 | .10 |

a=Significant at alpha level 0.05

Within-group comparisons revealed no significant differences between pre and post-test for the Control Group, F(1, 50)=3.21, $p$=.08, $d$=.50, or for the Literal Group, F(1, 58)=2.86, $p$=.10, $d$=.44, but a significant difference between pre and post-test and for the Analogous Group, F(1, 58)=11.22, $p$=.001, with a large effect size, $d$=.86.

#### 4.1.3.2 Production test

Table 4.10 and Figure 4.8 show descriptive statistics for the three groups (Analogous, Literal and Control) at both pre-test and post-test for the production of result clauses on the production test. Recall that for the result clauses, the score on pre and post-test is determined by the number of targetlike utterances (3 points), counterfactual modal with signs of pastness (2 points) and counterfactual modal with no signs of pastness (1 point). The sum is then divided

by the number of obligatory contexts. See *Section 3.6: Coding of Production Tests.*

**Table 4.10:** *Descriptive statistics for the use of result clauses in pre- and post-production test*

| | | Pre-test | | | | Post-test | | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | n | Mean[a] | SD | 95% CI (lower) | 95% CI (upper) | Mean | SD | 95% CI (lower) | 95% CI (upper) |
| Analogous | 33 | .46 | .58 | .26 | .67 | 1.17 | 1.12 | .77 | 1.57 |
| Literal | 31 | .58 | .73 | .31 | .85 | 1.13 | 1.02 | .75 | 1.5 |
| Control | 28 | .24 | .34 | .11 | .37 | .33 | .63 | .09 | .58 |

a) The maximum score was three points, e.g. Ten obligatory contexts and the use of ten TLUs; $Total\ Score = \frac{3points\ (TLU)\ X\ 10\ occurrences}{10\ obligatory\ contexts} = 3$. See 3.6 Coding of Production Tests for details.

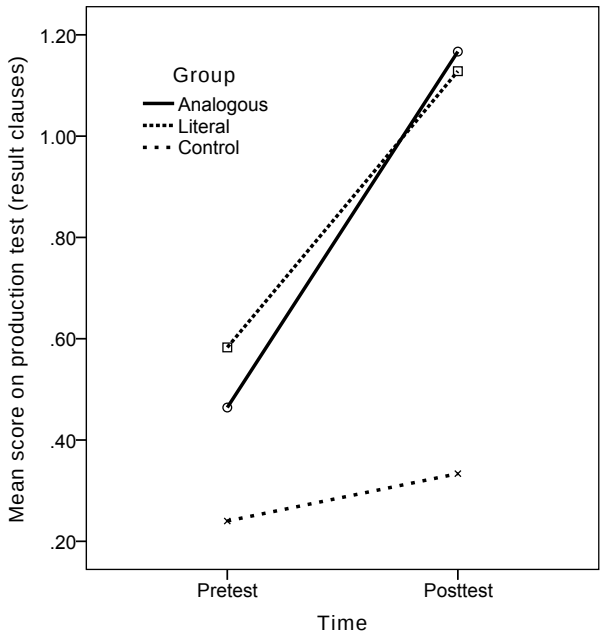



**Figure 4.8:** *Use of result clauses in production test*

Figure 4.9 and Table 4.11 show descriptive statistics for the Analogous and Literal groups at pre-test and post-test, but this time for condition clauses. Again, note that the scores for the condition clauses are determined by the number of targetlike utterances, the number of conditions with signs of counterfactuality and pastness and the number of conditional clauses with signs of

**Table 4.11:** *Descriptive statistics for the use of condition clauses in the production test*

| | | pre-test | | | | Post-test | | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | n | Mean[a] | SD | 95% CI (lower) | 95% CI (upper) | Mean | SD | 95% CI (lower) | 95% CI (upper) |
| Analogous | 33 | 1.14 | .77 | .87 | 1.41 | 1.51 | .95 | 1.17 | 1.85 |
| Literal | 31 | .87 | .50 | .69 | 1.06 | 1.67 | .87 | 1.35 | 1.99 |
| Control | 28 | 1.05 | .86 | .72 | 1.39 | 1.27 | 1.02 | .88 | 1.67 |

a) The maximum score was three points, e.g. Ten obligatory contexts and the use of ten TLUs; $Total\ Score = \frac{3points\ (TLU)\ X\ 10\ occurrences}{10\ obligatory\ contexts} = 3$. See 3.6 Coding of Production Tests for details.

counterfactuality OR pastness divided by the number of obligatory contexts. See *Section 3.6: Coding of Production Tests.*

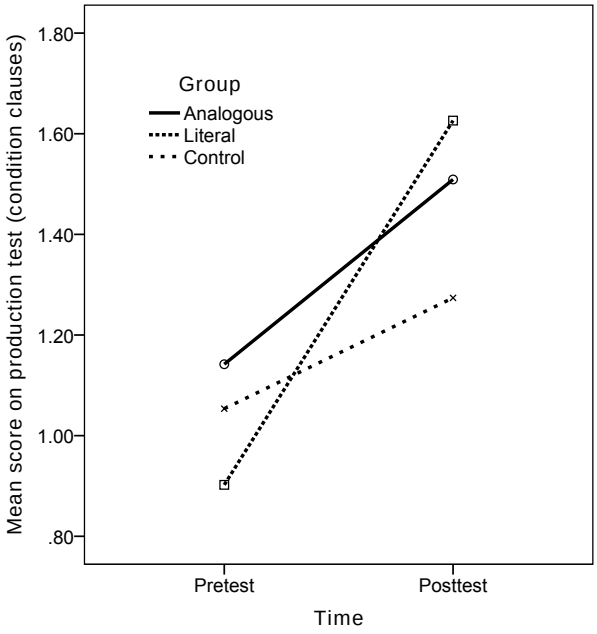



**Figure 4.9:** *Use of condition clauses in production test*

For both pre-tests and post-tests, a Kolmogorov-Smirnov test was used to test for normality and Levene's test for homogeneity of variance. Both tests were significant ($p < .05$), resulting in the use of non-parametric tests for between and within group comparisons.

A non-parametric Friedman's test showed a significant difference between

**Table 4.12:** *Within group differences for the production test*

| Group | Difference | Effect Size (Cohen's $d$) |
|---|---|---|
| **Condition Clauses** | | |
| Analogous | post-test>pre-test ($p$=0.031[a,b]) | 0.43 |
| Literal | post-test>pre-test ($p$=0.000[*]) | 1.13 |
| Control | post-test>pre-test ($p$=0.1) | 0.66 |
| **Result Clauses** | | |
| Analogous | post-test>pre-test ($p$=0.000[*]) | 0.8 |
| Literal | post-test>pre-test ($p$=0.000[*]) | 0.62 |
| Control | post-test>pre-test ($p$=0.425) | 0.18 |

a=Significant at alpha level 0.05
b=Exact tests were run to obtain significance levels and therefore no confidence intervals are reported

pre-test and post-test for condition clauses ($\chi^2(2) = 11.215, p = .01$) and result clauses ($\chi^2(2) = 15.211, p = .00$). To determine the location of the significance, first a within-group post-hoc analysis was conducted between pre-test and post-test using a Wilcoxon Signed Rank test. As illustrated in Table 4.12 for result clauses, there was a significant difference ($p$=0.000) between pre and post-test for the Analogous Group, with a large effect size of $d$=.8. The difference between pre and post-test for the Literal Group was also significant at $p$=0.000, this time with a medium effect size of $d$=.62. There was no signficant difference between pre and post-test for the Control Group, $p$=0.425, $d$=.18.

For condition clauses, there was a significant difference between pre and post-test for the Analogous Group ($p$=0.031), and the Literal Group ($p$=0.000), with medium and large effect sizes respectively (Literal Group: $d$=1.13, Analogous Group: $d$=.43). There was no significant difference between pre and post-tests for the control group, ($p$=0.1), although there was a medium effect size of $d$=.66.

Post-hoc Mann-Whitney tests were used on mean gain scores to examine between-group differences on the performance test, with a Bonferroni correction to control for type 1 errors. Table 4.13 summarises these results for both result and condition clauses between all groups. Both the Analogous and Literal

**Table 4.13:** *Between-group differences in gain scores on the production test*

| Group | Difference | Effect Size (Cohen's *d*) |
|---|---|---|
| **Condition Clauses** | | |
| Analogous X Literal Group | Analogous > Literal ($p$= .12[a] *n.s.*) U(62)=426, z=-1.6 | .18 |
| Analogous X Control | Analogous > Control ($p$= .3, *n.s.*) U(59)=427, z=-.5 | .24 |
| Literal X Control | Literal > Control ($p$= .04, *n.s.*) U(57)=321.5, z=-1.82 | .42 |
| **Result Clauses** | | |
| Analogous X Literal | Analogous > Literal ($p$= .44, *n.s.*) U(62)=500, z=-.16 | .04 |
| Analogous X Control | Analogous > Control ($p$= .004[b]) U(59)=279.5, z=-2.65 | .92 |
| Literal Group X Control | Literal > Control ($p$= .008[b]) U(57)=276, z=-2.41 | .94 |

a=Exact tests were run to obtain significance levels and therefore no confidence intervals are reported
b=Significant when alpha level is set to 0.0167 following a Bonferroni correction

groups showed significant differences with the Control Group on result clause measures, with large effect sizes ($d$=.92, $p$=.004 for the Analogous Group and $d$=.94, $p$=.008 for the Literal Group), but a negligible effect size when compared to each other ($d$=.04, *n.s*). For the condition clauses, the Literal and Analogous groups outperformed the Control Group with medium effect sizes, $d$=.24 for the Analogous Group and $d$=.42 for the Literal Group, although the differences between the Analogous and Control groups, and the Literal and Control groups were not significant.

### 4.1.4 Effects of exact use and repurposed use on tests

The previous section explored the relationship between condition (Analogous, Literal and Control) and gains on test scores without differentiating between the selection of past counterfactuals from the input and any subsequent repurposing of these selections to fit the task demands. This section teases apart selection and repurposing to give a more detailed picture of the correlations between on the one hand selections from the input and gain scores, and on the

other repurposing of selected tokens and gain scores.

As outlined in Section 4.1.2, while the Literal Group used significantly more exact phrases than the Analogous Group in their task productions, some participants in the Literal Group went beyond the rubric, and repurposed the phrases to more finely suit their intentions. It would seem while the Analogical Group were forced to repurpose the input due to the inadequacy of the input for their current task demands, some participants in the Literal Group chose to repurpose. To factor this in to an analysis of the effects of repurposing and selection on gain scores, no distinction was made between the Analogous and Literal groups in the regression analyses (n=64).

Hierarchical multiple regression analyses were used to examine to what extent the following predictors (102-106) could separately predict performance gains between pre-test and post-test on the acceptability judgment test and the production test.

102. Knowledge of past counterfactual constructions at pre-test (*Pre-test Scores*)
103. *Partial selection* of result/condition clauses
104. *Full selection* of result/condition clauses
105. *Exact use* of result/condition clauses in task production
106. *Repurposed use* of result/condition clauses in task production

Prior to each analysis, the data were assessed for the normality, linearity, multicollinearity and homoscedasticity of the sample.

#### 4.1.4.1 Predictors of acceptability judgment test scores

The first analysis examined how well selection of result clauses and subsequent use could predict gains on the acceptability judgment test, after initially controlling for the effect of pre-test scores. Initial analysis showed no violations after removing outliers using Mahalanobis distances. Pre-test scores for the production test were entered at step one of the model and explained 22% of the variance. Partial selection of input, full selection of input, exact use of

tokens and repurposed use of tokens were entered at steps two, three, four and five respectively. After all five predictors were entered into the equation, the model accounted for 39% of the variance in the acceptability judgment test gain scores. Both partial selection and full selection together accounted for only 2% of the variance, with exact use of the input in production (the use of past counterfactual tokens from the input with no changes) having a negligible effect. Only repurposed input reached significance after the other predictors had been accounted for (13% of the variance). Results are shown in table 4.14 with corresponding standardized beta coefficients ($\beta$).

**Table 4.14:** *Hierarchical regression analysis of the selection and use of result clauses as predictors of gain scores on acceptability judgment test*

| Model | $R^2$ | $\Delta R^2$ | B | SE B | $\beta$ |
|---|---|---|---|---|---|
| Step 1: pre-test scores | .22* | .22* | -.63 | .15 | -.47 |
| Step 2: partial selection | .22 | .00 | -.1 | .37 | .03 |
| Step 3: full selection | .25 | .02 | -.1 | .08 | .16 |
| Step 4: exact use | .26 | .01 | -.72 | .8 | -.1 |
| Step 5: repurposed use | .39* | .13* | -.43 | .12 | .43 |

* Significant at $p < 0.005$. With five predictors, model is significant, $F_{5, 60} = 6.4$, $p < 0.005$

The second regression analysis was identical to the first, but with the selection and use of *condition* not result clauses as predictors. Preliminary analysis showed high levels of heteroscedasticity for the partial selection of the input, and it was therefore removed from the model. No other violations were evident. Similar to the analysis of the result clauses, the pre-test scores accounted for 22% of the variance (the first step in the model) and of the remaining predictors only repurposed use of tokens reached significance, accounting for 17% of the total variance. Noticeably, the selection of input in both the condition and result models could account for very little of the gain scores, with full selection and exact use accounting for just 4% of the variance. Table 4.15 shows the results of the analysis.

**Table 4.15:** *Hierarchical regression analysis of the selection and use of condition clauses as predictors of gain scores on acceptability judgment test*

| Model | $R^2$ | $\Delta R^2$ | B | SE B | $\beta$ |
|---|---|---|---|---|---|
| Step 1: pre-test scores | .22* | .22* | -.63 | .15 | -.47 |
| Step 2: full selection | .23 | .01 | -.44 | .5 | .1 |
| Step 3: exact use | .26 | .03 | -1.1 | .75 | -.19 |
| Step 4: repurposed use | .43* | .17* | .6 | .15 | .45 |

* Significant at $p < 0.005$. With five predictors, model is significant, $F_{5, 60} = 6.4$, $p < 0.005$

#### 4.1.4.2 Predictors of production test scores

The second wave of regression models used the same predictors as the previous section (selection and use of the input), but this time the models were used to predict the gain scores on the production test for both condition and result clauses. Partial selection of the input was removed in the analysis of condition clauses for the production test due to heteroscedasticity. Outliers were removed from analysis in both regression models using Mahalanobis distances.

The first analysis (Table 4.16) examines the effect of the selection and use of condition clauses on gains between pre-test and post-test on the production test. In line with the previous regression analyses, the pre-test data were entered into the equation first to control for level of proficiency. This accounted for 26% of the variance. Again, repurposed input was significant, accounting for 10% of the variance, once all other predictors had been entered. A notable difference here is the effect that the predictor *exact use* had on the model, accounting for 12% of the variance. In total, the model accounted for almost 50% of the variance.

The final regression model examined the extent to which selection and use could predict the use of result clauses in the production test (Table 4.17). In the same manner as previous models, repurposed use (entered last in the equation) was significant, accounting for over 24% of the variance. The only other significant predictor was the selection of result clauses from the input,

**Table 4.16:** *Hierarchical regression analysis of the selection and use of condition clauses as predictors of gain scores on the production test*

| Model | $R^2$ | $\Delta R^2$ | B | SE B | $\beta$ |
|---|---|---|---|---|---|
| Step 1: pre-test scores | .26* | .26* | -.81 | .18 | -.51 |
| Step 2: full selection | .27 | .02 | .05 | .04 | .13 |
| Step 3: exact use | .39 | .12 | .19 | .06 | .38 |
| Step 4: repurposed use | .48* | .1* | .04 | .01 | .35 |

* Significant at $p < 0.005$. With four predictors, model is significant, F5, 60 =14.1, $p < 0.005$

**Table 4.17:** *Hierarchical regression analysis of the selection and use of result clauses as predictors of gain scores on the production test*

| Model | $R^2$ | $\Delta R^2$ | B | SE B | $\beta$ |
|---|---|---|---|---|---|
| Step 1: pre-test scores | .00 | .00 | -.00 | .20 | -.00 |
| Step 2: partial selection | .00 | .00 | .00 | .3 | -.01 |
| Step 3: full selection | .1* | .1* | .05 | .02 | .32 |
| Step 4: exact use | .1 | .00 | .00 | .07 | .01 |
| Step 5: repurposed use | .34** | .24** | .04 | .01 | .55 |

* Significant at p<0.05*. **Significant at p<0.005. With five predictors, model is significant, F5, 60 =5.6, $p < 0.005$

which accounted for 10% of the variance in gain scores. A marked difference between this model and the previous was the effect the pre-test had on the overall model. In previous models, knowledge of past counterfactuals at pre-test accounted for between 20 and 26% of the total variance, whereas in the model for result clauses and the production test, the pre-test had no effect.

#### 4.1.4.3 Overall proficiency as a predictor of test scores

A further series of regression analyses were run to account for any effects overall level, operationalised here as self-reported IELTS scores, had on the acceptability judgment test and the production test. Prior to each analysis, the data were assessed for the normality, linearity, multicollinearity and homoscedasticity of the sample.

As illustrated in Table 4.18, the first model in this series of regression models

uses IELTS scores as a predictor of variance in gain scores on the acceptability judgment test. Results of the analysis show that IELTS scores do not significantly predict gains on this test.

**Table 4.18:** *Regression analysis of IELTS score as a predictor of gain scores on the acceptability judgment test*

| Predictor | $R^2$ | B | SE B | $\beta$ |
|---|---|---|---|---|
| IELTS score | .013 | 2.87 | 3.38 | .116 |

Model is not significant, F1, 53=0.72, $p$=0.4

Tables 4.19 and 4.20 again show IELTS as a predictor, but this time as a predictor of variance in gain scores on the production of result clauses (Table 4.19) and condition clauses (Table 4.20) in the production test.

**Table 4.19:** *Regression analysis of IELTS score as a predictor of gain scores on the production of result clauses in the production test*

| Predictor | $R^2$ | B | SE B | $\beta$ |
|---|---|---|---|---|
| IELTS score | .008 | .14 | .21 | .09 |

Model is not significant, F1, 53=0.44, $p$=0.512

**Table 4.20:** *Regression analysis of IELTS score as a predictor of gain scores on the production of condition clauses in the production test*

| Predictor | $R^2$ | B | SE B | $\beta$ |
|---|---|---|---|---|
| IELTS score | .00 | -.04 | .28 | -.02 |

Model is not significant, F1, 53=.02, p = 0.9

Similar to the prediction of variance on the acceptability judgment test, IELTS scores could not significantly predict gains in the production of result or condition clauses in the production test. If we compare the significant effect of pre-test past-counterfactual knowledge on the variance in gain scores, as illustrated in Tables 4.14, 4.15 and 4.15, with the non-significant effects here,

it would seem that the knowledge of past-counterfactual knowledge is somewhat isolated from general proficiency, as measured by self-report IELTS scores.

### 4.1.5 The effects of cross-linguistic variation across L1s

As detailed in *Section 3.2.2.1.4: Cross-linguistic variation*, the L1s of the participants in the study were divided into three groups; in the first group (Verb Group), the L1s of the participants signal past counterfactuality with the Tense, Aspect, Mood (TAM) system, i.e. the back shifting of tense in English conditional clauses. For the L1s in the second group, the use of particles in conjunction with the TAM system is necessary to signal past counterfactuality (Particle Group), while in the third group, the L1s relied on both particles and pragmatic interpretation to infer a counterfactual meaning (Pragmatic Group). To investigate any differences in performance on pre- and post-tests, the performance of the three groups on the acceptability judgment and production test were compared. The relevant research question is listed below.

**Table 4.21:** *Research questions and hypothesis related to the development of past-counterfactuals*

| |
|---|
| *RQ8: Do cross-linguistic differences in the encoding of past counterfactuality in the L1 affect the development of counterfactual constructions in the L2?* |

Figure 4.10 and Table 4.22 report the descriptive statistics for the performances of the three groups (Verb, Particle and Pragmatic) on the acceptability judgment test. A repeated measures ANOVA was performed on the data with one between-group variable at three levels (Verb, Particle and Pragmatic groups), and one within-group variable at two levels (time 1: pre-test, and time 2: post-test). Results showed no significant time X group interaction, $F(2, 83)=.23$, $p=.80$, with a small effect size of $\eta_p^2=.005$.

Figure 4.11 and Table 4.23 report the descriptive statistics for the performances of the three L1 groups on the production of result clauses in the

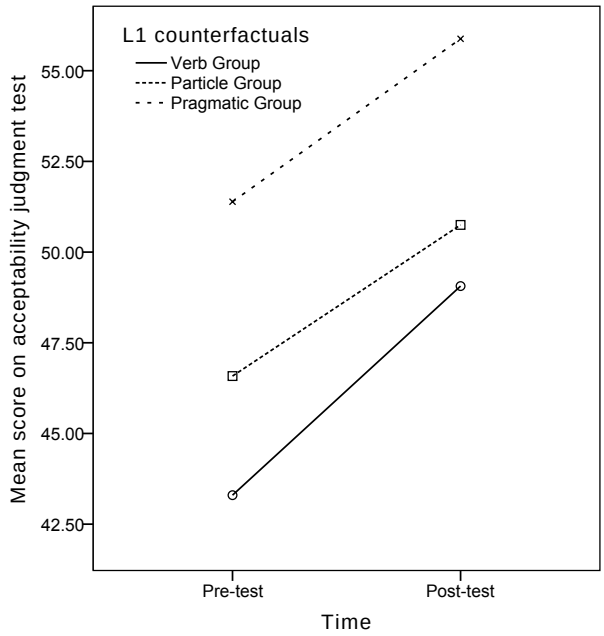



**Figure 4.10:** *Mean pre-test/post-test scores on the acceptability judgment test for the three L1 groups*

**Table 4.22:** *Descriptive statistics for pre- and post-test scores on the acceptability judgment test for Verb, Particle and Pragmatic L1 groups.*

| | | **Pre-test** | | | | **Post-test** | | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | n | Mean[a] | SD | 95% CI (lower) | 95% CI (upper) | Mean | SD | 95% CI (lower) | 95% CI (upper) |
| Verb | 20 | 43.30 | 5.82 | 40.57 | 46.02 | 49.06 | 6.64 | 45.96 | 52.17 |
| Particle | 51 | 46.58 | 7.28 | 44.54 | 48.63 | 50.75 | 6.64 | 48.53 | 52.97 |
| Pragmatic | 15 | 51.39 | 9.18 | 46.31 | 56.47 | 55.88 | 11.03 | 49.77 | 61.99 |

a) The maximum score for the acceptability judgment test was 100

production test. The data for all three groups showed a non-normal distribution in Kolmogorov-Smirnov tests, resulting in the use of non-parametric Kruskall Wallis tests to compare groups. Results showed that the type of L1 did not significantly affect the production of result clauses on the production test, H(2)=1.87, $p$=.40, with small effect sizes across all groups and 99% confidence intervals between .38 and .41 using the Monte Carlo method.

Figure 4.12 and Table 4.24 show descriptive statistics for the performances of the three groups on the production of condition clauses on the produc-

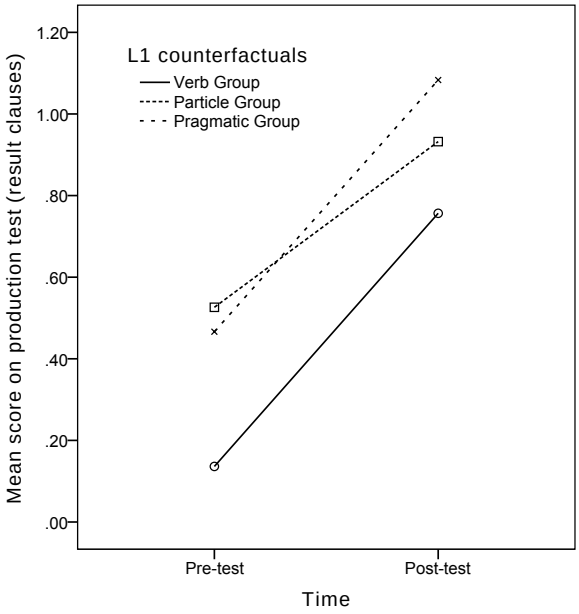



**Figure 4.11:** *Mean pre-test/post-test scores on production test for the three L1 groups (result clauses)*

**Table 4.23:** *Descriptive statistics for pre- and post-test scores on the production test for Verb, Particle and Pragmatic L1 groups (result clauses).*

| | | **Pre-test** | | | | **Post-test** | | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | n | Mean[a] | SD | 95% CI (lower) | 95% CI (upper) | Mean | SD | 95% CI (lower) | 95% CI (upper) |
| Verb | 20 | .14 | .24 | .02 | .25 | .76 | 1.03 | .28 | 1.24 |
| Particle | 51 | .53 | .68 | .34 | .72 | .93 | 1.00 | .28 | 1.24 |
| Pragmatic | 15 | .47 | .50 | .19 | .74 | 1.08 | .94 | .56 | 1.61 |

a) The scores on the pre- and post-tests were determined by the number of Target-Like Utterances (3 points), the number of counterfactual modals with signs of pastness (2 points) and the number of counterfactual modals with no signs of pastness (1 point). The sum was then divided by the number of obligatory contexts. See Section 3.7 Coding of Production Tests.

tion test. After non-normal Kolmogorov-Smirnov tests, a Kruskall Wallis test showed no significant effects of the type of past counterfactuality encoding in the L1 on the production of condition clauses, H(2)=1.71, $p$=.43, with small effect sizes between all groups and 99% confidence intervals between .42 and .45, using the Monte Carlo method.

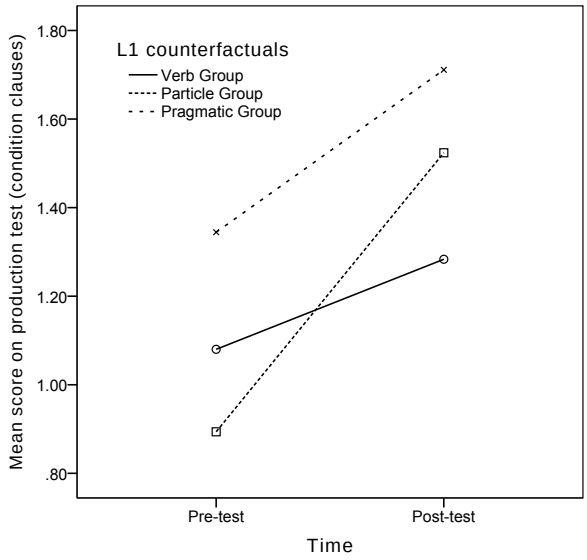



**Figure 4.12:** *Mean pre-test/post-test scores on production test for the three L1 groups (condition clauses)*

**Table 4.24:** *Descriptive statistics for pre- and post-test scores on the production test for Verb, Particle and Pragmatic L1 groups (condition clauses).*

| | | Pre-test | | | | Post-test | | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | n | Mean[a] | SD | 95% CI (lower) | 95% CI (upper) | Mean | SD | 95% CI (lower) | 95% CI (upper) |
| Verb | 20 | 1.08 | .77 | .72 | 1.44 | 1.28 | 1.03 | .80 | 1.76 |
| Particle | 51 | .89 | .61 | .72 | 1.06 | 1.52 | .87 | 1.28 | 1.77 |
| Pragmatic | 15 | 1.34 | .83 | .89 | 1.8 | 1.71 | 1.10 | 1.10 | 2.32 |

a) The scores on the pre- and post-tests were determined by the number of Target-Like Utterances (3 points), the number of counterfactual modals with signs of pastness (2 points) and the number of counterfactual modals with no signs of pastness (1 point). The sum was then divided by the number of obligatory contexts. See Section 3.7 Coding of Production Tests.

#### 4.1.5.1 Overall proficiency as a confounding variable

To isolate overall proficiency, operationalised as self-reported scores on IELTS tests, from any differences between the means of the Verb, Particle and Pragmatic groups on the acceptability judgment test and the production test, a post-hoc ANCOVA analysis was performed with self-reported IELTS scores as a covariate. Results from the analysis showed that for gains on the acceptability judgment test IELTS scores were not significantly related to gain

scores, F(1,47)=.738, $p$=.395, with a small-medium effect size, $\eta_p^2$=.015[5]. After controlling for IELTS scores, there was also no significant difference between means (on gain scores) for the Verb, Particle and Pragmatic L1 backgrounds, F(2,47)=.082, $p$=.922, with a small effect size, $\eta_p^2$=.003. For the production test, similar results were found: for result clauses, again self-reported IELTS scores were not significantly related to gain scores, F(1,47)=.110, $p$=.742, with a small effect size, $\eta_p^2$=.002, and there was also no significant difference between means for the Verb, Particle and Pragmatic L1 backgrounds after controlling for the IELTS scores, F(2,47)=.334, $p$=.718, with a small-medium effect size, $\eta_p^2$=.014. For condition clauses, very similar results were also apparent, F(1,47)=.001, $p$=.972, with a small effect size, $\eta_p^2$=.000 for the relationship between IELTS and gain scores, and F(2,47)=.330, $p$=.721, with a small-medium effect size, $\eta_p^2$=.014, for the difference between the means of the three groups after accounting for the IELTS scores.

### 4.1.6 Summary of statistical analyses

*H1: There will be no difference between a) the selection of tokens in the input when task demands require output that is identical to instances of counterfactuals in the input and b) the selection of tokens in the input when task demands require output that is analogous but not identical to instances of counterfactuals in the input.*

Hypothesis One is mainly supported. There were no significant differences between the Analogous and Literal groups on the number of selections of full result clauses, H(1)=1.1, $p$=.295, $d$=.31, on the number of selections of full condition clauses, H(1)=2.3, $p$=.13, $d$=.44, or on the number of selections of condition clauses, H(1)=.17, $p$=.68, $d$=.07. There was, however, a significant difference between group on the selection of partial result clauses, H(1)=4,

[5]Magnitude of effect sizes were interpreted following Cohen (1988) as small>.01, medium>.06 and large>.14

$p$=.02, $d$=.40, with the Literal Group selecting more partial result clauses than the Analogous Group.

*H2: The repurposing of tokens to fit task demands will be greater when task demands require output that is analogous but not identical to instances of counterfactuals in the input.*

Hypothesis Two is supported for the use of exact condition and exact result clauses in production, i.e. selected tokens were used verbatim in production. The Literal Group produced significantly more exact result clauses with a large effect size, U(62)=269.5, z=-3.5, $p$=.000, $d$=.97, and significantly more condition clauses, also with a large effect size, U(62)=249.5, z=-3.8, $p$=.000, $d$=1.03.

For the repurposing of tokens, Hypothesis Two is mainly supported. The Analogous Group significantly outperformed the Literal Group with a medium effect size on the number of repurposed condition clauses in task production, U(62)=370, z=-1.92, $p$=.028, $d$=.49. The difference between groups approached significance on the number of repurposed result clauses, again with a medium effect size, U(62)=409, z=-1.387, $p$=.083, $d$=.57.

*H3: The selection and repurposing of tokens will lead to superior performance on post-tests than selection and exact use of tokens.*

Hypothesis Three is largely supported by gains on the acceptability judgment test. The Analogous Group outperformed both the Literal Group and the Control Group, with a significant difference between the Analogous and the Literal groups (Analogous>Literal, $p$=.024, $d$=.69), and a difference approaching significance between the Analogous and Control groups (Analogous>Control, $p$=.09, $d$=.58). In terms of the production test, Hypothesis Three is not supported. Although the Analogous Group and the Literal Group significantly outperformed the Control Group on the production of result clauses with

large effect sizes (Analogous>Control, $p$=.004, $d$=.92, Literal Group>Control, $p$=.008, $d$=.92), the Analogous Group and the Literal Group were approximately equal in gains in the use of past-counterfactual tokens between pre and post-test. In the production of condition clauses, the Literal Group outperformed the Analogous Group, although not significantly, and there was no significant difference between the Analogous Group and the Control Group.

Regression models were used to isolate the effects of selection of past counterfactual tokens from the effects of repurposing of past counterfactual tokens in order to determine their individual effects of gain scores. The repurposing of selected phrases significantly affected gain scores on the acceptability judgment test ($\Delta R^2 = 13\%$, $p$<.005 for result clauses and 17%, $p$<.005 for condition clauses), even when scores, selection of input and exact use had been accounted for in the regression modals. This is supported by the effect repurposing had on the performance test ($\Delta R^2 = 24\%$, $p$<.005 for result clauses and 10%, $p$<.05 for condition clauses), again when the same factors had been accounted for.

Interestingly, in terms of the effect the selection of past counterfactual tokens had on gains in both tests, results from the regression model showed selection had little impact on the acceptability judgment test ($\Delta R^2 = 1\%$ for full selection of condition clauses; $\Delta R^2 = 2\%$ for both partial and full selection of result clauses), neither of which were significant. For the full selection of condition clauses as a predictor of gains on the production test, the effect is similar with only 2% of variance explained ($\Delta R^2$ =2%, *n.s.*). However, the exact use of condition clauses explained gains on the production test, although not at significant levels ($\Delta R^2$= 12%, *n.s.*), as did the full selection of result clauses ($\Delta R^2$= 10%, $p$<.05).

The discrepancy between the two tests is possibly due to the possibility of using selected tokens in the production test. That is, a token can be remembered, and then can be used in its entirety in the production test, leading to gains. This is not the case, however, in the acceptability judgment test, which demands the ability to go beyond memorised tokens and use schematic

(or rule-based) knowledge to decide on the acceptability of instances not seen in the input.

A result worth noting is the difference between the selection of partial past counterfactual elements and the full selection of clauses. It is clear from the regression models that partial selections that constituted a verb phrase (e.g. *'had'* or past participles) had almost no effect of the differences at test. Full selection had a much greater impact on the acceptability judgment test (although still only 2%) and the production test (10%, $p<.05$).

Finally, in terms of cross-linguistic variation, the way past counterfactuals are encoded in the L1s of the participants had little statistical effect on gain scores. On the acceptability judgment test, the repeated measures ANOVA revealed no time X group interaction, with a small effect size, $p=.80$, $\eta_p^2=.005$. Similarly, the differences in performance of the three groups on the performance test was non-significant for result clauses, $p=.40$, and for condition clauses, $p=.43$, both with small effect sizes.

## 4.2 Qualitative studies

### 4.2.1 Case Studies

The inferential statistics in the previous section had the benefit of being able to provide broad generalisations between and within groups. While this in itself is valuable, by definition the analyses do result in generalisations, and do not uncover the depth of individual performances. These case studies from five learners, therefore, offer a thicker description of the data by allowing the exploration of the use of input in task performance in general, and the analogical processes, in the form of repurposed operations, in particular. It also allows some insight into the subjective processes that accompany these performances, by cross-referencing performance with stimulated recall data. By uncovering aspects of learners' productions which the statistical analyses only partly investigated, these case studies provide additional data for the research questions

and shed light on the variability encountered in the quantitative analyses. The performances in these case studies are taken from both the Analogical and Literal groups, and were chosen as it was felt that they were representative of different types of task productions[6].

The first three case studies are presented as expert performances, although each exemplifies different traits: in Case Study One, there is a mix of both the exact use of selections and the repurposing of selections, while Case Studies Two and Three show more formulaic use and more repurposing respectively. Case Studies Four and Five offer task performances very different to the first three case studies. Case Study Four shows no gains between pre- and post-test, very few repurposing operations, but a large number of selections of past counterfactuals from the input. Case Study Five again shows no gain scores, no past counterfactual selections from the input, and therefore no use of past counterfactuals in the task performances.

#### 4.2.1.1 Case study one

Table 4.25 shows the statistical profile of ST14, a Thai learner progressing on to a PG course, along with transcriptions of her task performance in Table 4.26. To the right of the transcript is an analysis of the role input plays in her performance.

As can be seen in the statistical profile, ST14 achieved gain scores on both the acceptability judgment test and the production test. Of particular relevance is the gain in the quality of result clauses in the production test between pre-test and post-test. At pre-test, she produced nontargetlike utterances with counterfactual modals but no reference to past time. For example, in the third sentence in her pre-test in Table 4.26, she uses *'graduated'* and *'He could be studied'* (See Excerpt 1 in Table 4.26, sentence three), which suggests that she is

[6]As commented on in Section 4.1.4.1, although methodologically the two groups are distinct, in practice a number of participants in the Literal Group went beyond the input by repurposing Input in their own interpretation of the task. For this reason, both the Analogous and Literal groups are considered here.

attempting to encode past time reference in the sentence. This attempt at past time reference is supported by the rubric in which the protagonist of the story is dead. The use of modality with no past time reference is evident throughout her pre-test, resulting in six uses of modal plus no past time references and one nontargetlike use of modal plus past time reference (See Excerpt 2 in Table 4.26: *'He could met his family'*). The picture is very different in the post-test, which sees her result clauses move from modal plus no past time reference to targetlike performance (3 and 4) and the use of modals plus nontargetlike reference to past time (5, 6 and 7).

Moving to her task productions, Table 4.26 shows on the right-hand side tokens of past counterfactuals that were selected from the input. On the left-hand side are the task productions with underlined tokens corresponding to the tokens selected from the input. There is only one token in the task production identical to the selected tokens (Excerpt 11), although there are several utterances which have been modified. Excerpts 8, 9, 13, 14, 23 and 27 are based heavily on the selections from the input but with changed objects. Similarly, 15 and 21 have significant similarities to the selected input, but with a change of subject. There is also evidence that with some tokens both the subjects and the objects have been repurposed (18, 20, 21 and 29), but the phrase *'would have been'* has been retained. Other instances, 10, 12, 16, 22, 24 and 26 for example, are further removed from the input with only *'could have'* and *'would have'* retained. Of course, it could be argued that these were creatively constructed from prior knowledge as the modal plus auxiliary *'could have'* and *'would have'* are common to the vast majority of past counterfactuals. The pre-test, however, shows no evidence that she had control or knowledge of these structures in an obligatory context. It could be further argued that all that ST14 needed was past-counterfactual tokens that prompted her to allocate prior knowledge to this context. This could indeed be the case, as in the stimulated recall below (Excerpt 107), she does comment on the patterns being familiar when she men-

tions her teacher 'teach me many years ago', but she 'can't remember about them' because 'it's not in use in daily life'.

The stimulated recall also illuminated the conscious processing that she employed to move from input to her own production. In the parts underlined, it appears that she had an understanding of the pattern of which the instances were exemplars, and was able to 'put [her own] words in the patterns'. This supports what was seen with her repurposing of selected phrases in the task productions in Table 4.26 (See also 4.2.2 Stimulated Recall Interviews).

107. Stimulated recall excerpt for ST14

AH=Researcher

| | |
|---|---|
| AH: | So to fit your plan. You had your plan,<br>and you had the reading, and you put these together? |
| ST14: | Yeah, yeah. I... uh... mostly I use<br>the pattern, the sentence,<br>the pattern, after, the pattern of sentence, but my word. |
| AH: | But your words. |
| ST14: | Yeah, put my words in the patterns. |
| AH: | Yes, yes, yes. And before you did these tasks<br>did you know the patterns? |
| ST14: | No. |
| AH: | No? |
| ST14: | Maybe just the teacher teach me many years ago,<br>but I can't remember about it, because it's normally<br>it's haven't, it's not in use in daily life. Yeah. |

**Table 4.25:** *Statistical profile for ST14 for production and acceptability judgment tests.*

| Measure | Pre-test | Post-test |
|---|---|---|
| Acceptability Judgment Test[a] | 55 | 69 |
| Production Test | | |
| Condition Clauses[b] | | |
| OR[c] | 0 | 0 |
| AND | 1 | 0 |
| TLU | 0 | 2 |
| Result Clauses[d] | | |
| OR | 6 | 0 |
| AND | 1 | 3 |
| TLU | 0 | 7 |

a) Out of a possible 100

b) One obligatory context in pre-test, two in post-test

c) Or=signs of pastness or counterfactuality; AND=signs of pastness AND counterfactuality; TLU=Target-Like Utterance

d) Six obligatory contexts in pre-test, ten in post-test

**Table 4.26:** *Pre-test, post-test and task production for ST14*

Pre-test

. . . he had not move to London. Analysis from the data that he got the scholarship, he was an intelligence man. He could be studied in the 1st ranking university in Brazil, and graduated with high score[1]. He could be a high performance person in any business or government office. He might be married or stay single. Because of his intelligence and high performance profile, he could be a rich and successful man. On the other hand, his life might be more fullfill because it did not rush out and he could met his family members[2]. He might be more happier than living in London.

Post-test

. . . If he hadn't studied in the UK, his life would have been a lot diffent[3]. He would have graduated from the hightest university ranking in his country[4]. Because of his intelligence and performance, he could had had a good job[5], also in a good position. He could have earned a lot of money for his luxury lifestyle. He could been an outstanding person in society[6], so he could have married with a beautiful and intelligence woman. They would have spent their life together with happiness. If he had lived in his country, he would had had many relations[7]. He could have felt more comfortable than living in the UK. Unfortunately, he had died already, so he so he couldn't had spent in his luxury life.

| Performance | Selection from input |
|---|---|
| TASK ONE . . . I hadn't gave money to my friend, I would have been a very happy woman. I could have done useful things[8] not only for myself, but also for community. Firstly, if I had had a lot of money[9], I would have spent it for my life[10]. I have been able to buy new cars for myself and my parents. I could have travelled the world[11]. I have been able to invest for my own business, nursery, that fulfill my life. Secondly, if I had had money and time, I could have donated money for charity[12]. For example, I could sponsored in psychology research or started up my own foundation for special children. I have been able to give scholarships for poor students. Unfortunately, I gave all of my money to my friend and she ran away. If I hadn't had money, I couldn't done all of my wish.[13] | [8] I could have done all sorts of things<br>[9] If I'd had the money<br>[10] I would have bought a large house, I would have bought a boat, I also would have invested. . .<br>[11] I could have travelled the world<br>[12] I could have bought myself a company car, I could have travelled the world (plus nine other instances with 'could have')<br>[13] I could have done all sorts of things (plus two instance of 'wouldn't') |

TASK TWO ... If I hadn't studied in the UK[14], my life would have a lot different.[15]. My life skills, such as making decisions could have controlled by my parents.[16] If I had had problems, they would have solved them for me.[17] I could had own decisions and work out my own problems. I would have been a slow life person,[18] because I had had other person to help me such as cooking, doing housework and taking care of my son. Moreover in my career, if I hadn't studied in the UK, I wouldn't have an extra promotion from boss. I could be an ordinary person, and couldn't been out standing person. I probably would have had less connection with international friend.[19] Finally, my attitude of the UK would have had a lot of different.[19] I would have been visited here as a tourist. [20] If I hadn't studied in the UK, my life in career, life skills and friends would have been a lot of different thing.[21]

[14] If I hadn't studied in London
[15] things would have been a lot different
[16] Four instances of 'could have'
[17, 18, 19] Six instances of 'would have'
[19, 21] things would have been a lot different.
[20] four instances of 'wouldn't have been', things would have been a lot different.

TASK THREE ... If I hadn't made a right mess as I was younger, I would have had a normal life as other person.[22] If I had studied so hard in school,[23] I could have received a high school's certificated[24]. As a result, I could have futher studied, maybe in short vocational training.[25] I could have gotten a less complexed job such as a chef assistant[26], worker and cleaner. I could have earned enough money for my living and entertainment.[27] If I had saved money[28], I would married about bought a small house. I would have been more happy than now[29]. All in all, I done a gamble and mess of it all.

[22] Ten instances of 'would have'
[23] if I had studied a lot more at school
[24, 26] five instances of 'could have'
[25] I could then have entered a university and studied something
[27] and I could have earned good money.
[28] four instances of 'If I had...' and four instances of 'If I hadn't...'
[29] Two instances of 'would have been'

#### 4.2.1.2 Case study two

The statistical profile for ST4 (Table 4.27), a Russian PG student, shows a much lower pre-test score than ST14, and very little evidence for the ability to encode counterfactuality in pre-test production (Table 4.28: pre-test). As with ST14, the first sentence in the pre-test production (as well as the rubric of the task) suggests her intention is to show past counterfactuality (See Example 108). The second sentence, however, continues with the past simple. This is interpreted here as a lack of linguistic knowledge of the structure.

108. *. . . if Joey didn' go to UK and never been in London. So a 18 years old he tried to tak scholarship. . .*

The post-test, in a similar manner to ST14, shows areas of development, but unlike ST14, there is very little targetlike use apart from Excerpts 6 and 7 (note that the first sentence, Excerpt 2, is from the task rubric). There are, however, a number of differences between pre- and post-test towards the target language. In the condition clauses in 4 and 5, the use of *would* is erroneous but suggests a better understanding of how to signal counterfactuality in the past. The greater use of modals in comparison to the pre-test, in addition, shows use more in line with the expression of past counteractuality. There are also signs that she has taken the expression *could have* from the input, but was unable to attach a past participle. Overall, the performance gives the impression that she has acquired a number of items or chunks of past counterfactuality such as *could*, *could have* or *could have had*, and is developing the early stages of systematicity.

ST4's selections from the input and subsequent performance is considerably different to those of ST14 in terms of the repurposing of the selected phrases. For example, ST14 showed a large number of tokens that had been repurposed to fit the task demands. In ST4's production, many more selections are used wholesale in her production. In her condition clauses, 8, 15, 21, 25, 27 and 30 are full condition clauses from the input. There are also only slight changes to

the object in 13. Only one example (23) shows a change to the past participle, although the same phrase (*If I'd had...*) occurs in the input to Task One. The selection and rigid use of the condition clauses is mirrored by the selection and use of the result clauses. Excerpts 9, 11, 16, 21, 24, 26, 34 and 40 can be seen verbatim in the input, while there are minor changes to the objects of the clauses 17, 29 and 39. Interestingly, in 11, 18 and 32, it appears that she has taken the chunk *could have* but not attached a past participle, while on other occasions the auxiliary *have* has been dropped (28). On two occasions, *could have* is used creatively with a past participle (14).

There are similarities between how ST4 and ST14 described their processing of the input. In ST4's recall interview below, she mentions how she saw the structure of the past counterfactual clauses, and used her own words within the structure (double underlined). Something that ST14 did not refer to was the use of chunks of language; ST4 clearly used past counterfactual chunks in her own performances (single underlined), which seems to be echoed in both her task productions and her production.

109. Stimulated recall interview for ST4

AH=Researcher

ST4: The story was similar what I imagined about, and I used grammar structures, I use the beginning of the statements, and the...

AH: For example?

ST4: If I could have.

AH: OK, yeah. And did you use it exactly the same as the reading, or did you change little bits?

ST4: I take my own words, and look on the structure what can I use to my words.

AH: Ok.

**Table 4.27:** *Statistical profile for ST4 for production and acceptability judgment tests.*

| Measure | Pre-test | Post-test |
|---|---|---|
| Acceptability Judgment Test[a] | 18 | 38 |
| Production Test | | |
| Condition Clauses[b] | | |
| OR[c] | 1 | 0 |
| AND | 0 | 3 |
| TLU | 0 | 1 |
| Result Clauses[d] | | |
| OR | 1 | 3 |
| AND | 0 | 0 |
| TLU | 0 | 3 |

a) Out of a possible 100

b) One obligatory context in pre-test, four in post-test

c) Or=signs of pastness or counterfactuality; AND=signs of pastness AND counterfactuality; TLU=Target-Like Utterance

d) Fourteen obligatory contexts in pre-test, seven in post-test

**Table 4.28:** *Pre-test, post-test and task production for ST4*

Pre-test

. . . if Joey didn't go to UK and never been in London[1]. So a 18 years old he tried to tak scholarship but he couldn't do that. He was extremely worried about it and because of that he has a great mental disorder when he was 19. His family try to help him find a job but he didn't want to do anything. During next two years he drunk and leave his home. But 1970 he met his destiny. It was old woman from Chily and she Helped to him. The falled in love from each other. She had her own business and gave a job to joey. Everething was perfect, but in 1975 Sophia (chily woman) found another men. Joey was really disappoint of life and tried to find answer in religion. Today nobody knows where is Joey now.

Post-test

. . . he had not come to university in UK[2]. All his life could become very different[3]. He could have to find a job in Brazil and be a good person in this country. However, If he wouldn't have had some skills wich he had taken from UK university[4], he couldn't (XXX) a lot of money. May be in Brazil Joey could have to find some simple girl and got to marige with her. If he would have make a family and found a job he could have had a good life but not so perfect as in Brazil[5]. As the Best, in Brazil he could have had a perfect tan[6] and enjoed a good weather durin all year. Also he could have stayed with his family[7] and would be happy.

| Performance | Selection from input |
|---|---|
| TASK ONE . . . but if I hadn't done it[8], I could have done a lot of usefull things, and could had gone to different places.[9] At first, I had kept all my money on my bank accounts. Then I could have create some brilliant idea[10] and could have started my own business.[11] Best of all, I could have made my money to work.[12] If I hadn't spent all my time on business[13], I could have created some union of people and we all together could have made our country better.[14] If I hadn't given all money to my friend[15], she shouldn't have been bale to share it with the bank managere[16], and I would have felt myself with a great opportunities[17]. Morover I could have spend money to follow my education[18] and could have had a great success on my life[19]. In the end, I could have bought some special presents to my family.[20] | [8] If I hadn't done it,<br>[9] I could have done all sorts of things and gone to all sorts of places<br>[10, 14, 18, 19, 28 and 32] Eleven instances of 'could have'<br>[11] I could have started up my own business<br>[12] I could have made myself the managing director.<br>[13] If I hadn't spent it all on the business<br>[15] if I hadn't given all the money to my friend<br>[16] she wouldn't have been able to share it with Ernie the bank manager,<br>[17] and I wouldn't have felt stupid for so long.<br>[20] I could have bought myself a company car, I could have bought my mum a new house |

TASK TWO If I hadn't come to the UK when I was young thing would have been a lot different.[21] For a starem, If I hadn't gone to university in London, I would have been able to start education in my countre.[22 and 24] Probably, If I had had diploma[23] I would have been able to start job in my city.[24] If I hadn't moved to London[25] I wouldn't have studied english so much and made friend from all over the world.[26]. On the more positive side if I had stayed at home[27] I could seen all my friends and parents every day[28]. I would have been able to take some support from them.[29] I gained a lot of experiense. But if I hadn't come to UK[30] I would have been able to try to go somewhere else.[31] Probably I could have try to go to UK again but lik a tourist.[32]
To summurise it all I belive that I could be happy everywhere.

[21] I think that if I hadn't come to the UK when I was younger, things would have been a lot different.
[22] if I hadn't gone there
[23] Six instances of 'If I'd had...'; six instances of 'If I hadn't...'
[24] I wouldn't have been able to start a job in this area
[25] If I hadn't moved to London
[26] I wouldn't have studied English much and I wouldn't have made friends from around the world.
[27] if I had stayed at home
[29 and 31] I wouldn't have been able to...
[30] if I hadn't come to the UK

TASK THREE ... If I hadn't done this[33] everyething could be better. If I wouldn't have had a criminal record it would have found it a lot easier to get a job[34] and more my own legal money. If I had gone to school[35] I would have passed all my exam and go to university[36]. With university degree my cv would have looked better[37]. If I whould have a job I could have found I wife [38]. This would have change all my life.[39] If I had a family I could have felt myself more happy. Also I had a gambling problem, when I was younger. Without it I would have been able to save money[40] and bought a house. And I would have led to a better life with more happyness.[41]

[33]If I hadn't done this,
[34 and 38] I would have found it a lot easier to get a job.
[35] If I had gone to school
[36]I would have passed my exams
[37]my CV would have looked a lot better
[39]This would have changed everything
[40] I would have been able to save more money
[41] and maybe this would have led to a better life with more happy times.

#### 4.2.1.3 Case study three

The profile for ST7a, an Iranian PG student, is outlined in Table 4.29 with pre-test, post-test and task performances in Table 4.30. In a similar way to both ST4 and ST14, the acceptability judgment test results show positive gains. In terms of the production test, there are similarities with ST14, as both move from the use of modal verbs with no past time reference in the result clauses of their pre-test (Excerpts 1 and 2 in Table 4.30) to significant targetlike use in their post-test (Excerpts 3 to 9 in Table 4.30).

There are, however, differences between ST7a and the previous case studies. The most marked difference is between ST7a and ST4. ST4 stayed very close to the input, with the majority of clauses being used wholesale. With ST7a, on the other hand, at no point are phrases from the input used exactly. What ST7a seems to do is select phrases from the input and reuse *'would have'* and *'could have'*, but bring in his own subjects, objects and past participles to fit his intention (Excerpts 12 to 18 in Task Two, and 19 to 28 in Task Three, in Table 4.30). The argument could be made, as before, that the participant has prior knowledge of the grammar and that he is not using the input to form a new pattern. This line of argument, however, seems unlikely for two reasons. Firstly, his responses in the stimulated recall suggest this area of language is new and that he formed patterns from the input itself (See 110).

110. Stimulated recall interview for ST7a

AH=Researcher

| | |
|---|---|
| ST7a | And there is some kind of grammar here<br>I have never. . . uh. . . learnt before. |
| AH: | Which one, which bit? |
| ST7a: | I could have helped. |
| AH: | You'd never learnt that before. No?<br>And did you use that (AH points to the input)<br>in this (AH points to ST7's production)? |
| ST7a: | Yeah, because I understand that from the text. |
| AH: | Ah, from the text. |
| ST7a: | Yes, for example if I hadn't spent it all on |

the business I could have helped my family.
Maybe I would have helped my family.

Secondly, if he had prior knowledge of the pattern, it would be less likely that a marked trajectory would be seen across tasks. This trajectory, however, is clear in Table 4.30. In the first task, there are very few targetlike utterances; in Task Two there are more, and in Task Three, there are several targetlike result clauses. Similar to both ST4 and ST14, ST7a shows distinct signs that he explicitly grasped the structure of the result clauses, and either used this abstract structure to generate novel utterances (single underlined), or repurposed the selections (double underlined).

111. Stimulated recall interview for ST7a

AH=Researcher

AH: Can you give me some examples of
the words you used exactly?
I have never... uh... learnt before.

ST7: Exactly? (ST7 reviews his selections) I had not came,
he hadn't, I had not came. I... I would have, for example
... uh... he hadn't had the criminal record, he wouldn't
have got a very good job, also he would have been a rich man.

AH: So is this exactly from here (AH points to the input),
or have you changed it?

ST7: No, not exactly, I mean exactly from here
(ST7 points to where he had used 'would have been' in
his own work) 'I would have been'.

AH: Oh, OK, and you changed the last word.

ST7: Yeah, I changed the adjective, the rich man,
or... because that is mine, I didn't take it from here.

AH: OK, OK, so you took the would have been.

ST7: Yeah, something for grammar, just...

AH: Just something for grammar.

ST7: To help me identify the grammar of the sentences.

**Table 4.29:** *Statistical profile for ST7a for production and acceptability judgment tests.*

| Measure | Pre-test | Post-test |
|---|---|---|
| Acceptability Judgment Test[a] | 31 | 55 |
| Production Test | | |
| Condition Clauses[b] | | |
| OR[c] | 3 | 5 |
| AND | 0 | 0 |
| TLU | 0 | 0 |
| Result Clauses[d] | | |
| OR | 4 | 0 |
| AND | 0 | 2 |
| TLU | 0 | 8 |

a) Out of a possible 100

b) Four obligatory contexts in pre-test, five in post-test

c) Or=signs of pastness or counterfactuality; AND=signs of pastness AND counterfactuality; TLU=Target-Like Utterance

d) Six obligatory contexts in pre-test, ten in post-test

**Table 4.30:** *Pre-test, post-test and task production for ST7a*

Pre-test
. . . he stayed in his own country, he would not learn English, and also he will ever achieved his dream to move to UK, that might make him sad all his life. In addition if he was in Brazil, he would not get married his love[1], and if that happened to him, he might not have any children[2]
further more if still in his own country he will not be able to be assistant director, also he was not. But finally he was lockey man when he managed to move to UK.

Post-test
. . . he did not go to UK, he would not have studied a BA in business management at London University[3]. also he could not have met and lived with a lot of overseas students[4]. If he did not go to United Kingdom, his English would not have improved, and he could not got a really high mark in English Language[5]. In addition to this he would not got his first course in 1969[6]. If he was not able to speak English very well, he would not have got a good job in a small export company in London.[7] Furthermore, If he did study in London, he could not have met his love, and he could not have got married in 1970 and may be he would not have had his first daughter, Ramona.[8] If he did got a good job he would have been a poor man.[9]

| Performance | Selection from input |
|---|---|
| TASK ONE . . . I did not give all my money to my friend I would buy many incredible things and I would do many activities, such as travelling, shopping, also If I did not know that bad friend, all of that problems would not happened to my. if there had been some one to advice me[10] , I would still all my money and I had achieved all my dream, In addition if I was thinking and arranging very well, I would have been very happy[11] and I feel very happy. if my friend did not improvise me, I would be able to make my private business, Actually I have always wanted to travell to USA but if that did not happen to me I would had gone. I am usually thinking of buying a new big mantion, and I. . . | [10] If things had been different<br><br>[11] it would have been a lot better. |

TASK TWO ... If I did not go to UK, I would stay in my own country, In addition I could not have seen the Great Britain[12] . if I had not passed my main course in United Kingdom [13] , I would not have got a good job[14] . Furthermore, if I had not gone to the UK[15] , I would not have seen my incredible places there[16] , But fortunately I did go to some of these places such as, London eye, Oxford Circus and a big ben as well. One of most interesting point is that, If I would stayed in my country, I could not have seen and played in snow[17] . It is believed that, the UK has a strong reputation for higher education. So, if I did not study there, I would get an ordinary or a small job in my own country. In addition I made some relationship with some international and local student, when I was there, but if I had not made it[18] , I could not have known some cultures[19] . Moreover, My English is not too bad, because It has been evolved, due to speak to people too much, If I did not do that, I would not be able to speak that language.

[12, 17 and 19] 11 instances of 'could have'

[13 and 15] 3 instances of 'If I hadn't...'; 3 instances of 'If I had...'.

[14 and 16] 2 instances of 'wouldn't have...'; 3 instances of 'would have...'

[18] If I had made myself the managing director

TASK THREE ... he had not had the criminal record[19], he would have got a very good job[20], also he would have been a rich man[21]. also he used to fail his exams at school, when he was young, but if he had been punctual person[22], many things would not have happened[23], that is way he was not able to go to university, and he might be a successful in his life. In addition to this, he had played or gambled. So if he did not play, he would have gain alot of mony, and he could have bought whatever he like[24]. if he was intelligent he would have been a good man[25] and also, he could have got married, also he could have had children[26]. if he was rational person, he could have collected alot of money[27] rather than losing his mony. if he completed his studies, he could have been an efficient member in the society[28]. But unfortunately, if a lot of things have ever happened many changes would have occurred[28].

[19 and 22] 6 instances of 'If I hadn't...'; 2 instances of 'If I had...'

[20, 21, 24, 25 and 28]10 instances of 'wouldn't have'; 1 instance of 'would have'; 4 instances of 'wouldn't have been'

[23] Things would have been a lot different

[26]I could have bought a big flat and settled down

#### 4.2.1.4 Case study four

Table 4.31 and Table 4.32 show the statistical profile and task performances of a French UG student (ST8). In Table 4.32, there are also the selections which ST8 made from the input to the three tasks. From his profile, it is clear that no gains have been made either in the acceptability judgment test or in the production test. The initial frame in his pre-test *Imagine if... he stayed in Brazil all this life...* suggests that the following production will be counterfactual and in the past (See Excerpt 1, in Table 4.32). However, the use of first *will create* (2), and then *try* and *doesn't arrive* (3 and 4), and finally *will be* (5) makes both the factual nature of the production and the timeframe unclear. The post-test is clearly set in the past, as the past tense is used throughout (7-12), and the protagonist died (13). However, he shows little evidence that he can encode past counterfactuality in English. In Excerpt 6, *Imagine if... he wasn't in the UK* suggests counterfactuality, but it could refer solely to the past; the intention in 9 could be *he could have created* as the surrounding discourse refers to the past, although ultimately this is unclear.

In stark contrast to the previous case studies, ST8's engagement with the input is very limited: on four occasions (14, 15, 17 and 19) ST8 uses selections exactly, i.e. with no alterations, and on one occasion a part of an object is changed (18). There is one example of creativity within the verb phrase (16), in which the past participle is changed.

What is particularly interesting in ST8's work is the selections that he made. On all three tasks, he focused on several condition (e.g. *If I hadn't spent it all...*) and result clauses (e.g. *could have retired on the interest*). He also selects the full clause, which first confirms he has noticed the instances, and secondly suggests that he is aware the full clauses form a unit of meaning. However, although he uses some selections verbatim, he shows very little creativity with the phrases, and there is very little evidence that he can use these selected

instances analogously in order to generate other utterances[7]. This selection plus lack of development seems to fit with the earlier regression analyses. Recall that in these analyses, the selection of past counterfactual tokens explained very little of the variance in the acceptability judgment test and production test gains. Repurposing, however, explained a significant amount. While ST8 is only one example, that he selected many instances of past counterfactuals but did not manage to repurpose his selections suggests that his performances reflect the lack of correlation between the selection (and therefore the *noticing* of instances) and the development of counterfactual knowledge.

[7]Note that both ST8 and ST7b did not choose to take part in the stimulated recall interviews

**Table 4.31:** *Statistical profile for ST8 for production and acceptability judgment tests.*

| Measure | Pre-test | Post-test |
|---|---|---|
| Acceptability Judgment Test[a] | 42 | 41 |
| Production Test | | |
| Condition Clauses[b] | | |
| OR[c] | 1 | 1 |
| AND | 0 | 0 |
| TLU | 0 | 0 |
| Result Clauses[d] | | |
| OR | 1 | 1 |
| AND | 0 | 0 |
| TLU | 0 | 0 |

a) Out of a possible 100

b) One obligatory context in pre-test, one in post-test

c) Or=signs of pastness or counterfactuality; AND=signs of pastness AND counterfactuality; TLU=Target-Like Utterance

d) Five obligatory contexts in pre-test, nine in post-test

**Table 4.32:** *Pre-test, post-test, task production and selections from the input for ST8*

Pre-test
. . . he stayed in Brazil all this life he couldn't have the choice to learn and improve his English[1] . Whithout English today it's difficult to do commercial things abroad.
Maybe he will create his own company[2], because he looks like an enterprising person He try to learn English[3] but He doesn't arrive in Brasil.[4] His company will be very popular in Brazil but not abroad.[5] He lives a good life with his wife

Post-test
. . . He wasn't to the UK to study English.[6] He probably stayed[7] in Brazil and tried to learn English as much as possible but in different circumstance. He studied at university[8] of Sao Paulo.
He could try to create his own company[9] in Brazil in a new type and a new concept in restoration.
He tried to export his invention[10] all around the world but whithout English It was difficult for him to find negociation. He.
He finally diceded[11] to export his plan in Portugal because of the language.
He leaved in a good situation[12] with his daughter and their childrens in a big house near Rio de Janeiro.
He died at 82 years[13] old because of a car accident.

| Performance | Selection from input |
|---|---|
| TASK ONE . . . didn't give to him my money, all will be different, I have been too much (X) and I regret my choice. First with all this money, I took 4.4 million an bank to win the maximum of money with the interest. Secondly, with the 10 million remaining, I could have started up my own business[14] so directed the company and try to earn more money. | [14] started up my own business |

TASK TWO . . . I hadn't come the UK[15] , all my life would have been an other person.
First I would have been still with my family at home and I wouldn't have a good English I could have chosen to go to the US[16] to discover this country and improve my English. If I wouldn't speak English it will be very difficult to create my own company and working with other country. If I would stay at home, I probably studied economic subject and I probably never been adroad. To finish if I couldn't have the luck to improve my English are discover a new civilization and type of life.

[15] I hadn't come to the UK when I was younger
[16] I could have gone to USA

TASK THREE . . . He played a sport in competition at 14 to permited to have a hobbies to didn't have the idea to steal something.
After that if I don't missed school and if I was here at all clases I probably succeed an the exam. With my exam I could then have entered into a good university[17] and deepen my knowledge.
I could found a job more quickly because of my competences.
With that job, I could have earned lot of money[18] and not in a bad things.
If I hadn't gambled so much[19] I could be lots of thing with the money like travel abroad and spent my money differently

[17] I could then have entered a university
[18] I could have earned good money
[19] If I hadn't gambled so much

---

**Selections from Task One**
strated up my own business, kept only half of the money, I could have made myself the managing director, could have retired on the interest, directed the company, if I hadn't spent it all on the business, could have helped my family

**Selections from Task Two**
I wouldn't have gained my independence, If I had chosen a different course, I would have done a lot of different things, I could have gone to USA, I hadn't come to the UK when I was younger, I would have gone at home

**Selections from Task Three**
I would have found it a lot easier to get a job, I could then have entered a university, I could have earned good money, If I hadn't had a gambling problem, If I hadn't gambled so much

---

#### 4.2.1.5 Case study five

The statistical profile of ST7b, an Indian PG student, in Table 4.33 is similar to that of ST8 as both showed no gains on the acceptability judgment test or on the production test. As shown in Table 4.34, in both the pre- and post-tests, the past tense is used exclusively, with little reference to counterfactuality. In Excerpts 2 and 3 in Table 4.34, there are attempts to encode both the past tense with modal verbs (*can finished*, *can enjoyed*), but it is inconclusive whether the intention was *was able to* or *could have finished/enjoyed*. What is more clear are the similarities between pre- and post-test. Both employ the past tense and show no or few signs of counterfactuality.

ST7b's production (Table 4.34) is similar to ST8's, but it is a more severe case in terms of the type of tokens selected from the input. Firstly, the selected items mostly have little connection with counterfactuals. There is one example of a full phrase *wouldn't have been able to share*, and some examples of past counterfactual fragments such as past participles (e.g. *helped, travelled*), but in the majority of cases the selections are one, two or three-word noun phrases. Not surprisingly, in ST7b's task performances, there are no instances of the use of past counterfactuals from the input or repurposed phrases. Overall, the productions suggest that ST7b has bypassed the input and developed his own take on the task demands.

:

**Table 4.33:** *Statistical profile for ST7b for production and acceptability judgment tests.*

| Measure | Pre-test | Post-test |
|---|---|---|
| Acceptability Judgment Test[a] | 47 | 47 |
| Production Test | | |
| Condition Clauses[b] | | |
| OR[c] | 1 | 1 |
| AND | 0 | 0 |
| TLU | 0 | 0 |
| Result Clauses[d] | | |
| OR | 1 | 0 |
| AND | 0 | 0 |
| TLU | 0 | 0 |

a) Out of a possible 100

b) One obligatory context in pre-test, one in post-test

c) Or=signs of pastness or counterfactuality; AND=signs of pastness AND counterfactuality; TLU=Target-Like Utterance

d) Five obligatory contexts in pre-test, ten in post-test

**Table 4.34:** *Pre-test, post-test, task production and selections from the input for ST7b*

Pre-test
. . . Joey lived in brazil[1]. One of the finest country in the world he can finished[2] his study at his interest. he had better life[3] with his wife and children. All his families members missed when Joey died. he can enjoyed[4] his life with their parents. In another countries, you can not free your freedom like a home country. he was become a good business man[5], he might discover his langues in brazilian style. Its feel had someone died in younger aged becuuse he had just begun his carrer.

Post-test
. . . Joey was in UK, he lived there when he was 19[6] after that, he went for his further study, he completed his master[7] And start a new job with (X), he married with one woman and live his rest of life. They had have one children 'Romana', he had some difficuilties in his studies and language[8] he went to UDA, he had problem in (X) the courses, he got a new job. At later, he died in car accident.[9]

Performance

TASK ONE . . . I'm the luckiest person in the world 'YEAH!' I can do anything I want. I have 14.4 million. Firstly I have an idea so I'll create my business with my friends. I donate some moneys to build a new scelemimenty School for those are poor and can't afford it. I'm fascinating towards music so I opened a new music company and take franchise of 'HMV' I will bought a new house for my family, and take round a world trip in cruise. For example, I travelled a my bicycle for last 10 years now, I can but a new mclaren-mercedez car. I keep some moneys for my further education. I can depend myself, so my parents won't worry about my education.

TASK TWO . . . I think if I will not live in my country for a study it's quite a lots of problem and in (X) some and leant new experience. If I am away from house, It has lots of problem like accomodation problem, regular life style change, food habit, languages, etc, But, if you live with yours friends It maybe decrease your lonelyness. We have tried to cook yourself. You have to work for part-time. It very fast life probably most difficult experience in your life. My main (X) and all passport process to settled down.
It's more fun with independence live life you can hoose your own decisions and learn your mistake. Everybody wants better life to live.

TASK THREE . . . At this time, I've in a great position in my life. I had not stole that car, (X) should keep different things life learns a lot. if you make mistake you should not doing again change your bad things in yourself, life wants better idea if you have probably most importantly in your life is keep thinking positiv ways. For example I will do good in exam and get coud marks, I get proper job if I have some coud skills, to work, he should have a good job and honest life.
He can choose his favourite subjects that he likes and further studies, better thing and positive ideas, Repar your mistake you did.

---

**Selections from Task One**
Unfortunately, Obviously, Regret, all sorts of things, places, My own business, travelled around, helped my family as well, invested some money for both, holiday every year, for example, travelled on a bicycle, but best of all, felt stupid for so long, wouldn't have been able to share, I think that

**Selections from Task Two**
lots different, For a start, different course, economics and finance, in the area, of course, probably, join father's company, gained independence, own decisions, freedom, positive side, visa process, culture shock, travelled the, settled down, might chosen, Canada or USA

**Selections from Task Three**
But imagine, For a start, Probably, most importantly, easier to get a job, interested, subject, better, CV, good money, social life, gambling problem, no money, problem, better life, more happy times

---

#### 4.2.1.6 Discussion

The first characteristic that is evident from the case studies is the variability between learners. In the first three expert performances, there was an active engagement with the input, and a solid link between the input and their own production. This is not the case with the last two case studies, especially ST7b, who bypassed the input almost entirely and opted for his own interpretation of the task.

There is also variability between the first three case studies, although they are strongly related performances in terms of their engagement with the input. In the second case study, for example, ST4 heavily relies on the verbatim use of her selections, whereas both ST14 and ST7a use input tokens more constructively by producing tokens analogous to the input. The difference here could possibly be due to ST4's level of proficiency with the structure, which was noticeably lower than that of ST14 and ST4; staying closer to the input could well have been a crutch to lean on when there was limited productivity in her counterfactual knowledge. For ST14 and ST7a, the general feel is one of engagement with the input, but at a level which actively plays with the language, while still producing tokens that are at times heavily based on the input. ST14 is a good example here: there are some occasions when she uses the input wholesale, but in the majority of cases she performs operations on the input that allow her to match the input to her task demands.

Although there are a limited number of case studies documented here, that the participants ST14, ST7a and ST4 heavily engaged with the input and performed well in suggests a causal link. This is supported at least partially by the regression analyses earlier, which statistically linked repurposing operations to gain scores. Conversely and also supportive is the lack of engagement ST7b and ST8 had with the input, and the related poor performance on the post-tests.

In terms of the processes employed, both ST14 and ST7a comment explicitly

on the rules that they had inductively formed from the input. There is some evidence that there was prior knowledge of these rules based on previous formal instruction (ST14), although ST7a states that they are completely new, stating that he has inductively developed rules from the text. Whether his production was based on the rules he developed on selecting instances from the text, or he consciously drew analogies from the input, which subsequently led to conscious rules, is unclear. There is evidence that ST14 was well on the way to forming rules as she encountered instances in the input (One selection was *If I had+v3*), and had not laboured through creating analogies based on concrete instances. Ultimately, however, the extent to which conscious rules came before or after the repurposing operations or how rules, or rule-like behaviour, emerged from the process of analogy is not known.

### 4.2.2 Stimulated recall interviews

The purpose of documenting the case studies was to offer a thicker description of learner performances, especially in terms of the repurposing operations that they performed on the data. The stimulated recall interviews attempted to provide the learners' perspectives on these performances. More specifically, the coding of the interviews focuses on three areas. The first relates to the learners' subjective view of the selection of instances from the input, and the second the subsequent processing of these selections. This sheds light on why selections were made, and any explicit knowledge of the processes surrounding the repurposing operations on the input. The third area relates to the knowledge learners have of the linguistic target, especially in terms of the extent to which they could discern the pastness and counterfactuality of the constructions.

These three areas are broken down and reported in detail in Table 4.35.

**Table 4.35:** *Summary of stimulated recall interviews (raw counts)*

| Code | Raw Count[a] |
|---|---|
| Task demands and the selection of task-relevant input | |
| Task Demands affecting Selection of input | 18 |
| Input affecting plan/task demands | 8 |
| Processing of selected instances | |
| Chunking and juxtaposition | 11 |
| Abstract knowledge | 15 |
| Understanding of Past Counterfactuals | |
| Accurate knowledge of pastness[b] | 20 |
| Inaccurate knowledge of pastness | 6 |
| Accurate knowledge of counterfactuality[c] | 8 |
| Inaccurate knowledge of counterfactuality | 2 |

a) $n$=27 (14 from the Analogous Group (Group One), and 13 from the Literal Group (Group Two)
b and c) It is possible to have both accurate and inaccurate knowledge of pastness, as there exist both condition and result clauses, one of which may be considered a present tense, while the other is seen as past. In terms of counterfactuality, again, interpretations of the meaning of the result and condition clauses could differ.

#### 4.2.2.1 Selection of the input

Turning to the selection of instances from the input first, a recurring theme was the functional motivation behind the selection. For example, in Excerpt 112, it became clear that the planning and the task demands had established the idea of regret, and ST5 had mined the input for relevant instances to express these regrets.

112. Excerpt from stimulated recall interview for ST5

AH=Researcher
AH: Can you tell me why you selected the language? What were you thinking when you selected it?
ST5: I was thinking about a serious hurtful person, or a serious grateful person. . . not grateful, how to say the word? The one who feels shame for what he done.
AH: OK, regret, regret.
ST5: Regret
AH: Exactly. And can you give me examples of language which shows this?
ST5: If I haven't, which if I didn't done that, if I didn't do that, if I didn't win there, if I didn't give, and other languages.

A mapping of function to form, or form to function can also be seen in Excerpt 113. Here, ST7 uses both metalanguage and examples to describe the language he selected, and combines this with his interpretation of the past and counterfactual meaning: 'it's something your imagination to go back'. There are also signs that he needs the selections to complete his own task when he says 'I want talk about the situation it was like'.

113. Excerpt from stimulated recall interview for ST7

AH=Researcher
AH: So can you tell me something about what language you wanted, and how you used it?
ST7: Because I want talk about the situation it was like.
AH: Ok.
ST7: Past perfect tense.
AH: Ok.

ST7: So I. . . to put some sentences, like if I had, or I would have to, because it's something your imagination to go back.
AH: Ok.
ST7: So if I had this chance I would do like that.
AH: Yeah, yeah.

The link between the functional task demands and the selection of instances in the input is clear in Excerpt 114. In this excerpt, ST13 explicitly describes his lack of knowledge of the constructions he needs to complete the task. He also seems to have a clear idea of the meaning he would like to express when he mentions 'past' and 'imagine'.

114. Excerpt from stimulated recall interview for ST13

AH=Researcher
AH: OK, OK, fine. And then I gave you a, some reading, on a piece of paper, and you chose this language down here.
ST13: Yeah.
AH: Why? Can you tell me a bit about that?
ST13: Because I don't know, how can I describe? Of course I know in Japanese, but I don't know English how can I describe the past. Especially if I imagine something.
AH: OK, so you didn't know, when you were planning you didn't know this language, or you did know this language?
ST13: I didn't know.

In the following Excerpt (115), no direct reference to the meaning of past counterfactuals is made, but there is evidence that suggests that the selection is functionally motivated by the task demands and planning. For example, ST15 is looking for instances in the input that will help her with the organization of her own ideas. In Excerpt 116, ST6 selected instances in the input that were useful to his own production as 'they told me a lot, how to start with the sentence'.

115. Excerpt from stimulated recall interview for ST15

AH=Researcher
AH: OK, what were you thinking when you selected these pieces of language?
ST15: It helped me to organise the ideas.
AH: Ok.
ST15: To link one idea according to the topic I need to write.
AH: OK, and can you give me an example of something that helped you link the topic?
ST15: If I hadn't done it I could have travelled all my country.

116. Excerpt from stimulated recall interview for ST6

AH=Researcher
AH: Fine. And then you did the reading, and then you selected all of this language here.
ST6: Yeah.
AH: Why?
ST6: Because told me a lot, to how to start with the sentence, and...
AH: Can you give me an example?
ST6: Yeah, it's like if I hadn't done this...
AH: Yeah.
ST6: I would have to do this.
AH: Ok.

There are also a few cases in which the participants had changed their planning based on the input. In Excerpt 117, the planning was in the present tense, and on reading the text she changed her thoughts to the past. Interestingly, here there is evidence that the texts are used not solely as a resource to map on to planned meanings, but more prescriptively, as the 'correct' way to perform the task.

117. Excerpt from stimulated recall interview for ST4

AH=Researcher
AH: About the past. But your planning was all about the present.
ST4: The present, yes.

AH: So what happened there? What did you.. . . ?
ST4: Um. . . just I look how to write correctly and in the story was past, and I decide to use past.
AH: So you changed your plan to the past.
ST4: Yeah.

A further example comes from Excerpt 118. Here, again, the text alters the meanings that ST6 had intended. In a similar way, the text is seen not as a resource to map on to personal meanings, but as the 'suitable' way to complete the task.

118. Excerpt from stimulated recall interview for ST6

AH=Researcher
AH: OK. And two more questions, you said that your planning was in the future.
ST6: Yeah.
AH: Yeah, and where was your writing? Was it in the past, or future, or the present?
ST6: When I write I used the past tense.
AH: Yeah.
ST6: But it's. . . the idea is just it was becoming my brain, and I know that it's my future, I like that. But I don't know it's suitable for this writing or not.
AH: Ah, OK, so you planned for the future. . .
ST6: Yeah
AH: . . . but you wrote in the past with the ideas from your plan.
ST6: Yeah

#### 4.2.2.2 Processing of selected instances

This section considers the subjective experience of the learner while selecting and using past counterfactual tokens. The main themes that emerged were the use of tokens selected from the input to meet the demands of the task, and to a greater extent the explicit knowledge of past counterfactuals, either during selection of tokens or in task production.

There was some evidence that learners were using chunks of language in their production. For example, in Excerpt 119 the learner talks about using chunks at the beginning of her sentences. Whether this chunking is an example of what she refers to as 'grammar structures' is unclear. In Excerpt 120, juxtaposition of two concrete chunks can also be seen.

119. Excerpt from stimulated recall interview for ST4

AH=Researcher

AH: ...It's imaginary. And how did that, did that help with your writing? Did you use any of these phrases?

ST4: The story was similar what I imagined about, and I used grammar structures, I use the beginning of the statements, and the...

AH: For example?

ST4: If I could have.

AH: OK, yeah, yeah. And did you use it exactly the same as the reading, or did you change little bits?

ST4: I take my own words, and look on the structure

AH: Ok.

120. Excerpt from stimulated recall interview for ST16

ST16: Or ...don't exactly remember, but...or I mixed two sentence, for example, I have been able to ...I put, I would ...able to find a job, or I would have a lot of money.

However, many more learners comment on an abstract understanding of the input. This development of explicit knowledge can be seen in Excerpt 121 where ST7 suggests that he has gone beyond the selection and repurposing of the instances by developing a more abstract explicit understanding of how the past counterfactual functions.

121. Excerpt from stimulated recall interview for ST7

ST7: And second there's a very interesting point, I've chosen not words, just something (I 3.11) to the grammar. I didn't choose any word, if you see here, all of them are just tenses, not words.

This development of an abstract understanding is supported in Excerpt 122 by the same distinction between the surface structure *I could have given* and the underlying pattern: *just the grammar*. Although with different vocabulary this time, the same can be seen in Excerpt 123 with the use of *structure* and *construction*, and in Excerpt 124 with the use of *style*.

122. Excerpt from stimulated recall interview for ST14b

AH=Researcher
AH: OK, brilliant. So he should get, and I could have spent. Did you use these phrases exactly like in here (points to the selections), or did you change?
ST14b: No, just the grammar, and I change the verb of the sentence.
AH: If you can remember, can you give me an example of what you changed?
ST14b: I could have given, or something like that.

123. Excerpt from stimulated recall interview for ST12

AH=Researcher
AH: Yeah, OK, OK, and what about the difference between this one (points to selections in task one), and this one (points to selections in task two)?
ST12: Uh... this one, uh... I... it was not always, but we could find the same kind of expression, or...
AH: For example.
ST12: For example, the same structure. Construction.
AH: I wouldn't have...
ST12: I wouldn't have gained my independence.
AH: Right, OK.

124. Excerpt from stimulated recall interview for ST10

AH=Researcher
AH: OK, OK, OK, and did you take, did you take the sentences from here? Or did you take words from here, or..?
ST10: Yeah, yeah, yeah, I have taken some words and short, short words.
AH: Short words.
ST10: Yeah, because I think some, I think it's very beautiful verb,

and I take it, and I, if some... sentence there is very perfect, yeah, I... I will (I 2.48).

AH: You take it.

ST10: Not take it, it is just, you know, imitate.

AH: Imitate?

ST10: Yeah, yeah.

AH: OK, so did you imitate it exactly, or did you imitate approximately? So for example if I'd stayed at home, did you take that exactly, or did you change it?

ST10: Yeah, change it, change it. If they said, if I did stay at home, it's not (I 3.23) sentence, but I use this style.

AH: Oh, use this style, so can you give me an example of that, of how you use this as style, but you didn't use it exactly?

ST10: Uh, for example if I did stay at home, and if I had gone to be early, and if I had, if I had, if I have study in England, and I wouldn't, I wouldn't... something.

The most direct reference to explicit knowledge of a rule can be seen in Excerpt 125. ST14 had clearly seen the pattern in the input as in her selections from the input she includes what she describes as 'trees', e.g. *If I had+v3* and *I could have+v3*. Later on in the dialogue, she makes explicit reference to the operations she performed on the data based on the 'tree' that she developed from the input.

125. Excerpt from stimulated recall interview for ST14

AH=Researcher

AH: And did you take any language to say what you wanted to say? Did you take language from the... read... (reading)?

ST14: Yeah, yeah, I also take note about the pattern of a sentence, just have... if I had had ...

AH: Yeah, yeah.

ST14: Plus word trees, I like this (pointing to her 'trees' in her selections).

AH: I know, I saw, these three, these three.

ST14: Yeah, yeah. I (I 2.54) because it's very difficult, if I can, if I haven't record this one, maybe I could feel how to do, set them in order.

AH: And just two or three more questions. Did you, when you took that structure from the reading, did you change it to fit your plan? Or did you take it

exactly the same as the reading?
ST14: Ah, I changed, I change some words, it's quite simple, some words is the negative, the negative, if had not had something . . . negative and negative, but I change some, first sentence might be negative, the second a positive.
AH: OK.

As noted in the Case Studies section, there is an interplay between a number of variables. Learners are clearly explicitly selecting phrases, which they then use verbatim or repurpose to meet their needs. At some point in this process, a number of the learners are creating explicit rules. One question that this raises is whether learners first select, then repurpose, and subsequently develop explicit knowledge of rules, or first select, then inductively develop rules and subsequently use these rules to repurpose the selections. From Excerpt 125 above, it would seem that ST14 generated a rule in the process of selecting tokens, not as a consequence of the repurposing the tokens. However, a recurring theme in the quantitative results was the negligible correlation between selection and gain scores. If rules had been inductively generated from the instances, this would presumably have led to a correlation between selection of tokens and gain scores. However, it was the repurposing of tokens that significantly predicted gain scores, which suggests that any rules were consciously formulated during the repurposing of the tokens and not before.

#### 4.2.2.3 Knowledge of past counterfactuals

One concern throughout the study is the extent to which the students had grasped the meaning of counterfactuals, or had selected instances of counterfactual, repurposed instances of counterfactual, but had not explored how pastness and counterfactuality were encoded in the form. Therefore, this section reviews statements related to how participants viewed the meaning of the target form.

Some participants, as far as can be gleaned from the interviews, seemed to have a good grasp of pastness and counterfactuality. In Excerpt 126, for instance, the counterfactual meaning comes out on a number of occasions by juxtaposing the counterfactual with the factual.

126. Excerpt from stimulated recall interview for ST7

AH=Researcher
ST7: Yes, for example if I hadn't spent it all on the business I could have helped my family. Maybe I would have helped my family.
AH: Ok.
ST7: Because no, I could have helped, he didn't help his family, he would like to help, or he ...
AH: I know what you mean.
ST7: Yeah, I like to help.
AH: Yeah.
ST7: For example when you cut something up and you could have hurt yourself like that, oh, I could have hurt my ... you didn't, but you could. Yeah, that's what I understand from this.
AH: And you understand that from that, just reading that.
ST7: Yeah.

In Excerpt 127, although the participant does not have in this instance a full understanding of how to form counterfactuals, it is clear that she sees the meaning as counterfactual when she uses the terms *hypothetical, unrealistic and imagine*. In Excerpt 128, the learner is explicit in her understanding of the meaning.

127. Excerpt from stimulated recall interview for ST15

AH=Researcher
AH: OK, OK, fine. And do you see this as hypothetical, or do you see this as factual?
ST15: Hypothetical, because it's an unrealistic situation, I can imagine could be happened, if I do something before.

128. Excerpt from stimulated recall interview for ST4

AH=Researcher
AH: OK. And this language here (AH points to her selections), is it about real situations or imaginary situations?
ST4: Imaginary.

There are also cases where understanding of the past and the counterfactuality come together in the same excerpt. In both Excerpts 129 and 130, the learners are clear in their interpretation of both pastness and counterfactuality.

129. Excerpt from stimulated recall interview for ST5

AH=Researcher
AH: OK, and is this language about the past, or the future, or the..? I could have helped my family. Maybe I would have helped my family.
ST5: It's about the past.
AH: And is it about a real situation, or an unreal situation?
ST5: It's unrealistic but ...
AH: Did it happen? That's what I want to say.
ST5: No it didn't, it did not happen.

130. Excerpt from stimulated recall interview for ST13

AH=Researcher
AH: OK, and this language, do you know what it means? Do you know if it's about the future or the present or the past, or ...?
ST13: I know it's mean past.
AH: Past.
ST13: Yeah.
AH: And do you know imaginary or not imaginary?
ST13: So imagine.
AH: Imagine.
ST13: Yeah.

There were cases, however, where the meaning of other forms seemed to have polluted the meaning of the counterfactuals. For example, although in Excerpt 131 reference is made to the past, it is likely here that phrases such as *could have been, would have gone* have been interpreted as 'present perfect', due

presumably to the similarity between the forms, although this is not definitive; the reference to the present perfect could be a reference to the form *have + past participle*, not to the meaning.

131. Excerpt from stimulated recall interview for ST14

AH=Researcher
ST14: And I used the past.
AH: And you used the past.
ST14: And present perfect.
AH: Present perfect.

Metalanguage aside, on some occasions, it was clearer from the participants' descriptions that their interpretation was different to pastness and counterfactuality. In Excerpt 132, the counterfactual nature comes through, but her reference is to a present meaning. Similarly, Excerpt 133 translates *could have been* as the present tense.

132. Excerpt from stimulated recall interview for ST16

AH=Researcher
AH: Oh, I am going to ask you two questions. Is that tense about the past, the present, or the future?
ST16: Uh... present.
AH: Present. Yeah? The ones that you selected is about the present, yeah? OK. And is it real, or not real?
ST16: Uh... real.
AH: Real.
ST16: Not real?
AH: Yeah, is it real or not real? If I say if I hadn't gone to ... sorry, if I had gone to university, yes?
ST16: Yes.
AH: Is that talking about a real situation, or an imaginary situation?
ST16: Imagine.
AH: Imagine, yeah? So it's present and it was imaginary yeah?
ST16: Yeah.

133. Excerpt from stimulated recall interview for ST17

AH=Researcher
ST17: And I also, it was also good to see structure like this,
because I'm not very good with would or could, could have been,
don't know the tense, but I am not used to... use it, so ...
AH: Do you know what this means? Could have?
ST17: I could have. Yes, yes. LAUGHS
AH: Give me in French, what is it in French?
ST17: Je ne serais pas, je ne serais pas, je nous serais pas
[I wouldn't be..., we wouldn't be... c.f. Je n'aurais pas ete ...
(I wouldn't have been)]

Possibly connected to the confusion with the present perfect in Excerpt 131, Excerpt 134 sees the if-clause as referring to the past, but the result clause as having some consequence in the future.

134. Excerpt from stimulated recall interview for ST15

AH=Researcher
AH: OK. OK. And that language that you selected, do you know
if that, do you know what it means, that language?
Do you know if it's in the past, the present, the future?
ST15: If something happened in the past...
AH: In the past.
ST15: ... but has some consequence could happen in the future.
AH: In the future. So something happened in the past,
and it has a consequence in the future, yeah? So the ...
for example, if I hadn't done it, in your eyes is that in the past?
ST15: Yes, I don't it before.
AH: OK, and I could have done. Is this?
ST15: Possibly in the future.
AH: Possibly in the future. OK, great. And... so here you were,
were you writing in the past and in the future here?
ST15: Yes.

#### 4.2.2.4 Discussion

The idea of variability in the case studies is reflected in the stimulated recall interviews. This is especially the case in the knowledge of counterfactuals and pastness that participants had after the final testing had taken place. In some cases, there seemed to be a good knowledge of both pastness and

counterfactuality, while on other occasions, either counterfactuality or pastness in one or both of the clauses was misinterpreted. Clearly this is problematic as the statistical analyses focused on the form of the language, not its meaning, but evident here is a mismatch between form and meaning.

There is evidence in the stimulated recall interviews that the task demands and planning phase created a 'need to mean' (Samuda 2001), and that some participants were mining the texts for forms that would assist in how they expressed their meaning in their task performance. What was unexpected is that the texts had a dual function, one of satisfying a need for input, but also as something that guided or even changed the direction of the planned intentions.

Of particular interest here is the development of abstract rules, or types, from the input. As discussed earlier in the interview section, some learners were generating explicit rules at some point between planning and production. In some cases, the nature of the selections pointed to rule generation pre-production, i.e. the rules were being generated during the selection of instances, even though there was no task condition to do this. However, the selection of instances could not explain significant variance in the regression models, while the repurposing of tokens could. This suggests that the repurposing itself may well have led to explicit formation of rules, or both rules and operations were symbiotically related.

# Chapter 5

# Discussion

The literature review took as a starting point the proposal that grammatical development can at least in part be explained by a trajectory from initially entrenched formulaic chunks of linguistic material through partially abstract low-scope patterns to, depending on the construction in question, abstract constructions (e.g. Ellis, 2002; Tomasello, 2003). Although earlier empirical works such as Hanania and Gradman (1977), Shapira (1978) and Schmidt (1983) found little to support such a trajectory, more recent work by Eskildsen (2009, 2012) and Myles et al. (1998, 1999) demonstrate that when the methodological focus is on the breakdown of tokens from the environment not on the acquisition of abstract morphology formulaic chunks do seem to be developed into at least low-scope patterns. What is not clear from these later studies due to their longitudinal nature is why tokens in the environment are selected, how these selections are repurposed to meet communicative needs and what may be the outcome of such repurposing. With this in mind, the purpose of this study was to investigate the role analogical processing of selected tokens from the input may play in the acquisition of past counterfactual constructions. To this end, this section first discusses the extent to which the behaviour of the participants on the tasks can be interpreted as analogical processing, and to what extent the intervention in the study led to differences in the performance on the tasks, i.e. can the analogous group be interpreted as taking part in analogous processing and the literal group in literal processing, as defined in the literature review?

The discussion then employs both the results from the regression analyses and performances on the production and acceptability judgment tests to explore whether analogical and literal processing may lead to the acquisition of past counterfactual constructions.

This evidence from the behaviour of the participants in the study is then drawn upon to discuss the extent to which exposure to, the selection of, and the repurposing of tokens may correlate with the acquisition of past counterfactuals. These correlations are then discussed more broadly within a usage-based approach to language acquisition, in terms of whether the data support a trajectory from tokens to a partially abstract construction.

The data and the constructs of analogy and similarity are then interpreted within cognitive accounts of SLA such as Robinson's *elaborative processing* (2003) and Laufer and Hulstijn's (2001) *Involvement Load Hypothesis*, which predict that selection of tokens from the input and their subsequent *elaboration* may lead to the retention of formulas in the linguistic system. Of particular interest here is the way in which the notion of elaboration can be refined to include not only the rehearsal of tokens, but also the analogical processing of tokens.

Finally, the results are framed in pedagogic terms by bringing in the notion of task demands in Task-Based Learning and how the selection of task-relevant tokens can be seen as a resource for successful task completion.

## 5.1 Did the intervention promote analogical processing?

This section and *Section 5.2: The role of task demands in the breakdown of tokens* relates to RQs 3-4 and H2 on the analogous processing of tokens, operationalised as the repurposing of selected tokens. These research questions are reprinted in Table 5.1, along with the related hypothesis.

**Table 5.1:** *Research questions and hypotheses related to the exact and repurposed use of selected result and condition clauses in task production*

| |
|---|
| *RQ3: When learners are aware that task demands will require output that is identical to instances of counterfactuals in the input, does this lead to the repurposing of instances of task-relevant counterfactuals?* |
| *RQ4: When learners are aware that task demands will require output that is analogous but not identical to instances of counterfactuals in the input, does this lead to the repurposing on instances of task-relevant counterfactuals?* |
| *H2: The repurposing of tokens to fit task demands will be greater when task demands require output that is analogous but not identical to instances of counterfactuals in the input.* |

Gentner's structure mapping theory (1983) predicts that for an analogy to be formed, the relational structure of a base domain is mapped on to the relational structure of a target domain, and that this mapping is facilitated by similarities between objects in the base domain and the target domain (Gentner, 2003; Gentner and Rattermann, 1991). Gentner and colleagues (Gentner et al., 2003, 2004; Loewenstein et al., 2003, see also Gick and Holyoak, 1983) build on structure mapping theory by adding an explicit developmental angle referred to as *analogical encoding*. Here the analogical encoding of two similar cases results in the development of schemas from the common principles shared across the cases. These developing schemas in turn facilitate the subsequent categorisation of structurally similar cases.

Tomasello (2003) draws on and extends this theory to apply to a first language. From Tomasello's perspective, an analogy is formed if *both* relational structure such as the meaning X CAUSES Y TO MOVE $Z_{PATH/LOC}$ and categorial form such as S VP OB OBL are shared between the base and the target. At lower levels of proficiency when categorial form is not available, Tomasello opts for the term schematisation[1] to describe the abstraction of constructions as

[1]While Tomasello makes a two-way distinction between analogy and schematisation, Gentner et al. (e.g. Gentner and Rattermann, 1991) consider both to be forms of analogy with comparison between tokens based on similarity as predominant in novices, and comparison based on structural alignment predominant in later stages of development. In line with Gen-

a byproduct of the categorisation of two forms, citing the alignment of meaning and the similarity between surface forms as key factors in the mapping of base and target. Schematisation is illustrated in Figure 5.1, with shared attributes between tokens and predicted schematisation of differences between tokens

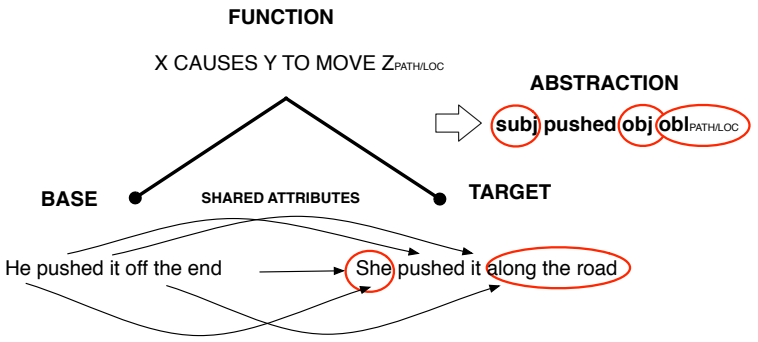



**Figure 5.1:** *Mapping of a base on to a target, based on Tomasello (2003)*

For the learners' behaviour in this study to be interpreted as analogical processing, there would need to be evidence that the form of the tokens in the output (the target) were derived from the selections in the input (the base), and that both the selected tokens and output tokens were mapped on to a past counterfactual meaning. Evidence for such a position comes from the exact repetition of selected tokens and also the repurposed use of selections from the input. With regard to exact repetitions first, Table 5.2 illustrates how the form of past counterfactual tokens in the output were heavily reliant on the tokens in the input, strongly suggesting that the tokens selected from the input were held in memory, and subsequently drawn upon to produce tokens in the output.

There is also considerable support for a link between target and base when there is a mismatch between input and output. Evidence of the link (see also Case Studies in Section 4.2.1) is illustrated in Table 5.3. The boxed elements are exact repetitions of the input, whereas the underlined sections illustrate changes. At the top of the table minor changes to adjuncts are apparent,

tner et al., no firm distinction is made here, although it is assumed that as the learners in this study are novices in terms of their knowledge of counterfactuals, similarity between tokens, *in addition to* functional comparisons dominates over purely structural comparisons.

**Table 5.2:** *Examples of exact repetitions of selected tokens*

| Type of operation | Base Token (selected from input) | → | Target Token (produced) |
|---|---|---|---|
| Exact repetition | If I hadn't done it | → | If I hadn't done it |
| | things would have been a lot different | → | things would have been a lot different |
| | I could've started up my own business | → | I could've started up my own business |

while more major changes can be seen with object and subject complements. Further down the table can be seen repurposing of the verb phrase and object complements, and finally the retention of *would have* with the changes of other elements within the token.

**Table 5.3:** *Examples of repurposing operations on selected tokens. Boxed elements are exact repetitions. Underlined sections illustrate changes between base and target.*

| Type of operation | Base Token (selected from input) | → | Target Token (produced) |
|---|---|---|---|
| Repurposing of adjuncts | If I had studied a lot more at school | → | If I had studied so hard in school |
| | I wouldn't have studied English much | → | I wouldn't have studied english so much |
| Repurposing of object complements | I could have earned good money | → | I could have earned enough money for my living and entertainment |
| | If I'd had the money | → | If I had had a lot of money |
| | I could have done all sorts of things and gone to all sorts of places | → | I could have done a lot of usefull things, and could had gone to different places |
| | I could have bought my mum a new house | → | I could have bought I could have bought some special presents to my family |
| | This would have change (d) everything | → | This would have change all my life |
| | I wouldn't have been able to make my own decisions | → | I wouldn't have been able to gain some freedom |
| Repurposing of subject complements | Things would have been a lot different | → | My life would have a lot different |
| | Things would have been a lot different | → | my life in career, life skills and friends would have been a lot of different thing |
| Repurposing of modal | I would n't have been able to start a job in this area | → | I would have been able to take some support from them |

| | | | |
|---|---|---|---|
| | I could have done all sorts of things | → | I could n't done all of my wish |
| Repurposing of past participle and object complements | I could then have entered a university | → | I could have futher studied, maybe in short vocational training |
| | I could have done a lot of different things with my life<br>I could have got a job in marketing<br>I could have earned good money<br>I could have bought a good place to live<br>I could have even had a better social life | | |
| | I probably would have ended up working for my father in his company<br>I probably would have studied a different course | → | I probably would have had less connection with international friend |
| | I would have had to listen to my father all the time<br>This control would have continued | → | I would have gone at home |
| Repurposing of subject, past participle and object complements | [combination of the above tokens in the input] | → | Finally, my attitude of the UK would have had a lot of different |

From the similarities between the input and output, it is highly unlikely that participants built from scratch tokens in their production that were an exact match for tokens in the input. The same can be argued for minor changes to adjuncts and changes to subjects and object complements, such as *I could have done all sorts of things* (base) →*I could have done a lot of usefull (sic) things* (target). When more changes are involved, it could be argued that the base input plays less of a role, with participants falling back on knowledge of past counterfactuals they had gained before the current intervention. While this cannot be ruled out completely, the knowledge displayed in pre-tests was non-existent in most cases, and minimal in only a few cases, suggesting that the production of past counterfactuals was based heavily on the input.

A considerable issue with the interpretations of the data as exemplars of analogical processing is the role that meaning plays in the process. For both Gentner and Tomasello, relational structure, such as X CAUSES Y TO MOVE $Z_{PATH/LOC}$ is a prerequisite to analogical processing. In this study, several factors pointed towards the interpretation that participants processed the tokens as exemplars of past counterfactuality. Firstly, the task demands included space builders such as *imagine*, *hypothetical*, and *the past* as well as considerable context which, if interpreted as such, were likely to induce a past counterfactual interpretation of tokens in the input. The textual analysis of the task productions also gave the impression that the output tokens were interpreted as having past counterfactual meaning. For example, in 135, although the participant has produced non-standard forms, the lexical reference to *the past* and the use of *if* suggests an understanding of past counterfactuality. This reference to counterfactuality is also evident in 136 where the participant explicitly states that the situation is *contrary to London* with the cotext also pointing towards an situation in the past.

135. I could have achieved success if there were no mess in my early life... My life could be better if i had no criminal activities in past.

136. I would have travelled around the world, perhaps, but I would not have meet a lot of friend, contrary to London.

Furthermore, the stimulated recall interviews often pointed to a past counterfactual interpretation, although not conclusively. In 137, for example, the task demands created a context in which the participant was forced to interpret his production as both in the past and imaginary.

137. Of course I know in Japanese, but I don't know English how can I describe the past. Especially if I imagine something.

If the evidence for analogy is taken as the mapping between a base form and a target form, and both target and base are interpreted as fulfilling the same function, then the selection and use of tokens can be framed in a similar way to Figure 5.1. This can be seen in Figure 5.2, which illustrates a mapping of past counterfactual forms selected from the input onto identical forms used in task production. In Figure 5.3, the same mapping is evident between forms selected from the input and an output form, although in this case changes have been made to the input token.

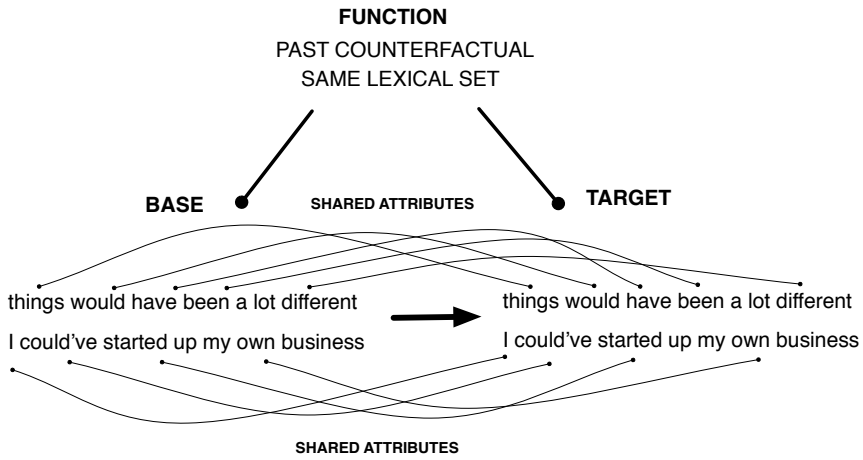



**Figure 5.2:** *Mapping base selections on to targets (exact)*

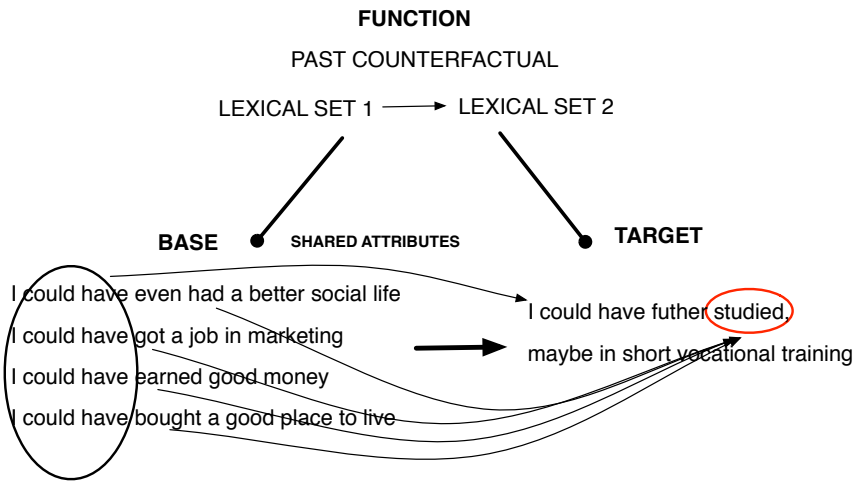



**Figure 5.3:** *Mapping base selections on to targets (repurposed)*

## 5.2 The role of task demands in the breakdown of tokens

While Figures 5.2 and 5.3 framed both the Literal and Analogous groups as interpretable within the same framework, the hypotheses predicted differences in terms of how the tokens selected from the input were repurposed to meet the task demands. More specifically, Figure 5.2 shows the relationship between the task demands and the use of exact tokens by specifying the overall constructional meaning (PAST COUNTERFACTUAL) and an identical lexical set for both base and target tokens (Literal Group). This is in contrast to Figure 5.3 in which the same constructional meaning is present but this time with different lexical sets between the base and target (Analogous Group). It was hypothesised that the relationship between the task demands and the input would drive the repurposing of the input for the Analogous Group, but not for the Literal Group. As is summarised in Table 5.4, this was borne out by the use of selections that were exactly the same as the input. Participants in the Literal Group mapped significantly more base tokens on to *identical* target tokens than the participants in the Analogous Group ($p<.005$, $d=1.03$ for condition clauses, and $p<.005$, $d=.97$ for result clauses).

**Table 5.4:** *Summary of the statistical differences between the use of exact clauses by the Analogous and Literal groups*

| Difference between groups | Significance | Effect Size ($d$) |
|---|---|---|
| Literal > Analogous (Result Clauses) | $p<.005$ | .97 |
| Literal > Analogous (Condition Clauses) | $p<.005$ | 1.03 |

The difference between groups is also borne out by the difference in the number of repurposed selections, although the results are more complex. As seen in Table 5.5, the Analogous Group repurposed significantly more elements of condition clauses, $p<.05$, $d=.49$, than the Literal Group, and more result clauses with a medium effect size, $d=.57$, although not at a significant level ($p<.1$). This lack of significance between the groups on the repurposing of result clauses could well be the result of the methodology used. Although the Analogous Group repurposed more tokens than the Literal Group, it became apparent that while the Literal condition was originally designed to push participants into literal processing (i.e. the use of selected tokens with no alterations), the tasks allowed for considerable repurposing of the input. This can be seen in the productions of a number of participants who used the literal input as a springboard for their own take on their performance of the task.

**Table 5.5:** *Summary of the statistical differences between the use of repurposed clauses by the Analogous and Literal groups*

| Difference between groups | Significance | Effect Size ($d$) |
|---|---|---|
| Analogous > Literal (Result Clauses) | $p=.083$ | .57 |
| Analogous > Literal (Condition Clauses) | $p=.028$ | .49 |

Interesting parallels can be drawn between the behaviour of the two groups and the studies by Myles et al. (1998) and Myles et al. (1999). Myles et al. conclude that communicative need drives the breakdown of initially rigid formulas. From this perspective, a communicative need which is beyond the

capabilities of the formula forces the analysis of relevant elements of the formula, i.e. a communicative situation that necessitates third person reference while the nearest relevant formula has second-person reference. The need for third person reference, they argue, forces the analysis of the pronoun slot to fit third-person reference. A consequence of recursive analyses of this type is the 'freeing up' of the second-person pronoun slot, which will allow for alternative subjects such as third-person pronouns and, in principle, other pronouns and inanimate subjects. Such a process can be interpreted within a analogous processing framework: pupils in the study by Myles were making use of established formulas in novel ways to meet the exigencies of the communication. If we assume, and it seems reasonable to do so, that in their study participants had memorised formulas that were mapped on to an appropriate, albeit approximate, meaning, e.g. [ASK NAME] = [*Comment t'appelles-tu?*] (Myles, 2004), and they had established the intention they needed to express in the current situation, the entrenched meaning of the formula would need to be mapped on to the the current intention, and the entrenched formula, in formal terms, would need to be mapped on to a token that expresses this intention. Interpreting the Myles et al. data in such a way leads to Figure 5.4.

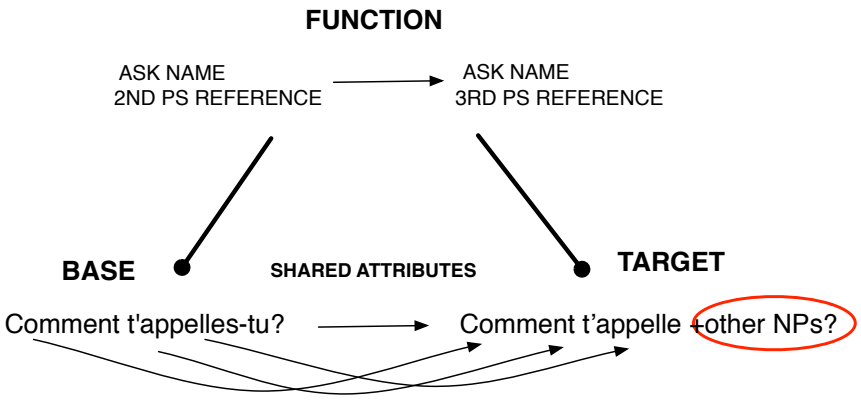



**Figure 5.4:** *Mapping of a base on to a target (data from Myles et al. (1998))*

Relating Figure 5.4 to the present study, the Literal Group were supplied with tokens that matched the task demands in terms of grammatical structure,

function and lexical items, and therefore did not have the need to break the tokens down to fit any additional demands of the task (although as discussed earlier some participants in the Literal Group did go beyond the task demands). This saturation of the task demands was not the case for the Analogous Group who were supplied with tokens that did not fully meet the task demands, and therefore, to complete the task successfully tokens needed to be partially analysed so that elements such as past participles could be substituted. What the results on the acceptability test seem to suggest is that this substitution of elements within tokens based on communicative need led to the ability to outperform the Literal Group, which, in turn suggests that the intervention led to a different type of knowledge to the Literal Group. This is considered in more detail in *Section 5.4: How abstract was the knowledge?*.

One difference between the study by Myles and this study here is the status of the base. In Myles, entrenched formulas can be interpreted as the base, which were subsequently broken down to fit the communicative need. In this study, novel tokens were selected from the input and subsequently repurposed to meet the communicative needs, suggesting that this repurposing can operate on newly registered tokens.

## 5.3 Does analogical processing lead to acquisition?

This section relates to RQs 5-6 and Hypothesis Three on the correlation between analogous processing and the development of past counterfactual knowledge. Both RQs 5-6 and H3 are reprinted below in Table 5.6.

That the task demands were a causal factor in the repurposing of tokens for the Analogous Group and the exact use of selections for the Literal Group, while interesting, does not suggest that the selection or repurposing led to the acquisition of past counterfactuals. As can be seen from the summary in Table 5.7, however, results from the acceptability judgment test do indicate that the

**Table 5.6:** *Research questions and hypothesis related to the development of past-counterfactuals*

| |
|---|
| *RQ5: Does the selection and use of exact tokens correlate with gains in past counterfactual knowledge?* |
| *RQ6: Does the selection and repurposing of tokens correlate with gains in past counterfactual knowledge?* |
| *H3: The selection and repurposing of tokens will lead to superior performance on post-tests than selection and exact use of tokens.* |

Analogous Group outperformed both the Literal Group and the Control Group. Specifically, the Analogous Group significantly outperformed the Literal Group with a medium effect size ($d$=.69), and while there was no significant difference between the Analogous Group and the Control ($p$=.09, with alpha levels set at .05), the Analogous Group outperformed the Control Group with a medium effect size ($d$=.58).

**Table 5.7:** *Summary of the statistical differences between the Analogous, Literal and Control groups on the acceptability judgment test*

| Difference between groups | Significance | Effect Size ($d$) |
|---|---|---|
| Analogous > Literal | $p$=.024 | .69 |
| Analogous > Control | $p$=.09 | .58 |
| Control > Literal | $p$=.927 | .10 |

As both groups were exposed to an almost identical number of past counterfactual tokens and selected a similar number of tokens from the input, the better performance of the Analogous Group on the acceptability judgment test suggests firstly that the exposure to tokens in the linguistic environment and the subsequent selection of tokens was not sufficient to explain any gains, since the Literal Group also benefited from this exposure and selection process. In support of this, the regression analyses showed that the partial and full selection of past counterfactuals from the input were poor predictors of gain scores

on the acceptability judgment test, with $\Delta R^2$=.02 for selection of full result clauses and $\Delta R^2$=.01 for the selection of full condition clauses (selection of partial result clauses had a even more negligible effect).

That the exposure to and selection of tokens from the input, while necessary, is not sufficient in the short-term acquisition of past counterfactual conditionals has interesting ramifications for motivation from the linguistic environment. Recall that both Wray (2002, 2007) and Ellis and Ferreira-Junior (2009a,b) cite type and token frequency of the input as drivers of acquisition, Wray in terms of the necessary evidence to 'loosen' elements in formulas and Ellis and Ferriera in terms of the correlations between Zipfian profiles of Verb Argument Constructions (VACS) in the input and acquisition and subsequent development of VAC prototypes. However, at least in very short interventions such as this one, sheer frequency of exposure in the linguistic environment would not appear to be sufficient, as although the input was heavily seeded with past counterfactuals (over 50 result clauses, and 30 condition clauses), and the tokens were registered (registration being demonstrated by selection), this level of evidence from the linguistic environment had little effect on post-tests.

However, in contrast to the Literal Group, the Analogous Group repurposed more tokens from the input to construct novel, task-relevant tokens in the output, and this repurposing was a significant predictor of gains on the acceptability judgment test with $\Delta R^2$=.13 (significant at $p$<.005) for the repurposing of result clauses and $\Delta R^2$=.17 (again, significant at $p$<.005) for the repurposing of condition clauses, even after initial test results and selections had been accounted for in the models. This strongly suggests that while exposure and selection are not sufficient to explain the results, the analogous processing of selected tokens, operationalised as the repurposing of selections, led to significant gain scores.

As summarised in Table 5.8, the results from the performance test would appear to paint a different picture to the one outlined above, as *both* the Analogous Group and the Literal Group showed significant within-group differences

between pre-test and post-test with large or medium effect sizes (the Analogous Group showed gains in the use of result clauses, $d$=.8, $p$<.0005 and in condition clauses, $d$=.43, $p$<.05, as did the Literal Group on result clauses $d$=.62, $p$<.0005 and condition clauses $d$=1.13, $p$<.0005). Both the Literal and Analogous groups also outperformed the Control on between-group measures of result clauses, with large effect sizes, Analogous>Control, $d$=.92, $p$<.005 and Literal Group > Control $d$=.94, $p$<.05. Both groups also outperformed the Control on the production of condition clauses, although not significantly.

**Table 5.8:** *Summary of the statistical differences between the Analogous, Literal and Control groups on the production test*

| Difference between groups | Significance | Effect Size ($d$) |
|---|---|---|
| Analogous > Literal (Condition Clauses) | $p$=.12 | .18 |
| Analogous > Control (Condition Clauses) | $p$=.3 | .24 |
| Literal> Control (Condition Clauses) | $p$=.04 (*n.s.*) | .42 |
| Analogous > Literal (Result Clauses) | $p$=.44 | .04 |
| Analogous > Control (Result Clauses) | $p$=.004* | .92 |
| Literal> Control (Result Clauses) | $p$=.008* | .94 |

*=Significant at alpha level 0.0167

While this similarity between the performances of the Analogous and Literal groups is at odds with the between-group differences on the acceptability judgment test, the results may well be the result of the type of test employed. The production test employed the same type of task as the intervention, and therefore it is plausible that selections from the input in the interventions could have been memorised and used verbatim in the production test. This rote use of single phrases would have been detected either by the targetlike or interlanguage measures, allowing for high gain scores. This suggestion that unanalysed formulas might have been used in the production test would seem to be at least partially borne out in the regression analyses based on the production test data, as the use of exact condition clauses in the intervention tasks was a good predictor of gain scores ($\Delta R^2$=.12), and also the selection of full result clauses

was a significant predictor ($\Delta R^2$=.1, $p$<.05). Conversely, for gains on the acceptability judgment test, the selection of past counterfactual tokens from the input in the interventions as well as the exact use of selections in task performance were poor predictors of gains ($\Delta R^2$=.02 for the selection of full result clauses, $\Delta R^2$=.01 for the selection of full condition clauses and $\Delta R^2$=.01 for the exact use of selected result tokens and $\Delta R^2$=.03 for the exact use of selected condition tokens as a predictor of gains).

As the counterfactual test items on the acceptability judgment test did not resemble the tokens in the input during the intervention, participants could not fall back on memorised formulas to determine the acceptability of the sentences in the test. While not conclusive, this supports the hypothesis that the Analogous Group outperformed the Literal Group on the acceptability judgment test because the Analogous group possessed more abstract knowledge and the Literal Group more lexically specific knowledge, and therefore the knowledge that the Analogous group developed was more likely to be able to 'sanction... novel expressions' (Langacker, 2000, p.26). The abstractness of knowledge is discussed in more detail in *Section 5.4: How abstract was the knowledge?*, which discusses an interpretation of the findings from a usage-based perspective.

## 5.4 Did analogical processing lead to abstract constructional knowledge?

In this section, the analogous processing of instances and gains on test scores are related to the proposed development of abstract counterfactual constructions. RQ7 is reprinted below.

**Table 5.9:** *Research question related to the development of abstract past-counterfactual constructions*

| |
|---|
| *RQ7: Do repurposing operations result in an abstract past counterfactual construction?* |

The theoretical underpinnings of this study draw on usage-based approaches which, in part, predict that tokens in the linguistic environment are categorised, with suitable type and token frequency, via categorisation based on analogy and similarity, and that this categorisation leads to the development of fully or partially abstract patterns. In more specific terms, the memory of entrenched formulas, or as is the case in this study recently-detected counterfactual tokens, allows for the categorisation of these tokens with either similar tokens in the input or with repurposed tokens. This categorisation of the established base and the novel target, it is argued, leads to the entrenchment of elements shared between base and target, and the schematisation of elements that differ, i.e. the categorisation of the base *I would have gone* and the target *I would have seen*, allows for the entrenchment of *I would have*, and the schematisation of *been* and *seen*. To explore to what extent this may be the case, it is useful to first summarise the main findings.

1. Both the Literal and Analogous groups showed similar knowledge pre-treatment, and selected a similar number of tokens from the input.

2. Both the Literal and Analogous groups showed significant and positive within-group differences in gain scores on the production test.

3. Both groups significantly outperformed the Control on gain scores in the production of result clauses in the production test with large effect sizes, and outperformed the control (but not significantly) on condition clauses.

Clearly, these results show some kind of gain in knowledge between pre and post-test for the Literal and Analogous groups, based on the selection and use of past counterfactual clauses. However, the acceptability judgment test showed a distinct difference between the groups:

4. The Literal Group showed no significant within-group gains on the acceptability judgment test. Conversely, the Analogous showed significant gains with large effect sizes.

5. The Analogous Group significantly outperformed the Literal Group in gain scores on the acceptability judgment test.

6. The full selection of past counterfactual clauses and the exact use of past counterfactuals in the tasks explained very little of the variance in acceptability judgment test scores. Repurposing of past counterfactual clauses explained a significant amount of the variance.

To explain the differences between the two groups, first, as argued in Section 5.3: *Does analogical processing lead to acquisition?*, there is the possibility that in the production test the groups could have memorised tokens from the input, which they then put to use in the production test, resulting in successful completion of task based on unanalysed tokens. The acceptability judgment test, however, used sentences that were unrelated in lexis to the input in the tasks, and therefore it would seem that more abstract knowledge than the retention of concrete tokens from the input was needed to classify the sentences as acceptable or not. These results suggest that analogies between input tokens and output tokens led to the development of partially abstract knowledge, illustrated as a byproduct to the right of Figure 5.5, which could be used to interpret unfamiliar tokens in the acceptability judgment test. If this is the case, then this would firstly lend support to Langacker's view that the simple categorisation of usage events and formulas/more abstract constructions in the learner's linguistic system results in the schematisation of related but not identical elements within the construction (Langacker, 2000).

To discuss further the question of whether drawing analogies between tokens in the input and tokens in the output led to low-scope patterns, studies from a usage-based perspective in FLA are useful comparisons, albeit with the proviso that there are clear cognitive differences between young children and adults. Empirical work by Dabrowska and Lieven (2005) and Lieven et al. (2003, 1997) traced the output tokens of two two-year olds and one three-year old to input tokens from care givers, to gauge the extent to which the tokens

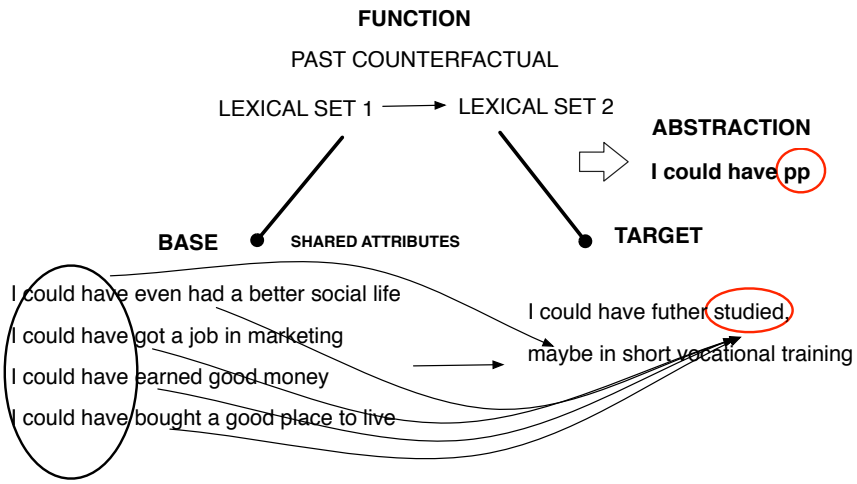



**Figure 5.5:** *Selections from the input, repurposed output and hypothetical low-scope past counterfactual patterns*

that the children produced were either identical to the input tokens or similar. Summarising the studies, they found that the majority of output tokens could be traced back to identical tokens in the input, and those that were not identical could be derived from previous utterances by means of just one 'substitute' or 'add-on' operation, e.g *Shall I* [PROCESS]? and *open that* are superimposed on each other to create *shall I open that?*.

Although the present study is focused on intervention in SLA not naturalistic FLA, there are clear parallels here between these studies and the present one. In this study and both Dabrowska and Lieven (2005) and Lieven et al. (1997), there were substantial similarities between the input and output, and where there did exist differences, the majority of these could be accounted for by the simple mechanism of substitution. Importantly, however, Lieven et al. (1997) warn against assuming that tokens in the output are indeed generated by a lexically-specific grammar at a very low level of abstraction (such as *Shall I* [PROCESS]), and prefer to see citing a lexical-specific grammar only as a favourable interpretation of the data.

> All we can say on the basis of the present study is that many of Annie's [one of the children in the study] novel utterances are closely

> related to utterances that have been said previously, that these previous utterances are highly frequent, and have the appearance of containing variable slots. And that, in so far as some of children's utterances are accepted as arising from the registration of strings in memory and the creation of slots in those strings, our results suggest that this may account for a considerable proportion of the data for an English-speaking child at this stage of development. (Lieven et al., 1997, p.366)

The same conclusion can be drawn for the present study; while the data would seem to point towards the development of patterns, ultimately the link between the selection of past counterfactuals and subsequent repurposing and the development of a lexically-specific schema such as *I could have X* is, like the studies by Dabrowska and Lieven (2005) and Lieven et al. (2003, 1997), an assumption based on the behaviour of the participants during the study and their performance on post-tests after such behaviour. It may be the case that the repurposing of selected tokens led to incipient interlanguage rules, in the sense of 'if x, then y', and not the emergence of lexically specific schemas. To investigate the psychological reality of a lexically-specific grammar with schematic slots, tests that probe the scope of past counterfactual representations would need to be employed.

Of relevance here is the thorny issue of whether any gains in the knowledge of schemas (or rules) was implicit or explicit, and what the relationship might be between implicit teaching, and implicit and explicit learning. In terms of the materials used in the intervention, it is clear that the focus was implicit, as there was no reference to any metalinguistic rules at any point. However, as implicit teaching does not take into account the agency of learners, implicit teaching and implicit learning can be seen as separable constructs (R.Ellis et al. 2009, Schmidt 1994). In support of this separation, the stimulated recall interviews clearly show that many of the participants could report knowledge of 'patterns', 'structure', 'construction[s]' and 'style' during their selection and use of tokens from the input.

As the acceptability test was designed to tap implicit knowledge of past counterfactuals, the question is whether this sort of explicit learning can lead to implicit knowledge, or whether both types of learning took place during the intervention. If we take implicit knowledge as the end point of the implicit tallying of frequency information (N.Ellis 2002; 2005a), and that the time scale for this sort of aggregation is in the thousands of hours (ibid.), then this relatively short intervention would be unlikely to produce any significant changes in implicit knowledge.

This disparity between the likelihood of implicit learning and the results on the timed acceptability judgment test is difficult to reconcile. One possible way out is to conclude that the timing of the test (9 seconds) allowed for the use of explicit knowledge in deciding the acceptability of test sentences. This could well be the case as 9 seconds is at the top end of the time frame predicted to exclude the use of explicit knowledge (Loewen, 2009). A second related explanation could be connected to the use of the term 'explicit knowledge' in second language research. The dominant description is to differentiate between explicit knowledge of instances, and the explicit knowledge of rules (e.g. R.Ellis et al. 2009, Schmidt 2001). However, the type of knowledge suggested here is a mixture of concrete morphological material and low-scope pattern (e.g. *would have* + past participles), not an abstract rule, such as verb-subject agreement or the dative alternation. It could be argued that mapping abstract metalinguistic rules such as verb-subject agreement on to concrete morphological material involves more processing, when compared to fast explicit access to both concrete lexical items and lexical items that filled the slots in a *would have + pp* construction. A reasonable conclusion from the data in this study therefore is that participants had explicit knowledge of low-scope patterns, not implicit knowledge of more abstract constructions, and that the acceptability judgment test allowed for search and selection of such patterns within the allotted time frame.

An interesting hypothetical corollary to the type of explicit knowledge cited above is the extent to which it might act as an attractor in further encounters with instances of past counterfactuals. This is essentially N. Ellis' (2005a) suggestion when he notes that the explicit learning of concrete linguistic material such as formulas, and the less concrete learning of slot-and-frame patterns guide the implicit learning of structure:

> Explicit memories can ... guide the conscious building of novel linguistic utterances through processes of analogy. *Formulas, slot-and-frame patterns*, drills, and declarative pedagogical grammar rules all *contribute to the conscious creation of utterances whose subsequent usage promotes implicit learning and proceduralization* (p. 305, emphasis added).

Although Ellis makes no distinction between formulas and slot and fill patterns (termed low-scope patterns here), the difference between the potential that both offer in terms of the categorisation of subsequent past counterfactual tokens could be considerable. If we follow the common distinction between the entrenchment of linguistic knowledge on the one hand, and the schematisation of linguistic knowledge on the other (Bybee, 2008, 2010; Hudson, 2008; Langacker, 2000, *inter alia*), a formula entrenched in the learner's mind has limited scope to categorise similar tokens in the environment, e.g. the formula *if it hadn't been for X* can possibly only categorisation almost identical tokens. The low-scope pattern *If it hadn't + pp X* can categorise many more. Therefore, if indeed the outcome of analogous processing is low-scope patterns, as suggested above, such knowledge, due to the level of abstraction, has the capacity to attract a larger number of tokens from the linguistic environment than formulas alone. Ultimately, however, the utility of the explicit learning of low-scope patterns over formulas is open to conjecture. Further research would be needed to determine whether, and how successfully, these types of knowledge acted as attractors to past counterfactual tokens in the input, and therefore the slow aggregation of implicit counterfactual knowledge.

## 5.5 Implications for cognitive accounts of SLA

Analogous and literal processing as defined here can be situated within cognitive accounts of language acquisition that propose that selection of tokens from the linguistic environment in addition to the mental operations that act upon the input correlate with their retention (Laufer and Hulstijn, 2001; Moonen et al., 2006; Robinson, 1995, 2003). Robinson (2003), in his elaboration of Schmidt's *Noticing Hypothesis* (1990; 2001) defines noticing as 'the subset of detected information that receives focal attention, enters working memory, and is rehearsed' (p.655). Of particular interest here is not the selection, and the issue of awareness at the point of selection - clearly, in the current study participants were aware of the tokens as they had selected them - but the rehearsal of tokens in working memory. Based on a distinction by Jacoby (1983) between data-driven and conceptually-driven processing, Robinson differentiates between maintenance rehearsal and elaborate rehearsal. Maintenance rehearsal involves the rehearsal of instances (or tokens here) in working memory without conceptual integration, whereas elaborative rehearsal involves some form of elaboration in terms of embedding or integrating with conceptual knowledge.

While such a model is explicit about the selection and rehearsal of tokens, one key issue is how these processes may lead to the development of more abstract constructions. In Schmidt's original formulation of the Noticing Hypothesis (Schmidt, 1990), he claims that the conscious detection, or noticing of information in the input is 'the necessary and sufficient condition for converting input to intake' (1990, p.129). In later developments of the hypothesis, Schmidt (2001) is very clear that this conversion from input to intake is at the level of surface features, not at the level of understanding. Noticing at the level of surface features is the perception of an item in the linguistic environment, while noticing 'at the level of understanding' evokes the existence of an underlying system:

> Noticing refers to surface level phenomena and item learning, while understanding refers to deeper level of abstraction related to (semantic, syntactic or communicative) meaning, system learning. (Schmidt, 2001, p.29)

This divorcing of noticing items in the input and system learning suggests that acquisition takes place in two distinct stages: the first is the noticing of the surface features and the second is the transformation of these discrete items into a natural grammar. If noticing is taken as the conscious detection of instances of surface form, and their elaboration (Robinson, 2003), as noted by Truscott (1998) and Truscott and Sharwood Smith (2011), this leaves the development of more abstract constructions open to question .

In a similar emphasis on elaboration in working memory, Laufer and Hulstijn (2001) draw on work by Craik and Tulving (1975) and Craik and Lockhart (1972) on the level and depth of processing of new information in proposing that the retention of vocabulary is conditional on what they term *task-induced involvement*. This involvement is broken down into three motivational-cognitive constructs: *need*, *search* and *evaluation*. Need is the motivational aspect of task involvement and is defined as 'a drive to comply with the task requirements' (p.14). Search and evaluation describe the search for a meaning of a word and subsequent 'attention to the form-meaning relationship' (ibid.), such as choosing a specific word and deciding if it fits a specific context or choosing one word over other related words. Using Laufer and Hustijn's examples, a task that involves reading and comprehension questions in which target words are glossed in the text and are not necessary for successful task completion has a low involvement load as the target items are not required to complete the task (-need), there is no requirement to search for a meaning of a target word as the words are glossed in the text (-search) and there exists no subsequent elaboration. However, a task such as writing a composition in which words are needed for the completion of the task (+need), and are then integrated into

the learner-generated discourse (+ search for meaning and + elaboration) has a high degree of involvement load.

There are clear parallels between the Involvement Hypothesis and Robinson's elaboration of tokens in working memory, as both models focus on the selection, elaboration and retention of tokens. However, such a focus on the retention of tokens raises the same questions as those raised by Schmidt's divorce between token and system. That is, if one selects and retains tokens, based on elaboration, what is the relationship between this selection and the development of more abstract constructions?

Relating the above discussion to the current thesis, on the basis that the performance on the tasks in the current study can be broadly separated into selection plus exact repetition to meet task demands (Literal) and selection plus repurposing to meet the same task demands (Analogous), the results, as a first approximation, can be interpreted as reflecting the amount of involvement in the task in terms of the number and types of operations that are performed on selected tokens. For the participants who selected tokens from the input and used them verbatim in the task production, it is assumed that a need was an integral part of the task as the tokens in the input were essential to successful task completion, and that participants searched for relevant tokens in the input that would perform the meanings that they had intended. In terms of elaboration, participants selected tokens that they then had to contextualise into their own take on the task demands, i.e. if they used a token verbatim, they needed to decide if it was appropriate for the concept that they intended in the production, and whether there was a fit with the on-going discourse. This in itself meets Laufer and Hulstijn's (2001) '+need, +search, +elaboration' conditions within their Involvement Load Hypothesis.

That the Literal Group made significant within-group gains on the performance test with medium and large effect sizes (result clauses, $d$=.62, $\rho$<.0005 and condition clauses, $d$=1.13, $\rho$<.0005) and significantly outperformed the Control Group with large effect sizes (Literal Group > Control, $d$=.94, $\rho$<.05)

tentatively supports Laufer and Hulstijn's (2001)'s position that with high involvement loads, items are more likely to be retained, although any stronger support would need a research design that varied the amount of load. It could be argued that the Involvement Load Hypothesis relates to individual vocabulary items, and not to larger tokens. However, the status of individual items is not seen here as qualitatively different to larger tokens. Tokens that were selected from the input were used verbatim in task production, and therefore it is assumed that no break down of tokens took place, and that the selections were processed as large chunks of material. This would seem to fit N.Ellis' description of formulaic phrases as 'big words' (Ellis, 1996, p.111) with no structure.

The actions by the participants who constructed analogies between tokens (i.e. the repurposing of tokens selected from the input) could be interpreted as satisfying the same conditions as participants who selected tokens and used them verbatim in their production, i.e. the tasks demands led to a need for the selection of items from the input that met their intended meaning, and that the selected tokens were contextualised into their own take on the task. However, there is clearly differences between the ways in which the Analogous Group and the Literal Group processed the selected tokens. That is, the literal use of tokens in task production involved the selection and exact use of tokens, and the analogical use of tokens involved the selection and the repurposing of selected tokens.

The difference between exact use and repurposed use is seen here as the type of elaboration involved. For analogical processing, elements within the selected tokens, e.g. object and subject complements, past participles and modal verbs, were replaced with different lexical items to meet the demands of the task. This replacement not only involves the selection and mapping of tokens on to concepts as in the Literal condition, but also a more complex form of elaboration than solely deciding how it fits into their production, i.e. the participants needed to compare tokens in the input with their intended

meaning to decide which elements within the token needed to be changed. One conclusion that can be drawn from this difference is that elaboration can be divided into two processes depending on the outcome of these processes. The first process is the mapping of meanings to tokens (individual items or longer formulas), as in Robinson's elaborate rehearsal and Laufer and Hulstijn's search plus elaboration, which results in a retained (or entrenched) token. The second type of process is the elaboration of tokens in working memory by drawing analogies between base and target tokens. The end product of such a process hypothetically is not solely the entrenchment of concrete tokens from the input but the creation of variables within the token, e.g. *Subj would have PP NP*.

One counter-argument for including analogy in the notion of elaboration is the argument that grammatical development and the retention of vocabulary items are two separate processes. If we assume, however, that there exists a cline between concrete lexical phrases and abstract constructions (e.g. Cowie, 1998; Culicover, 1999; Ellis, 2008; Goldberg, 2003; Howarth, 1998; Jackendoff, 2002; Nattinger and DeCarrico, 1992; Pawley and Syder, 1983), and that concrete exemplars develop into abstract constructions based on type and token frequency, then it is reasonable to discuss involvement load as relevant to constructional development. For example, Table 5.10 illustrates how the lexical item 'creation' may be processed within Hulstijn and Laufer's framework. Here, a learner presumably has no knowledge of how to express a concept, for example the lexical item 'creation', and therefore searches the linguistic environment for an appropriate token to express the intention. The token is then elaborated on by multiple decisions on if and how the token is appropriate for a certain context, meaning repeatedly mapping a token to a concept. Within a usage-based perspective, the result of this recursive mapping in the entrenchment of a form-function mapping based on token frequency. As mentioned above, longer tokens are seen in the same light, and therefore undergo the same entrenchment as individual lexical items (illustrated by the second example in Table 5.10).

Although elaboration can be seen as the entrenchment of either individual lexical items or 'big words', it can also be seen as developing a type. As illustrated in the third example in Table 5.10, if, for example, a token such as *give it me* or *I could have done it* is selected from the linguistic environment, but the intention that needs to be expressed is GIVE IT HER or I WOULD HAVE GONE, tokens would need to be generated such as *give it her* or *I would have gone* based on an analogy with *give it me* or *I could have done it*. Here, the emphasis is not on the entrenchment of individual tokens based on the frequency of the token, but on the comparisons of tokens as the same type. Taking this further, a learner could possess knowledge of the type *give it x*, and need to express the concept SEND IT THEM. If the token *send* is employed to meet the need, both *send it them* and *give it x* are categorised together as the same type. A usage-based perspective predicts that this repeated categorisation leads to more and more abstract types.

**Table 5.10:** *Type and token frequency and the Involvement Load Hypothesis*

| Type of selected token | Need | Example Selection | Elaboration | Product |
|---|---|---|---|---|
| Individual lexical items (based on Laufer and Hulstijn, 2001) | Item needed to express intended meaning in a composition (e.g. need: CREATION) | *Creation* | Evaluative: is this the best word for the current purpose; does item semantically and grammatically fit context in composition → repeated conceptual integration of new item. | Entrenched tokens |
| Tokens longer than the individual word | Token needed to express intended meaning in a composition (e.g. need: GIVE IT ME, I WOULD HAVE GONE) | *Give it me, I would have gone* | Evaluative: is this the best token for the current purpose; does token semantically and grammatically fit context in composition → repeated conceptual Integration of new item. | Entrenched tokens |
| Tokens longer than the individual word | Token needed to express intended meaning in a composition (e.g. need: SEND IT ME, I COULD HAVE DONE IT) | *Give it me, I would have gone* | Evaluative: is this the best token for the current purpose; does item semantically and grammatically fit context in composition.<br><br>What additional changes to elements in the token are needed to fit intention? → repeated conceptual integration of new item; repeated analogical processing of tokens+formulas. | Low-scope pattern based on processing *send it me* + *give it me* or *I would have gone* and *I could have done it* as the same constructional type |

## 5.6 Pedagogical implications

Framing selection and repurposing within a framework of analogous processing of newly acquired tokens and tokens in the linguistic environment is particularly relevant for research into Task-Based Learning in second language classrooms. One primary factor in any mainstream description of TBL is that there is some built in pragmatic outcome to the task (Bygate et al., 2001; Long, 1985; Prabhu, 1987; Skehan, 1996). In the Bangalore project, for example, Prabhu (1987) emphasised the notion that a task includes some form of non-linguistic outcome, and some form of activity is necessary to reach this outcome.

> An activity which required learners to arrive at an outcome from given information through some process of thought, and which allowed teachers to control and regulate that process, was regarded as a 'task'. (p.24)

A more recent definition of tasks by Bygate et al. (2001, p.11) succinctly brings together the primacy of meaning, and the use of language in the attainment of an objective.

> A task is an activity which requires learners to use language, with emphasis on meaning, to attain an objective.

A common theme in these descriptions of tasks is the idea of an non-linguistic outcome. This emphasis on the pragmatic use of language to reach an outcome is often criticised as generating highly contextualised, and therefore impoverished elliptical and lexicalised interaction (Batstone, 2002; Seedhouse, 1999; Widdowson, 1990, 2003). Seedhouse (1999) exemplifies this outcome-based nature of tasks by citing a NNS-NNS task production, with the aim of labelling a geometric shape.

L1: What?
L2: Stop.
L3: Dot?
L4: Dot?

L5: Point?
L6: Dot?
LL: Point point, yeah.
LI: Point?
L5: Small point.
L3: Dot.

(Lynch, 1989, p.124 cited in Seedhouse, 1999, p.153)

Seedhouse's argument is that language is used indexically to *point to* shared context, and as task-based interaction is highly contextualised, a minimal amount of language is needed to reach a successful outcome. He concludes this type of interaction treats 'linguistic forms [...] as a vehicle of minor importance' (Seedhouse, 1999, p.153). Both Batstone (2002) and Widdowson (1990) view this indexical communication as problematic from the perspective of language learning. In Widdowson's view, *language as communication*, i.e. a focus on a pragmatic outcome and the indexical use of language to achieve that outcome, 'denies [the learner] access to the necessary data for learning' (Widdowson, 1990, p.113). Similarly, Batstone (2002) sees indexical communication as necessary, but at the same time 'fundamentally opposed' (p.14) to learning needs.

The dichotomy between a focus on a pragmatic outcome on the one hand and a focus on linguistic forms on the other, with the learner choosing to divide her attention between the two, can be seen in much of the work on establishing a focus on formal elements in both task-based research and elsewhere. VanPatten (VanPatten, 1990, 2004), for example, argues that within a limited capacity attentional system a focus on form is in competition with a focus on meaning, with a focus on meaning as the default setting. From this perspective, if a balance is to be struck between form and meaning, a focus on linguistic form has to be artificially constructed. This competition is also noticeable in accounts of attention allocation during task performance. Skehan (Skehan, 1998; Skehan and Foster, 2001) argues that the degree of complexity of task demands

will deplete available resources, 'and what free attentional resources are likely to be left over ... could be directed towards noticing' (Skehan, 1998, p.51). Similar to the competition between form and meaning in VanPatten's account, Skehan sees the pragmatic focus of a task as disassociable from a focus on formal elements, with the task depleting the amount of attention available to the learner, and the remaining resources directed towards form. It is unclear, however, why, once the purpose of the task is achieved, attentional resources are directed elsewhere and what the purpose may be for directing these resources elsewhere.

An alternative view is that the pragmatic meaning established by task demands and a focus on form are not in competition but *complementary*. Ortega (2005), in a study which focused on attention to form during task planning, found that learners 'pay[ed] attention to the inextricable relationship between form and meaning, simultaneously holding in long-term memory considerations regarding the message to be conveyed and the essential formal resources to convey it' (p.106), and goes on to propose that *because of* the task demands, learners were directed towards form that would allow them to meet the task demands. Both Robinson's cognitive complexity of tasks (Robinson, 2001, 2007) and Samuda's 'need to mean' (2001) are relevant here as both view task demands as establishing the need for a pragmatic outcome, and that task demands drive the uptake of task-relevant resources from the linguistic environment.

In line with a complementary view of form and meaning, the participants in the present study selected tokens from the input such as *I could have done all sorts of things*, *If I'd had the money* and *I would have bought a large house*, which were then either situated wholesale into their own production, or repurposed to fit the task demands more accurately. Rather than these selections detracting from the production of the task, the selection of tokens from the linguistic environment can be seen as enabling the pragmatic interpretation of

task demands. In this sense the provision of tokens is seen as a *resource* that equips the learner with the tools for successful task completion.

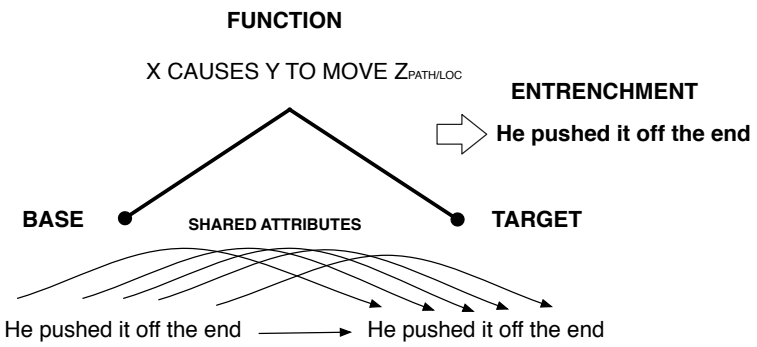
That participants in this study were using the tokens to meet the pragmatic task demands is also seen in the stimulated recall interviews, in which participants described how they selected tokens from the input that filled gaps between their conceptualisation of the task demands and their own production. It is also apparent from the interviews that some participants in the pre-task phase were actively going beyond the tokens to formulate their own conceptualisation of the task over and above the repurposing of selections by formulating their own structure.

While the selection of tokens based on the conceptualisation of the task demands is interpreted here as fulfilling a need for resources that allow for task production, a more fine-grained description would be to differentiate between the selection and exact uptake of tokens, and the selection and repurposing of tokens. Although the two share the same status as resource for task production, the repurposing involves additional work to match the conceptualisation of the task demands by forming analogies between input tokens and output. Tentatively, it is this that allows for an interesting interaction between task demands, language as resource and processes such as similarity and analogy that act upon both the tokens and related concepts to meet these demands.

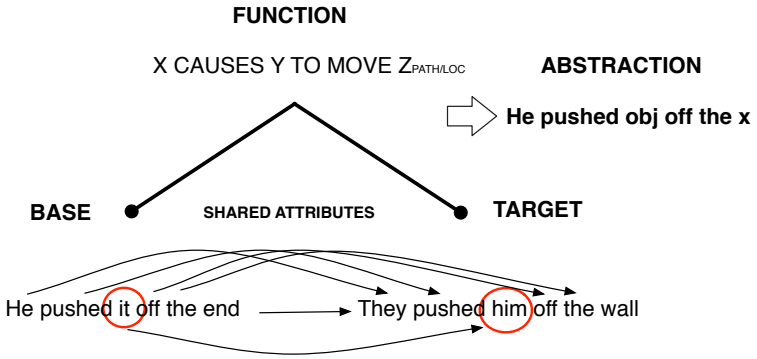
## 5.7 Limitations and future directions

The distinction in this study was made between task demands that attempted to push learners into the exact use of selected tokens on the one hand, and the repurposed use to fit the task demands on the other. While this was satisfactory for the aims of the current study, there is scope for a more fine-grained approach which grades functional distance between base tokens and targets. This type of grading is illustrated in Figure 5.6.

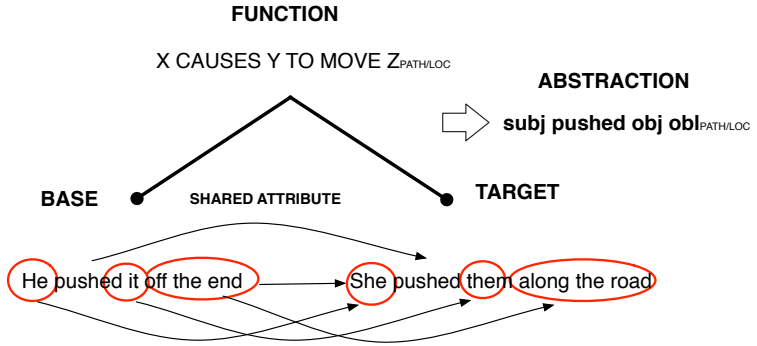
(a) *Target and base identical*




(b) *Differences between objects*




(c) *Differences between subjects, objects and obliques*




(d) *Differences between subjects, verbs, objects and obliques*

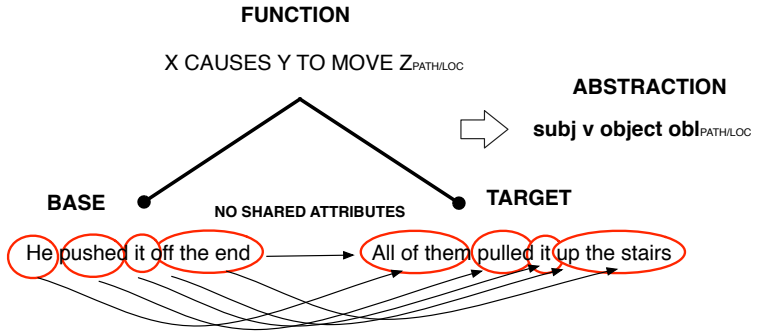



**Figure 5.6:** *From small to large differences between base and target*

In Figure 5.6a the task demands match the necessary output in the same way as the Literal Group in the current study. However, Figures 5.6b, 5.6c and 5.6d attempt to grade the distances between input and necessary performance under a constructional caused-motion meaning. A usage-based approach predicts that the outcome of each model of analogical encoding depends on the distance between tokens. For 5.6a the predicted outcome is entrenchment, while in 5.6b, 5.6c and 5.6d the greater the distance between the tokens the greater the abstraction.

Figure 5.6, as discussed above, suggests variable outcomes for the individual models. However, they can also be interpreted as sequential interventions starting with Figure 5.6a and ending with 5.6d. This progression attempts to map the development of constructions from formulas (or tokens) on to a 'career of similarity' (e.g. Gentner and Rattermann, 1991) from an initial stage that draws upon similarity between tokens and then based on the outcome of such a process draws more and more on analogies between tokens. As illustrated in Table 5.11, a first step in such a progression would be the development of tasks that initially entrenched tokens by requiring their selection and exact use in production (in a similar way to the Literal Group in this study). This would be followed by tasks that necessitated limited repurposing of selections to meet the task demands. Hypothetically, it would be at this point that formulas entrenched in the initial stage would be drawn upon and repurposed, allowing for a limited break down of the formulas. A final stage would necessitate more major repurposing of formulas, and therefore a greater type frequency. Differences in terms of constructional development with certain permutations of these tasks, e.g. 1-2-3, 3-2-1, 2-3-1 would offer evidence for the utility of an initial skewed type frequency in the anchoring of formulas (cf. Casenhiser and Goldberg, 2005; Elio and Anderson, 1984), and subsequent development of types based on established formulas.

From a methodological perspective, the current study was quasi-experimental, using intact classes, in the participants normal setting, which arguably offers

**Table 5.11:** *Task sequence from the entrenchment of formulas to the development of patterns*

| Task | Distance between base tokens (input) and necessary target tokens (output) for task completion | Type of process | Current knowledge | Product |
|---|---|---|---|---|
| 1 | Base approximately the same as necessary target for task completion | Categorisation of almost identical forms (base+target) under the same constructional meaning | None | Entrenchment of tokens, e.g. *he put it on, he pulled it off* |
| 2 | Task demands necessitate minor changes to base tokens in order to meet task demands | Categorisation of entrenched tokens and repurposed target same constructional meaning | Entrenched formulas (from tokens in task 1) | The development of low-scope patterns, e.g. *subj pull x off/on* |
| 3 | Task demands necessitate major changes to base tokens in order to meet task demands | Categorisation of low-scope patterns and repurposed target under same constructional meaning | Low-scope patterns (from tasks 1 and 2) | The development of abstract constructions, e.g. Subj VP Ob Obl. |

a good level of ecological validity. As Hulstijn and de Graaff (1994) point out, however, studies in a classroom setting run the risk of allowing interfering variables into the design, resulting in a loss of internal validity. In this study, the non-randomness of students within the classes, the lack of a delayed post-test, as well as variability between teachers and the lack of control over input external to the study all posed a risk to internal validity and replicability. In future research, one way of protecting against interfering variables while maintaining a level of ecological validity is to develop a 'twin' approach (Hulstijn and de Graaff, 1994), in which quasi-experiments in, for example, classroom settings, are twinned with experiments that have more control over

variables, such as those in a laboratory setting using semi-artificial languages. Developing a more controlled research design in many ways would not present many challenges, as the tasks and testing instruments are easily adapted to a computer-controlled design.

One clear potential of (and also limitation to) this study is its portability to other contexts. The current focus was past counterfactuality and the development of paradigmatic slots within a construction (i.e. the past participles that could fit in the post-auxiliary slot) which is abstract in places, but retains substantial morphological material across exemplars (e.g. *would/could*, *have*, *if*, *had*). Further support for analogy in SLA would need to focus firstly on constructions whose trajectory progresses from initial rote-learned formulas to low-scope patterns to more abstract constructions, such as passive constructions, caused-motion constructions and ditransitives. Secondly, research which targeted syntagmatic aspects of constructions such as inversion in question forms and relative clauses would allow for a broader application of analogy and similarity in the acquisition of constructions from formulas. Moving beyond the sentence to areas such as move structures in academic writing (e.g. the move structures in dissertations, outlined by Swales, 1994) could also allow research designs to replicate studies by Gentner and colleagues (specifically, Gentner et al., 2003, 2004; Loewenstein et al., 2003) more closely than in the current study, while simultaneously allowing for a broader discoursal view of the status of analogy in SLA.

More in terms of application to second language classrooms, this study introduced tokens pre-task, and these tokens were available during the performance of the task. This design conflates pre-task strategic planning and in-task integrative planning (see for examples studies in Ellis, 2005b). Pulling these two apart in future research designs would provide a a more fine-grained understanding of the relative benefits of focusing on the micro, in-task formation of analogies or a more macro focus pre-task.

# Chapter 6

# Conclusions

The most important conclusion that can be drawn from this study is the application of research into analogy and similarity to the area of second language acquisition. Substantial research over the past 30 years, largely by Gentner and colleagues, has culminated in a wealth of empirical evidence supporting a core role for analogy in general cognition. This evidence allows for a cohesive model that predicts how analogies are formed (Gentner, 1983), how analogy can lead to the learning of concepts (Gentner et al., 2004), and the variable importance of similarity and analogy at different levels of competence (Gentner and Rattermann, 1991).

In language acquisition from a usage-based perspective, both analogy and similarity are often cited as key mechanisms that allow for the progression from rote-learned formulas to structural regularities. As Tomasello (2003) notes, however, in FLA there is a very limited amount of research on how children form analogies at later stages of development based either on categorial and conceptual knowledge or, more interestingly, how children in the earlier stages of FLA may use similarities between tokens to bootstrap themselves into more abstract structure. In SLA, to my knowledge, there is no empirical research into analogical processing, as formulated by Gentner, although, similar to FLA, it is at times relied upon to explain a way in which formulas form the basis of structure (e.g. Ellis, 2005a; Eskildsen, 2012).

This study, albeit tentatively and with clear limitations (See Section 5.7),

provides both empirical support for the role of analogy and similarity in SLA and also a framework that brings together work on task demands, performance and the breakdown of tokens from the input. In terms of the current data, there is a good case for considering exact use of selections as examples of analogical events, within which both input and output tokens are mapped on to a constructional meaning.

Of particular interest, however, is not the exact repetition of tokens selected from the input but the repurposing of tokens to fit the task demands. In the same manner as the exact repetition of selections, the repurposing of tokens is interpreted as an analogical event, as the output was clearly based on the input and the tokens were interpreted, although not exclusively, as past counterfactuals. The difference, however, between exact repetition and repurposing is the alterations to the input tokens, such as changes to object complements, subjects and verb phrases, for example, made to the tokens. That these repurposing operations correlated with gains on testing instruments suggests that the analogical processing of tokens led to knowledge that could be drawn upon in the processing of sentences on post-tests, tentatively supporting usage-based theories of acquisition that posit that structure emerges piecemeal from the categorisation of like forms (e.g. Langacker, 2000; Robinson and Ellis, 2008). I propose this tentatively here as the behaviour of the participants points towards low-scope patterns developing from the analogical processing of tokens as a reasonable explanation, although the data presented here is only one form of support for such a position.

Certain parallels can be drawn between the predicted outcomes of analogical processing and work in SLA on the selection and elaboration of tokens from the environment. Both Hulstijn and Laufer's Involvement Load Hypothesis (2001) and Robinson's elaboration of the Noticing Hypothesis (Robinson, 2003, based on Schmidt 1990, 1995, 2001) correlate the elaboration of tokens with retention. The current thesis, however, goes beyond the retention of tokens by arguing that the elaboration of tokens can be subdivided into repetition which leads

to entrenchment and analogous processing which leads to the schematisation of tokens. This perspective on elaboration can be related specifically to the function of task demands; when task demands are in line with tokens in the input, the selection and use of these resources allows for the elaboration of the token, but tokens remain intact. When there exists a gap between task demands and resources, the same elaboration is present, but tokens are also broken down to meet the task demands. Seen this way, this sort of framework allows for task demands and input as resource to be varied in interesting ways depending on the qualities of the input and their relationship with the task demands (See Section 5.7).

Analogical and literal processing can be situated within a framework that brings together work on task demands, the pragmatics of task performance, tokens in the input and the breakdown of tokens. If we construe the conceptualisation of a task as creating a need to mean (Samuda, 2001) and the complexity of task demands as leading to the selection of forms from the input (Robinson, 2001), then input can be considered a resource that can be selected to meet these needs. Drawing on Myles et al. (1999), when current resources are insufficient to meet the exigencies of a current communicative situation, tokens (or established formulas in the case of Myles et al.'s data) are broken down to express the necessary intention. In relation to the current study, widening the gap between the input as resource for task performance and the intention that needs to be expressed forces the repurposing of selected tokens. Bridging this gap between input and output is considered here a case of analogy, as the input as resource is mapped on to performance tokens, based on the similarity between tokens and the meanings that they express.

This sort of framework has considerable appeal in the area of pedagogy, most notably the area of task-based learning. In the past 30 years significant progress has been made in relation to the performance on tasks in terms of the *Complexity, Accuracy and Fluency* (CAF) of learner production (e.g. the special issue of Applied Linguistics, Housen and Kuiken, 2009) and the

effects of task-based interaction (e.g. the meta-analysis of interaction studies in Keck et al., 2006). Task-Based Learning (TBL), however, has been criticised as having an overly-pragmatic approach to SLA in which the functional task demands overshadow the construction and subsequent development of new language (Batstone, 2002; Cook, 2008; Swan, 2005; Widdowson, 2003). In line with much work on TBL, this thesis draws on the conceptualisation of task demands and pragmatic communication, but demonstrates at the least that these demands can motivate the selection of task-relevant input as a resource for communication. If we consider the repurposing of task-relevant language as effecting the development of constructions, I would like to suggest that this research might be the starting point of explorations into the interaction of input, task demands and the development of constructions from tokens.

# References


Barlow, M. (2000). Usage, blends and grammar. In Barlow, M. and Kemmer, S., editors, *Usage-based models of language*. CSLI Publications, Stanford, Calif.

Barlow, M. and Kemmer, S. (2000). *Usage-based models of language*. CSLI Publications, Stanford, Calif.

Batstone, R. (2002). Making sense of new language: A discourse perspective. *Language Awareness*, 11(1):14–29.

Bhatt, R. (1997). Counterfactuality in Indo-Aryan, notes assembled for Sabine Iatridou's counterfactuals project. Available at http://people.umass.edu/bhatt/papers/sjv.pdf. Accessed 29th October 2012.

Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Pearson ESL.

Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning*, 33(1):1–17.

Boyd, J. and Goldberg, A. (2009). Input effects within a constructionist framework. *The Modern Language Journal*, 93(3):418–429.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 82(4):711.

Bybee, J. (2008). Usage-based grammar and second language acquisition. In Robinson, R. and Ellis, N. C., editors, *Handbook of Cognitive Linguistics and Second Language Acquisition*, pages 216–236. Routledge, New York.

Bybee, J. L. (2007). *Frequency of use and the organization of language.* Oxford University Press, New York; Oxford.

Bybee, J. L. (2010). *Language, usage and cognition.* Cambridge University Press, Cambridge.

Bygate, M., Skehan, P., and Swain, M. (2001). *Researching pedagogic tasks : second language learning, teaching and testing.* Longman, Harlow.

Bystrov, I. S. and Stankevič, N. V. (2005). Conditional constructions in Vietnamese. In Xrakovskij, V. S., editor, *Typology of conditional constructions*, pages 498–511. Lincom.

Cameron-Faulkner, T., Lieven, E., and Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science*, 27(6):843 – 873.

Casenhiser, D. and Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science*, 8(6):500–508.

Celce-Murcia, M., Larsen-Freeman, D., and Williams, H. A. (1999). *The grammar book: an ESL/EFL teacher's course.* Heinle and Heinle, Boston, 2nd edition.

Chomsky, N. (1965). *Aspects of the theory of syntax.* M.I.T. Press, Cambridge.

Chomsky, N. (1972). *Language and mind.* Harcourt Brace Jovanovich, New York.

Chomsky, N. (1995). *The minimalist program.* M.I.T. Press, Cambridge Mass.

Clark, R. (1974). Performing without competence. *Journal of Child Language*, 1(01):1–10.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates, Hillsdale N.J., 2nd edition.

Comrie, B. (1986). Conditionals: a typology. In Traugott, E. C., Muelen, A., Reilley, J. S., and Ferguson, C., editors, *On conditionals*, pages 77–99. Cambridge University Press.

Cook, V. (2008). *Second language learning and language teaching*. Hodder Education, London, 4th edition.

Council of Europe (2001). *Common European Framework of Reference for Languages : learning, teaching, assessment*. Cambridge University Press, Cambridge U.K.

Cowie, A. P. (1998). Phraseological dictionaries: some east-west comparisons. In Cowie, A. P., editor, *Phraseology : theory, analysis, and applications*. Clarendon press; Oxford university press, Oxford; New York.

Craik, F. I. and Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6):671 – 684.

Craik, F. I. M. and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104(3):268–294.

Croft, W. (2001). *Radical construction grammar : syntactic theory in typological perspective*. Oxford University Press, Oxford.

Croft, W. and Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge University Press, Cambridge.

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition*, 11(4):367–383.

Culicover, P. W. (1999). *Syntactic nuts : hard cases, syntactic theory, and language acquisition*. Oxford Univ. Press, Oxford.

Culicover, P. W. and Nowak, A. (2003). *Dynamical grammar : minimalism, acquisition, and change*. Oxford University Press, Oxford; New York.

Cummins, R. (1983). *The nature of psychological explanation*. MIT Press, Cambridge, Mass.; London.

Dabrowska, E. (2001). From formula to schema: The acquisition of English questions. *Cognitive Linguistics*, 11(1-2):83–102.

Dabrowska, E. (2006). Low-level schemas or general rules? the role of diminutives in the acquisition of Polish case inflections. *Language Sciences*, 28(1):120–135.

Dabrowska, E. and Lieven, E. (2005). Towards a lexically specific grammar of children's question constructions. *Cognitive linguistics*, 16(3):437–474.

Dancygier, B. and Sweetser, E. (2005). *Mental spaces in grammar : conditional constructions*. Cambridge University Press, Cambridge.

Dulay, H., Burt, M., and Krashen, S. (1982). *Language two*. Oxford University Press, New York.

Elio, R. and Anderson, J. R. (1984). The effects of information order and learning mode on schema abstraction. *Memory  Cognition*, 12(1):20–30.

Ellis, N. C. (1996). Sequencing in SLA. *Studies in Second Language Acquisition*, 18(01):91–126.

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in second language acquisition*, 24(02):143–188.

Ellis, N. C. (2005a). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27(02):305–352.

Ellis, N. C. (2008). Phraseology: The periphery and the heart of language. In Meunier, F. and Grainger, S., editors, *Phraseology in language learning and teaching*, pages 1–13. John Benjamins.

Ellis, N. C. (2009). Optimizing the input: Frequency and sampling in usage-based and form-focused learning. In Long Michael, H. and Doughty, C., editors, *The Handbook of Language Teaching*, pages 139–158. Wiley-Blackwell.

Ellis, N. C. (2010). Construction learning as category learning. In Pütz, M. and Sicola, L., editors, *Cognitive processing in second language acquisition : inside the learner's mind*. John Benjamins Pub., Amsterdam; Philadelphia.

Ellis, N. C. and Ferreira-Junior, F. (2009a). Construction learning as a function of frequency, frequency distribution, and function. *The Modern Language Journal*, 93(3):370–385.

Ellis, N. C. and Ferreira-Junior, F. (2009b). Constructions and their acquisition islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics*, 7(1):187–220.

Ellis, N. C. and Larsen-Freeman, D. (2009). Constructing a second language: Analyses and computational simulations of the emergence of linguistic constructions from usage. *Language Learning*, 59(1):90–125.

Ellis, R. (1984). Formulaic speech in early classroom second language development. In Handscombe, J., Orem, R. A., and Taylor, B. P., editors, *On TESOL '83: The Question of Control.*, pages 60–72, Washington, DC. TESOL.

Ellis, R. (2005b). Planning and task-based performance: Theory and research. *Language Learning and Language Teaching*, 11:3–36.

Ellis, R. and Barkhuizen, G. P. (2005). *Analysing learner language*. Oxford University Press, Oxford.

Ellis, R., Loewen, S., Elder, C., Erlam, R., and Philp, J. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching*. Second language acquisition. Multilingual Matters, Bristol, UK.

Eskildsen, S. and Cadierno, T. (2007). Are recurring multi-word expressions really syntactic freezes. Second language acquisition from the perspective of usage-based linguistics. In Nenonen, M. and Niemi, S., editors, *Collocations and idioms 1, papers from the First Nordic Conference on Syntactic Freezes*, pages 86–99, Joensuu: Finland. Joensuu University Press.

Eskildsen, S. W. (2009). Constructing another language—usage-based linguistics in second language acquisition. *Applied Linguistics*, 30(3):335–357.

Eskildsen, S. W. (2012). L2 negation constructions at work. *Language Learning*, 62(2):335–372.

Eubank, L. and Gregg, K. (2002). News flash - Hume still dead. *Studies in Second Language Acquisition*, 24(02):237–247.

Færch, C. and Kasper, G., editors (1987). *Introspection in second language research*. Multilingual Matters, Clevedon, U.K.

Fauconnier, G. (1997). *Mappings in thought and language*. Cambridge University Press, Cambridge.

Fillmore, C., Kay, P., and O'Connor, M. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language*, 64(3):501–538.

Fodor, J. D. and Crowther, C. (2008). Understanding stimulus poverty arguments. *The Linguistic Review*, 18(1-2):105–145.

Foster, P. and Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18(3):299–324.

Foster, P. and Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, 3(3):215–247.

Gass, S. and Mackey, A. (2000). *Stimulated recall methodology in second language research*. Lawrence Erlbaum Associates, Mahwah; London.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2):155 – 170.

Gentner, D. (2003). Why we're so smart. In Gentner, D. and Goldin-Meadow, S., editors, *Language in mind: Advances in the study of language and thought*, pages 195–235. MIT Press, Cambridge, MA.

Gentner, D., Loewenstein, J., and Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development*, 8(3):285–307.

Gentner, D., Loewenstein, J., and Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, 95:393–408.

Gentner, D., Loewenstein, J., and Thompson, L. (2004). Analogical encoding: Facilitating knowledge transfer and integration. In Forbus, K., Gentner, D., and Regier, T., editors, *Proceedings of the 26th Meeting of the Cognitive Science Society*, pages 452–457.

Gentner, D. and Markman, A. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52(1):45–56.

Gentner, D. and Markman, A. B. (2005). Defining structural similarity. *Journal of Cognitive Science*, 6(1):1–20.

Gentner, D. and Medina, J. (1997). Comparison and the development of cognition and language. *Cognitive Studies: Bulletin of the Japanese Cognitive Society*, 4(1):112–149.

Gentner, D. and Medina, J. (1998). Similarity and the development of rules. *Cognition*, 65(2-3):263 – 297.

Gentner, D. and Namy, L. (2006). Analogical processes in language learning. *Current Directions in Psychological Science*, 15(6):297.

Gentner, D. and Rattermann, M. J. (1991). Language and the career of similarity. In Gelman, S. A. and Brynes, J. P., editors, *Perspectives on Language and Thought: Interrelations in Development*, pages 225–277. Cambridge University Press, London.

Gentner, D. and Smith, L. (2012). Analogical reasoning. In Ramachandran, V. S., editor, *Encyclopedia of Human Behavior*, pages 130–136. Elsevier, Oxford, UK, 2nd edition.

Gick, M. L. and Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15(1):1 – 38.

Goldberg, A. E. (1995). *Constructions: a construction grammar approach to argument structure*. University of Chicago Press, Chicago.

Goldberg, A. E. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences*, 7(5):219–224.

Goldberg, A. E. (2006). *Constructions at work : the nature of generalization in language*. Oxford University Press, Oxford.

Goldberg, A. E. and Jackendoff, R. (2004). The English resultative as a family of constructions. *Language*, 80(3):pp. 532–568.

Gregg, K. (2003a). SLA theory: Construction and assessment. In Doughty, C. and Long, M. H., editors, *Handbook of Second Language Research*, pages 831–865. Blackwell, London.

Gregg, K. R. (2003b). The state of emergentism in second language acquisition. *Second Language Research*, 19(2):95–128.

Grønn, A. (2008). An amazing come-back: a counterfactual imperfective in Russian. *Scando-Slavica*, 54(1):5–31.

Hakuta, K. (1974). Prefabricated patterns and the emergence of structure in second language acquisition. *Language Learning*, 24(2):287–297.

Hanania, E. A. S. and Gradman, H. L. (1977). Acquisition of English structures: A case study of an adult native speaker of Arabic in an English-speaking environment. *Language Learning*, 27(1):75–91.

Hasada, R. (1997). Conditionals and counterfactuals in japanese. *Language Sciences*, 19(3):277 – 288.

Hawkins, R. (2008). Can innate linguistic knowledge be eliminated from theories of SLA? *Lingua*, 118(4):613 – 619.

Housen, A. and Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition [introduction to special edition]. *Applied Linguistics*, 30(4):461–473.

Howarth, P. (1998). The phraseology of learners' academic writing. In Cowie, A. P., editor, *Phraseology : theory, analysis, and applications*. Clarendon press; Oxford university press, Oxford; New York.

Huang, J. and Hatch, E. (1978). A Chinese child's acquisition of English. In Hatch, E. M., editor, *Second language acquisition: a book of readings*. Newbury House Publishers, Rowley Mass.

Hudson, R. (2008). Word grammar, cognitive linguistics and second language learning and teaching. In Ellis Nick, C. and Robinson, P., editors, *Handbook of Cognitive Linguistics and Second Language Acquisition*, pages 89–113. Lawrence Erlbaum.

Hulstijn, J. H. and de Graaff, R. (1994). Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *AILA Review*, 11:97–112.

Iatridou, S. (2000). The grammatical ingredients of counterfactuality. *Linguistic Inquiry*, 31(2):231–270.

Izumi, S. and Bigelow, M. (2000). Does output promote noticing and second language acquisition? *Tesol Quarterly*, 34(2):239–278.

Izumi, S., Bigelow, M., Fujiwara, M., and Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition*, 21(3):421–452.

Jackendoff, R. (2002). *Foundations of language : brain, meaning, grammar, and evolution*. Oxford University Press, Oxford.

Jacoby, L. L. (1983). Remembering the data: analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior*, 22(5):485 – 508.

Kay, P. (1997). *Words and the grammar of context*. CSLI Publications, Stanford Calif.

Kay, P. and Fillmore, C. (1999). Grammatical constructions and linguistic generalizations: The what's x doing y construction? *Language*, 75(1):1–33.

Keck, C. M., Iberri-Shea, G., Tracy-Ventura, N., and Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition. In Norris, J. and Ortega, L., editors, *Synthesizing research on language learning and teaching*, pages 91–131. John Benjamins, Amsterdam.

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, 56(5):746–759.

Krashen, S. and Scarcella, R. (1978). On routines and patterns in language acquisition and performance. *Language Learning*, 28(2):283–300.

Lakoff, G. (1987). *Women, fire, and dangerous things : what categories reveal about the mind*. University of Chicago Press, Chicago, Ill.; London.

Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–74.

Langacker, R. W. (1987). *Foundations of cognitive grammar*. Stanford University Press, Stanford, Calif.

Langacker, R. W. (2000). A dynamic usage-based model. In Barlow, M. and Kemmer, S., editors, *Usage-based Models of Language*, pages 1–64. CSLI Publications, Stanford.

Lardiere, D. (1992). On the linguistic shaping of thought: Another response to Alfred Bloom. *Language in Society*, 21(02):231–251.

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. Routledge, New York; London.

Laufer, B. and Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. *Applied Linguistics*, 22(1):1–26.

Li, C. N. and Thompson, S. A. (1989). *Mandarin Chinese: a functional reference grammar*. University of California Press, Berkeley.

Lieven, E., Behrens, H., Speares, J., and Tomasello, M. (2003). Early syntactic creativity: a usage-based approach. *Journal of Child Language*, 30(02):333–370.

Lieven, E., Pine, J., and Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of child language*, 24(01):187–219.

Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In Ellis, R., Loewen, S., Elder, C., Erlam, R., and Philp, J., editors, *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*. Multilingual Matters, Bristol UK.

Loewenstein, J., Thompson, L., and Gentner, D. (2003). Analogical learning in negotiation teams: Comparing cases promotes learning and transfer. *Academy of Management Learning  Education*, 2(2):119–127.

Long, M. H. (1985). A role for instruction in second language acquisition. In Hyltenstam, K. and Pienemann, P., editors, *Modelling and Assessing Second Language Acquisition*. Multilingual Matters, Clevedon Avon.

Lynch, T. (1989). Researching teachers: Behaviour and belief. In Brumfit, C. and Mitchell, R., editors, *Research in the Language Classroom*. Modern English Publications and the British Council, London.

Lyons, J. (1968). *Introduction to theoretical linguistics*. CUP, London.

Mace, J. (2002). *Persian Grammar: For Reference and Revision*. Routledge, London.

MacWhinney, B. (2004). A multiple process solution to the logical problem of language acquisition. *Journal of Child Language*, 31(04):883–914.

Markman, A. B. and Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology*, 25(4):431–467.

McDonough, K. and Kim, Y. (2009). Syntactic priming, type frequency, and EFL learners' production of wh-questions. *The Modern Language Journal*, 93(3):386–398.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20(1):83–107.

Mochizuki, N. and Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research*, 12(1):11.

Moonen, M., Graaff, R. d., and Westhoff, G. (2006). Focused tasks, mental actions and second language learning. cognitive and connectionist accounts of task effectiveness. *ITL - International Journal of Applied Linguistics*, 152:35–53.

Myles, F. (2004). From data to theory: the over-representation of linguistic knowledge in SLA. *Transactions of the Philological Society*, 102(2):139–168.

Myles, F., Hooper, J., and Mitchell, R. (1998). Rote or rule? exploring the role of formulaic language in classroom foreignlanguage learning. *Language Learning*, 48(3):323–364.

Myles, F., Mitchell, R., and Hooper, J. (1999). Interrogative chunks in french L2. *Studies in Second Language Acquisition*, 21(01):49–80.

Nakagawa, S. (2005). The logical constructions of conditional sentences and the functions of conditional expressions: A contrastive study of Japanese and Thai. In Sidwell, P., editor, *SEALS XV Papers from the 15th Annual Meeting of the Southeast Asian Linguistics Society 2003*, pages 41–49. Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University.

Nattinger, J. R. and DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford University Press, Oxford; New York.

Norris, J. and Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3):417–528.

Norris, J. M. and Ortega, L. (2003). Defining and measuring SLA. In Doughty, C. J. and Long, M. H., editors, *Handbook of second language acquisition*, pages 717–761. Blackwell, Malden, MA.

Ortega, L. (2005). 'What do learners plan? Learner-driven attention to form during pre- task planning. In Ellis, R., editor, *Planning and Task Performance in a Second Language*. John Benjamins, Amsterdam, NL.

Pawley, A. and Syder, H. (1983). Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In Richards, J. and Schmidt, R., editors, *Language and communication*. Longman, London; New York.

Perdue, C. (1993). *Adult language acquisition : cross-linguistic perspectives: Field Methods*, volume 1. Cambridge University Press, Cambridge ; New York, NY.

Peters, A. M. (1985). *The units of language acquisition*. Cambridge Univ. Press, Cambridge.

Pinker, S. (1989). *Learnability and cognition the acquisition of argument structure*. MIT Press, Cambridge, Mass.

Pinker, S. (1998). Words and rules. *Lingua*, 106(1-4):219–242.

Pinker, S. (2004). Clarifying the logical problem of language acquisition. *Journal of Child Language*, 31(04):949–953.

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford University Press, Oxford.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. Pearson Longman, London.

Riemsdijk, H. C. and Williams, E. (1986). *Introduction to the theory of grammar*. MIT Press, Cambridge Mass.

Robinson, P. (1995). Attention, memory, and the 'Noticing' Hypothesis. *Language Learning*, 45(2):283–331.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22(1):27–57.

Robinson, P. (2003). Attention and memory during SLA. In Doughty, C. and Long, M. H., editors, *Handbook of Second Language Acquisition*, chapter 19, pages 630–678. Blackwell, Oxford.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *IRAL - International Review of Applied Linguistics in Language Teaching*, 43(1):1–80.

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *Iral*, 45:193–214.

Robinson, P. and Ellis, N. C. (2008). *Handbook of cognitive linguistics and second language acquisition*. Routledge, New York.

Samuda, V. (2001). Guiding relationships between form and meaning during task performance: the role of the teacher. In Bygate, M., Skehan, P., and Swain, M., editors, *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*, pages 119–140. Longman, Harlow.

Schmidt, R. (1983). Interaction, acculturation and the acquisition of communicative competence. In Wolfson, N. and Judd, E., editors, *Sociolinguistics and language acquisition*, pages 137–174. Newbury House, Rowley, MA.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2):129–158.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. In Schmidt, R. and Hulstijn, J., editors, *Consciousness and second language learning: Conceptual, methodological and practical issues in language learning and teaching*, volume 11, pages 11–26. AILA Review.

Schmidt, R. (1995). *Consciousness and foreign language learning: a tutorial on the role of attention and awareness in learning*, volume Technical Report No. 9. University of Hawai'i, Second Language Teaching Curriculum Center, Honolulu, HI.

Schmidt, R. (2001). Attention. In Robinson, P., editor, *Cognition and second language instruction*. CUP.

Schumann, J. H. (1976). Social distance as a factor in second language acquisition. *Language Learning*, 26(1):135–143.

Seedhouse, P. (1999). Task-based interaction. *ELT J*, 53(3):149–156.

Shapira, R. (1978). The non-learning of English: Case study of an adult. In Hatch, E., editor, *Second language acquisition : a book of readings*. Newbury House Publishers, Rowley Mass.

Shirai, Y. (2007). The aspect hypothesis, the comparative fallacy and the validity of obligatory context analysis: a reply to Lardiere, 2003. *Second Language Research*, 23(1):51–64.

Sinclair, J. M. (1991). *Corpus, concordance, collocation.* Oxford Univ. Press, Oxford.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17(1):38–62.

Skehan, P. (1998). *A cognitive approach to language learning.* Oxford Univ. Press, Oxford.

Skehan, P. and Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1(3):185–211.

Skehan, P. and Foster, P. (2001). Cognition and tasks. In Robinson, P., editor, *Cognition and second language instruction.* Cambridge Univ. Press.

Song, M. and Suh, B. (2008). The effects of output task types on noticing and learning of the English past counterfactual conditional. *System*, 36(2):295–312.

Sorace, A. (1996). The use of acceptability judgments in SLA research. In Ritchie, W. C. and Bhatia, T. K., editors, *Handbook of second language acquisition*, pages 375–409. Academic Press, San Diego.

Swales, J. M. (1994). *Academic writing for graduate students : essential tasks and skills : a course for nonnative speakers of English.* University of Michigan Press, Ann Arbor.

Swan, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics*, 26(3):376–401.

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, 11(1/2):61–82.

Tomasello, M. (2003). *Constructing a language : a usage-based theory of language acquisition.* Harvard University Press, Cambridge, Mass., London.

Tomasello, M. (2004). What kind of evidence could refute the UG hypothesis?: Commentary on Wunderlich. *Studies in Language*, 28(3):642–645.

Trahey, M. and White, L. (1993). Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition*, 15(2):181.

Truscott, J. (1998). Noticing in second language acquisition: a critical review. *Second Language Research*, 14(2):103–135.

Truscott, J. and Sharwood Smith, M. (2011). Input, intake, and consciousness. *Studies in Second Language Acquisition*, 33(04):497–528.

Tummers, J., Heylen, K., and Geeraerts, D. (2005). Usage-based approaches in cognitive linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory*, 1(2):151–310.

VanPatten, B. (1990). Attending to form and content in the input. *Studies in Second Language Acquisition*, 12(03):287–301.

VanPatten, B. (2004). *Processing instruction theory, research, and commentary*. L. Erlbaum Associates, Mahwah, N.J.

Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied Linguistics*, 16(2):180–205.

White, L. (2003). *Second language acquisition and universal grammar*. Cambridge University Press, Cambridge.

Widdowson, H. G. (1990). *Aspects of language teaching*. Oxford University Press, Oxford.

Widdowson, H. G. (2003). *Defining issues in English language teaching*. Oxford University Press, Oxford.

Wierzbicka, A. (1997). Conditionals and counterfactuals: Conceptual primitives and linguistic universals. In Athanasiadou, A. and Dirven, R., editors, *On Conditionals Again*, pages 15–59. John Benjamins, Amsterdam.

Wilkinson, L. and the Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54(8):594–604.

Wong Fillmore, L. (1976). *The second time around: cognitive and social strategies in second language acquisition.* PhD thesis, Stamford.

Wong Fillmore, L. (1979). Individual differences in second language acquisition. In Fillmore, C. J., Kempler, D., and Wang, W. S. Y., editors, *Individual Differences in Language Ability and Language Behavior.* Academic Press.

Wray, A. (2002). *Formulaic language and the lexicon.* Cambridge University Press, Cambridge; New York.

Wray, A. (2007). 'Needs only' analysis in linguistic ontogeny and phylogeny. In Lyon, C., Nehaniv, C., and Cangelosi, A., editors, *Emergence of Communication and Language*, pages 53–70. Springer.

Wray, A. and Perkins, M. R. (2000). The functions of formulaic language: an integrated model. *Communication*, 20(1):1–28.

Yorio, C. (1989). Idiomaticity as an indicator of second language proficiency. In Hyltenstam, K. and Obler, L. K., editors, *Bilingualism across the lifespan: aspects of acquisition, maturity, and loss.* Cambridge University Press, Cambridge.

Yorio, C. A. (1980). Conventionalized language forms and the development of communicative competence. *TESOL Quarterly*, 14(4):pp. 433–442.

Zipf, G. K. (1949). *Human behavior and the principle of least effort an introduction to human ecology.* Addison-Wesley Press, Cambridge Mass.

# Appendices

# Appendix A

# Consent form and biographical questionnaire for the study

**Andy Harris. Dphil Research.** A.j.harris@uel.ac.uk

## Consent Form

I understand that the results of the research study I am participating in might be published and presented at professional conferences, but that my name and personal details will be kept confidential and will not appear anywhere.

I understand that any information collected about me will be electronically encrypted, stored in a locked office, and will be destroyed at the end of the project.

I understand that my voice may be recorded both during the class and during short interviews after the class.

I understand that I can leave the study at any time without giving a reason. I also understand that my participation in this study will have NO effect on my assessment or studies at the university now or in the future.

I have been given the opportunity to ask questions, and any questions have been answered. I also know how to complain if I have any problems with the study.

Please tick (☑) the following box, if you agree.

**I AGREE to take part in the research** ☐

Full name: _____

Signature: _____

| **Name:** | |
|---|---|
| **Email:** | |

| | | | |
|---|---|---|---|
| Age: | | Nationality: | |
| Male ☐ Female ☐ | | Next course: | PG ☐ UG ☐ None ☐ |
| First Language: | | Second/third languages (please include English here): | 1st 2nd 3rd 4th |
| How good are you in your second language? | Excellent<br>Good<br>Ok<br>A Little Weak<br>Very Weak | How good are you in your third language? | Excellent<br>Good<br>Ok<br>A Little Weak<br>Very Weak |
| Years and months studying English in a classroom: | ___years<br>___months | Years and months living in an English speaking country: | ___years<br>___months |
| IELTS Score | | Other test: | |

| How good do you think your English is (Circle One) | | | | | |
|---|---|---|---|---|---|
| Writing | Very Good | Good | OK | A Little Weak | Very Weak |
| Speaking | Very Good | Good | OK | A Little Weak | Very Weak |
| Listening | Very Good | Good | OK | A Little Weak | Very Weak |
| Reading | Very Good | Good | OK | A Little Weak | Very Weak |
| Vocabulary | Very Good | Good | OK | A Little Weak | Very Weak |
| Grammar | Very Good | Good | OK | A Little Weak | Very Weak |

This research has been cleared by the Oxford CUREC system

# Appendix B

# Ethics approval confirmation for the study (CUREC)

SOCIAL SCIENCES & HUMANITIES
INTER-DIVISIONAL RESEARCH ETHICS COMMITTEE

Hayes House, 75 George Street, Oxford. OX1 2BQ
Tel: +44(0)1865 614871 Fax: +44(0)1865 614855
ethics@socsci.ox.ac.uk www.socsci.ox.ac.uk

Co-ordinator of the IDREC
Social Sciences Divisional Office

UNIVERSITY OF OXFORD

Ref. SSD/2/3/IDREC

07 December 2009

Andrew Harris,
111 Station Road
London
N3 2SP

Dear Andrew Harris,

**Application for research ethics approval**

**Ref No.: SSD/CUREC1/09 – 247**

**Title: The effects of task demands on form-function detection in second-language classrooms**

The above application has been considered on behalf of the Social Sciences and Humanities Inter-divisional Research Ethics Committee (IDREC) in accordance with the procedures laid down by the University for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to the IDREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly approval has been granted.

Should there be any subsequent changes to the project, which raise ethical issues not covered in the original application, you should submit details to the IDREC for consideration.

Yours sincerely,

Chris. Ballinger

Dr Chris. Ballinger

cc: Dr Catherine Walter, Department of Education
Lisa Currie, Department of Education

# Appendix C

# Pre-study information for participants

**Andy Harris. Dphil Research. A.j.harris@uel.ac.uk**

Dear Student,

My name is Andy Harris. I am an English language lecturer at The University of East London. I am also a doctoral (DPhil) student at The University of Oxford. For my DPhil, I am investigating how students learn a second language, and I would like to do some classroom research in this area. To do my research, I need students to take part in some classroom activities.

You will take part in five sessions, which will last approximately 40 minutes each. During these classes, I will be giving you some very intensive exercises to complete that will help you with your English language.

The data I collect will be used for my DPhil, and might be published and presented at professional conferences. However, only I will know your name and personal details; they will not appear in my DPhil or in any other piece of work. The data I collect will be electronically encrypted, stored in a locked office, and will be destroyed at the end of the project.

Your participation in this study will have NO effect on your assessment or studies at the university now or in the future.

If you would like any more information, please feel free to contact me at a.j.harris@uel.ac.uk. If you feel I haven't dealt with the study professionally, you can contact the secretary of the Department's Research Ethics Committee, Dr C. Ballinger:

Social Sciences Division
75 George Street,
Oxford
OX1 2BQ
Email: ethics@socsci.ox.ac.uk

Thank you very much for your help.

Andy Harris

This research has been cleared by the Oxford CUREC system.

# Appendix D

# Production test

**Andy Harris. Dphil Research. A.j.harris@uel.ac.uk**

# Production Exercise – imagining a different past time

**The situation**

Joey was born and grew up in a small town in Brazil. During his school years, his one aim was to move to the UK to study and experience another culture. In 1966, at the age of 18, he was given a scholarship and he moved to London, and between 1966 and 1969, Joey studied for a BA in business management at a London university. While he was studying, he met (and lived with) a lot of students from a lot of different nationalities. Because of this, his English improved and he got a really high mark in the English test he took in his final year. During his time at university, he studied a lot, and so in 1969 he got a first in his course. Because of this result, he got himself a good job in a small import/export company in London. He also met the love of his life on the BA course and in 1970 they got married and 2 years later, they had a daughter who they called Ramona. In 1981, at the age of 43, he was promoted to assistant director, and was earning a lot of money. He bought a lot of things, but by far the best thing was a flat he bought in South London with a view over the river. Joey died in 2009 - he was a little young, but he'd had a great life.

**Your task: Joey's story was true. Your task is to imagine that Joey had stayed in his own country and had not come to the UK and had not studied Business management in London. Describe below the imaginary differences in his life. Use the information above and your own imagination to do this. Start with the sentence at the top of the box. You have 10 minutes to do this.**

**Joey did very well at university in the UK and was very successful after he finished. However, imagine if…**

# Appendix E

# Acceptability judgment test

**Andy Harris. Dphil Research. A.j.harris@uel.ac.uk**

# Acceptability Judgement Exercise

## Instructions

Look at the following sentences.

1. Decide if the English feels right or wrong or somewhere in the middle. Tick ONE of the boxes.

## Example

**Sentence 0. I am go to go to my friends house tonight**

| | How good does the sentence feel? → | | | | | |
|---|---|---|---|---|---|---|
| Sentence Number | Feels very wrong | Feels wrong | Feel a little wrong | Feels OK | Feels right | Feels very right |
| Sentence 0. | | ✓ | | | | |

## Practice

| | How good does the sentence feel? → | | | | | |
|---|---|---|---|---|---|---|
| Sentence Number | Feels very wrong | Feels wrong | Feel a little wrong | Feels OK | Feels right | Feels very right |
| Sentence 1. | | | | | | |
| Sentence 2. | | | | | | |
| Sentence 3. | | | | | | |
| Sentence 4. | | | | | | |

**Andy Harris. Dphil Research. A.j.harris@uel.ac.uk**

## Acceptability exercise

| | How good does the English feel? | | | | | |
|---|---|---|---|---|---|---|
| Sentence Number | Feels very wrong | Feels wrong | Feel a little wrong | Feels OK | Feels right | Feels very right |
| Sentence 1. | | | | | | |
| Sentence 2. | | | | | | |
| Sentence 3. | | | | | | |
| Sentence 4. | | | | | | |
| Sentence 5. | | | | | | |
| Sentence 6. | | | | | | |
| Sentence 7. | | | | | | |
| Sentence 8. | | | | | | |
| Sentence 9. | | | | | | |
| Sentence 10. | | | | | | |
| Sentence 11. | | | | | | |
| Sentence 12. | | | | | | |
| Sentence 13. | | | | | | |
| Sentence 14. | | | | | | |
| Sentence 15. | | | | | | |
| Sentence 16. | | | | | | |
| Sentence 17. | | | | | | |
| Sentence 18. | | | | | | |
| Sentence 19. | | | | | | |
| Sentence 20. | | | | | | |
| Sentence 21. | | | | | | |
| Sentence 22. | | | | | | |
| Sentence 23. | | | | | | |
| Sentence 24. | | | | | | |
| Sentence 25. | | | | | | |
| Sentence 26. | | | | | | |
| Sentence 27. | | | | | | |
| Sentence 28. | | | | | | |
| Sentence 29. | | | | | | |
| Sentence 30. | | | | | | |
| Sentence 31. | | | | | | |
| Sentence 32. | | | | | | |
| Sentence 33. | | | | | | |
| Sentence 34. | | | | | | |
| Sentence 35. | | | | | | |
| Sentence 36. | | | | | | |
| Sentence 37. | | | | | | |
| Sentence 38. | | | | | | |
| Sentence 39. | | | | | | |
| Sentence 40. | | | | | | |

## The End. Thank you

EXAMPLE ACCEPTABILITY JUDGMENT TEST SENTENCES

1. I love people who are friendly.
2. The man sit next to me smells.
3. I cannot stand it when people talk loudly on their
4. I am loving it.

MAIN ACCEPTABILITY JUDGMENT TEST SENTENCES

1. He is the guy who I helped with the test.
2. If I was born in France, I would obviously have spoken French when I was growing up.
3. She stopped me from going on holiday because of the tests.
4. By the time you get home tonight, will you phone her?
5. If we had had a lot more help from other people, we earn a lot more money.
6. You had worked really hard on your writing this month.
7. I would have had more job opportunities if I had taken French at school.
8. We have been waiting here for ages. Where have you been?
9. The decision which it was taken was wrong.
10. I would have accepted that new assignment at work if I had had more time.
11. You could call me last Saturday, if you had wanted to.
12. We went out for a meal, and then after that, we had gone for a drink.
13. I must have left it in the office on the top floor.
14. If I had not gone on holiday last week, I would go mad last weekend.
15. There is a place where I do not want to go to.
16. If Sam had not gone out last weekend, he could have translated the letter for you.
17. If Mark had got the job instead of Johnny, he would have moved to the USA.
18. If I had worked harder when I was at school, I could pass my exams when I left.
19. She would have finished her assignments last September, if she had had a little more time.
20. Would you like to coming out tonight?
21. The sign says we mustn't park here, even for ten minutes.
22. We could have gone on a trip to Scotland, if my parents had not arrived.
23. If I had jump on the train this morning, I would not have had to wait.
24. If we had had more money last year, we would have go on holiday to The Maldives.
25. Do you think you will have finished it by next Monday?
26. I've been waiting for this moment for the whole of my life.
27. Last week, they have been shopping in the town centre for hours.
28. There is three people who can do it.
29. I will have gone with you last Christmas, if you had asked me.
30. If John had not wasted his money on designer clothes, he would have been able to take a holiday.
31. It was 10 o'clock before I realized I had forgotten my keys.
32. He could have told me that he was going out.
33. She would not have needed to get a visa, if she had been born in the United Kingdom.
34. If you did not spend all your money last week, you would have been able to go out last night.
35. When I was younger, I used to smoke a lot.
36. If I had won the lottery, I would have been able to help you.
37. I would have been able to do the essay last week, if you had not interrupted me all the time.
38. If I had had more money, I could have gone out last night.
39. When I was growing up, we use to play in the fields all summer.
40. What would you do in my situation?

# Appendix F

# Materials for the Analogous Group (Group One, Task One)

## Group 1: task 1

Andy Harris. DPhil Research. A.j.harris@uel.ac.uk

**Name:**

## Your Task: 'Money and the Bank Manager'

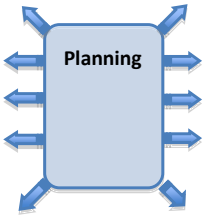
**Scenario:** *Ten years ago you won the lottery - £14.4 million pounds to be exact. You were, obviously, very happy and made a lot of plans to buy a football team, buy new houses, stop your job and travel the world. So that you didn't have to give any money to the taxman, you gave all your money – all £14.4 million - to your best friend. Two weeks later, she/he ran away with the bank manager and all your money. Obviously you feel very stupid and regret what you did.*

I would like you to look back ten years and imagine a hypothetical situation. In this situation I would like you to imagine that you hadn't given your friend the money and imagine the changes in your life ten years ago. In 15 minutes time, I would like you to write as much as you can on the piece of paper provided. Begin your writing with:

**"Ten years ago I won 14.4 million pounds, and I gave it all to my best friend. Unfortunately, after I gave him/her the money, she/he ran away with the bank manager, and my money. Obviously I really regret doing this, but imagine if…"**

Before you start writing, I would like you to first plan your writing and then find some useful phrases for your writing.

**Planning:** You have 8 minutes to plan what you are going to write. Read the task again very carefully and write down ideas, words and phrases that will help you when you write it properly. You can use a dictionary if you want to, and you can keep this piece of paper when you write.



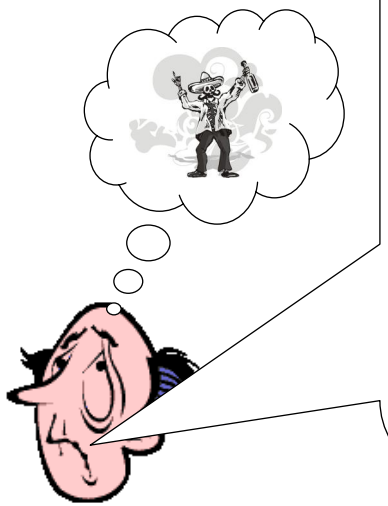
## Group 1: task 1

Andy Harris. DPhil Research. A.j.harris@uel.ac.uk

### The Scenario

Johnny went out last night and regrets it now. On the right hand side is what he's saying now about the situation.

**Language:** Read what Johnny is saying about the situation and decide on between 10 and 15 words or phrases that will help you with YOUR 'Money and the Bank Manager' task. When you write, you will not be able to see this piece of paper so carefully copy the words or phrases on to the next piece of paper. You will be able to use your 10 to 15 words and phrases and your planning when you write the task. Please use a dictionary if you want to. You have 12 minutes to do this.



I really regret going out last night because I didn't do the things that I wanted to do. What's more, because I was very tired, I did, and said, some stupid things that are making me feel pretty bad. If things had been different, it would have been a lot better. For example, If I'd stayed at home and hadn't been so tired, I wouldn't have said those terrible things to my friend – like he should get a better job, and I wouldn't have told my girlfriend that I didn't love her. I wouldn't have told everyone what my boss said to me about Mark - she said Mark was useless. Also, I agreed to go to India this summer with Joey – If I hadn't been partying, I wouldn't have done that, and I wouldn't have agreed to pay for his ticket.

If I'd stayed at home, and if I had gone to bed early, I wouldn't have fallen asleep on the bus on the way home, and I wouldn't have woken up tired this morning. What's more, I spent all my money on getting in the club and on drinks. If I had stayed at home, I wouldn't have spent it all, and so I could've spent it on other things, like some CDs, a new tent, or I could have gone out for a meal instead. Or, better, I could have gone to the cinema and seen the new Star Trek film.

I did some stupid things as well like dancing on the stage, and falling over in the street. If I hadn't gone out, I wouldn't have fallen over and hurt my hand, or ripped my jeans. And I definitely wouldn't have embarrassed myself on the dance floor.

## Group 1: task 1

Andy Harris. DPhil Research. A.j.harris@uel.ac.uk

**10-15 Words or Phrases**

1._____

2._____

3._____

4._____

5._____

6._____

7._____

8._____

9._____

10._____

11._____

12._____

13._____

14._____

15._____

**Writing:** **"Ten years ago I won 14.4 million pounds, and I gave it all to my best friend. Unfortunately, after I gave him/her the money, she/he ran away with the bank manager, and my money. Obviously I really regret doing this, but imagine if…"**

# Appendix G

# Materials for the Literal Group (Group Two, Task One)

## Group 2: task 1

Andy Harris. DPhil Research. A.j.harris@uel.ac.uk

**Name:**

## Your Task: 'Money and the Bank Manager'

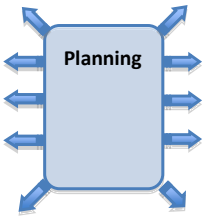
**Scenario:** *Ten years ago you won the lottery - £14.4 million pounds to be exact. You were, obviously, very happy and made a lot of plans to buy a football team, buy new houses, stop your job and travel the world. So that you didn't have to give any money to the taxman, you gave all your money – all £14.4 million - to your best friend. Two weeks later, she/he ran away with the bank manager and all your money. Obviously you feel very stupid and regret what you did.*

I would like you to look back ten years and imagine a hypothetical situation. In this situation I would like you to imagine that you hadn't given your friend the money and imagine the changes in your life ten years ago. In 15 minutes time, I would like you to write as much as you can on the piece of paper provided. Begin your writing with:

**"Ten years ago I won 14.4 million pounds, and I gave it all to my best friend. Unfortunately, after I gave him/her the money, she/he ran away with the bank manager, and my money. Obviously I really regret doing this, but imagine if…"**

Before you start writing, I would like you to first plan your writing and then find some useful phrases for your writing.

**Planning:** You have 8 minutes to plan what you are going to write. Read the task again very carefully and write down ideas, words and phrases that will help you when you write it properly. You can use a dictionary if you want to, and you can keep this piece of paper when you write.



## Group 2: task 1

Andy Harris. DPhil Research. A.j.harris@uel.ac.uk

### The Scenario

Johnny gave all the money to his friend who then ran off with the bank manager. Obviously he feels very stupid and regrets what he did.

**Language:** Read what Johnny is saying about the situation and decide on between 10 and 15 words or phrases that will help you with YOUR 'Money and the Bank Manager' task. When you write, you will not be able to see this piece of paper so carefully copy the words or phrases on to the next piece of paper. You will be able to use your 10 to 15 words and phrases and your planning when you write the task. Please use a dictionary if you want



Ten years ago I won 14.4 million pounds, and I gave it all to best friend. Unfortunately, after I gave her/him the money, she/he ran away with the bank manager, and my money. Obviously I really regret doing this. If I hadn't done it, I could have done all sorts of things and gone to all sorts of places. For a start, if I had kept only half of the money, I could have started up my own business and I could have made myself the managing director.

If I had made myself the managing director, I could have directed the company for half of the time, and I could have travelled the world for the other half. If I had done this, I would have bought a boat and travelled round the Caribbean. But, unfortunately, I did give her the money, so I've been on holiday in Derby for the past ten years

I travelled around on a bicycle for 10 years. If I'd had the money, I could have bought myself a company car – a Porsche, for example, and maybe I would have bought a large house in Scotland. I could have gone walking in the summer and skiing in the winter.

If I hadn't spent it all on the business, I could have helped my family as well. I could have bought my mum a new house and I could have taken her on holiday every year. I also would have invested some money for both me and my mum so we could have retired on the interest.

But best of all, if I hadn't given all the money to my friend, she wouldn't have been able to share it with Ernie the bank manager, and I wouldn't have felt stupid for so long.

## Group 2: task 1

Andy Harris. DPhil Research. A.j.harris@uel.ac.uk

**10-15 Words or Phrases**

1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
11.
12.
13.
14.
15.

**Writing:** **"Ten years ago I won 14.4 million pounds, and I gave it all to my best friend. Unfortunately, after I gave him/her the money, she/he ran away with the bank manager, and my money. Obviously I really regret doing this, but imagine if…"**

# Appendix H

# Materials for the Control Group (Group Three, Task One)

## Group 3: task 1

**Andy Harris. DPhil Research. A.j.harris@uel.ac.uk**

**Name:**

## Your Task: 'Money and the Bank Manager'

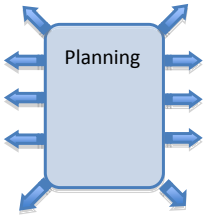
**Scenario:** *Ten years ago you won the lottery - £14.4 million pounds to be exact. You were, obviously, very happy and made a lot of plans to buy a football team, buy new houses, stop your job and travel the world. So that you didn't have to give any money to the taxman, you gave all your money – all £14.4 million - to your best friend. Two weeks later, she/he ran away with the bank manager and all your money. Obviously you feel very stupid and regret what you did.*

I would like you to look back ten years and imagine a hypothetical situation. In this situation I would like you to imagine that you hadn't given your friend the money and imagine the changes in your life ten years ago. In 15 minutes time, I would like you to write as much as you can on the piece of paper provided. Begin your writing with:

**"Ten years ago I won 14.4 million pounds, and I gave it all to my best friend. Unfortunately, after I gave him/her the money, she/he ran away with the bank manager, and my money. Obviously I really regret doing this, but imagine if…"**

Before you start writing, I would like you to first plan your writing.

**Planning:** You have 8 minutes to plan what you are going to write. Read the task again very carefully and write down ideas, words and phrases that will help you when you write it properly. You can use a dictionary if you want to, and you can keep this piece of paper when you write.



## Group 3: task 1

Andy Harris. DPhil Research. A.j.harris@uel.ac.uk

**Writing:** **Ten years ago I won 14.4 million pounds, and I gave it all to my best friend. Unfortunately, after I gave her/him the money, she/he ran away with the bank manager, and my money. Obviously I really regret doing this, but imagine if…**

# Appendix I

# Data for the partial and full selection of result and condition clauses from the input

**Table I.1:** *Raw counts and percentages for partial and full selection of condition and result clauses from the input for Analogous Group (3 tasks).*

| | | **Selection of Result Clauses** | | | | **Selection of Condition Clauses** | | | |
|---|---|---|---|---|---|---|---|---|---|
| Participant | Total Selections | Partial (raw) | Full (raw) | Partial (%) | Full (%) | Partial (raw) | Full (raw) | Partial (%) | Full (%) |
| ST1 | 32 | 0 | 23 | 0 | 71.88 | 0 | 7 | 0.00 | 21.88 |
| ST2 | 15 | 2 | 1 | 13.33 | 6.67 | 0 | 0 | 0.00 | 0 |
| ST3 | 27 | 3 | 3 | 11.11 | 11.11 | 0 | 5 | 0.00 | 18.52 |
| ST4 | 14 | 0 | 6 | 0 | 42.86 | 1 | 4 | 7.14 | 35.71 |
| ST5 | 21 | 2 | 0 | 9.52 | 0 | 1 | 0 | 4.76 | 4.76 |
| ST6 | 28 | 3 | 6 | 10.71 | 21.43 | 1 | 4 | 3.57 | 14.29 |
| ST7 | 30 | 10 | 0 | 33.33 | 0 | 0 | 0 | 0.00 | 0 |
| ST8 | 15 | 0 | 8 | 0 | 53.33 | 1 | 8 | 6.67 | 46.67 |
| ST9 | 21 | 1 | 11 | 4.76 | 52.38 | 0 | 9 | 0.00 | 42.86 |
| ST10 | 48 | 8 | 0 | 16.67 | 0 | 1 | 0 | 2.08 | 0 |
| ST11 | 3 | 0 | 1 | 0 | 33.33 | 0 | 1 | 0.00 | 33.33 |
| ST12 | 29 | 7 | 7 | 24.14 | 24.14 | 1 | 0 | 3.45 | 3.45 |
| ST13 | 2 | 1 | 1 | 50 | 50 | 0 | 1 | 0.00 | 50 |
| ST14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 | 0 |
| ST15 | 28 | 0 | 3 | 0 | 10.71 | 3 | 2 | 10.71 | 14.29 |
| ST16 | 23 | 0 | 7 | 0 | 30.43 | 0 | 7 | 0.00 | 30.43 |
| ST17 | 28 | 4 | 5 | 14.29 | 17.86 | 11 | 1 | 39.29 | 32.14 |
| ST18 | 13 | 0 | 1 | 0 | 7.69 | 0 | 0 | 0.00 | 0 |
| ST19 | 25 | 10 | 1 | 40 | 4 | 1 | 2 | 4.00 | 12 |
| ST20 | 16 | 3 | 4 | 18.75 | 25 | 1 | 2 | 6.25 | 6.25 |
| ST21 | 23 | 6 | 5 | 26.09 | 21.74 | 0 | 2 | 0.00 | 8.7 |
| ST22 | 20 | 1 | 10 | 5 | 50 | 0 | 3 | 0.00 | 15 |
| ST23 | 22 | 3 | 5 | 13.64 | 22.73 | 0 | 4 | 0.00 | 18.18 |
| ST24 | 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 | 0 |
| ST25 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 | 0 |
| ST26 | 37 | 7 | 11 | 18.92 | 29.73 | 0 | 2 | 0.00 | 5.41 |
| ST27 | 11 | 0 | 6 | 0 | 54.55 | 0 | 2 | 0.00 | 18.18 |
| ST28 | 32 | 3 | 9 | 9.38 | 28.13 | 0 | 6 | 0.00 | 18.75 |
| ST29 | 22 | 0 | 13 | 0 | 59.09 | 0 | 3 | 0.00 | 13.64 |
| ST30 | 15 | 2 | 3 | 13.33 | 20 | 0 | 1 | 0.00 | 6.67 |
| ST31 | 21 | 1 | 4 | 4.76 | 19.05 | 0 | 0 | 0.00 | 0 |
| ST32 | 26 | 0 | 3 | 0 | 11.54 | 0 | 1 | 0.00 | 3.85 |
| ST33 | 41 | 1 | 2 | 2.44 | 4.88 | 0 | 3 | 0.00 | 7.32 |

**Table I.2:** *Raw counts and percentages of partial and full selection of condition and result clauses from the input for Literal Group (3 tasks).*

| | | **Selection of Result Clauses** | | | | **Selection of Condition Clauses** | | | |
|---|---|---|---|---|---|---|---|---|---|
| Participant | Total Selections | Partial (raw) | Full (raw) | Partial (%) | Full (%) | Partial (raw) | Full (raw) | Partial (%) | Full (%) |
| ST1 | 33 | 6 | 4 | 18.18 | 12.12 | 1 | 2 | 3.03 | 3.03 |
| ST2 | 45 | 5 | 0 | 11.11 | 0 | 1 | 0 | 2.22 | 0 |
| ST3 | 14 | 0 | 8 | 0 | 57.14 | 0 | 3 | 0.00 | 21.43 |
| ST4 | 26 | 1 | 7 | 3.85 | 26.92 | 0 | 6 | 0.00 | 23.08 |
| ST5 | 36 | 6 | 12 | 16.67 | 33.33 | 4 | 0 | 11.11 | 8.33 |
| ST6 | 46 | 5 | 0 | 10.87 | 0 | 0 | 0 | 0.00 | 0 |
| ST7 | 31 | 3 | 0 | 9.68 | 0 | 1 | 0 | 3.23 | 0 |
| ST8 | 39 | 9 | 1 | 23.08 | 2.56 | 0 | 1 | 0.00 | 2.56 |
| ST9 | 11 | 4 | 1 | 36.36 | 9.09 | 0 | 0 | 0.00 | 0 |
| ST10 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 | 0 |
| ST11 | 39 | 10 | 4 | 25.64 | 10.26 | 0 | 0 | 0.00 | 0 |
| ST12 | 41 | 3 | 0 | 7.32 | 0 | 0 | 0 | 0.00 | 0 |
| ST13 | 22 | 2 | 0 | 9.09 | 0 | 0 | 0 | 0.00 | 0 |
| ST14 | 49 | 7 | 1 | 14.29 | 2.04 | 0 | 0 | 0.00 | 0 |
| ST15 | 18 | 0 | 0 | 0 | 0 | 0 | 1 | 0.00 | 5.56 |
| ST16 | 24 | 6 | 6 | 25 | 25 | 0 | 2 | 0.00 | 8.33 |
| ST17 | 28 | 0 | 12 | 0 | 42.86 | 1 | 9 | 3.57 | 28.57 |
| ST18 | 45 | 1 | 0 | 2.22 | 0 | 0 | 0 | 0.00 | 0 |
| ST19 | 33 | 9 | 11 | 27.27 | 33.33 | 1 | 2 | 3.03 | 3.03 |
| ST20 | 36 | 6 | 8 | 16.67 | 22.22 | 2 | 6 | 5.56 | 11.11 |
| ST21 | 45 | 4 | 22 | 8.89 | 48.89 | 5 | 7 | 11.11 | 22.22 |
| ST22 | 21 | 3 | 2 | 14.29 | 9.52 | 1 | 5 | 4.76 | 28.57 |
| ST23 | 27 | 3 | 15 | 11.11 | 55.56 | 0 | 9 | 0.00 | 33.33 |
| ST24 | 40 | 21 | 2 | 52.5 | 5 | 1 | 0 | 2.50 | 2.5 |
| ST25 | 34 | 8 | 4 | 23.53 | 11.76 | 0 | 0 | 0.00 | 0 |
| ST26 | 32 | 9 | 11 | 28.13 | 34.38 | 2 | 3 | 6.25 | 15.63 |
| ST27 | 26 | 3 | 5 | 11.54 | 19.23 | 2 | 5 | 7.69 | 15.38 |
| ST28 | 48 | 5 | 17 | 10.42 | 35.42 | 0 | 9 | 0.00 | 18.75 |
| ST29 | 38 | 5 | 3 | 13.16 | 7.89 | 5 | 1 | 13.16 | 13.16 |
| ST30 | 37 | 4 | 12 | 10.81 | 32.43 | 1 | 6 | 2.70 | 13.51 |
| ST31 | 39 | 14 | 5 | 35.9 | 12.82 | 1 | 0 | 2.56 | 2.56 |