University of Stirling

Faculty of Social Sciences

MSc in Teaching English to Speakers of Other Languages and Computer Assisted Language Learning

**A pilot study for an investigation into the effectiveness of receptive-based and productive-based instruction for absolute beginners of Scottish Gaelic using CALL techniques.**

August 2017

St­udent number: 2536385

***Do Chaluim***

*Fear cho tapaidh ’s brosnachail ris an tachair mi riamh.*

*Cha b’ urrainn dhomh Gàidhlig a thogail cho math no cho luath gun do chuideachadh, misneachadh agus – thar a h-uile sìon eile – foighidinn.*

*Cha robh neach eile ann an Dùn Èideann a rinn an urad riut-sa airson luchd-ionnsachaidh a’ bhaile.*

*Tha an obair seo na pàirt de do dhìleab, agus tha mi an dòchas gum bi cothrom dhomh dèanamh airson daoine eile na rinn thusa air mo shonsa.*

*Bidh sinn gad ionndrainn.*

***Malcolm “Calum” Cameron***

***22nd December 1923 - 19th January 2017***

Abstract

Online language learning has become a major growth area in the educational sector in recent years, but as it is dominated by large companies and closed systems, it is not yet widely researched. Much literature exists that is relevant to the styles of teaching delivered online, in particular research into the relative benefits of production-based instruction (PBI) and receptive-based instruction. However, this research rarely addresses the needs of the true beginner, and much internet-based material is targeted at this group.

This paper describes a pilot study conducted in July 2017 which involved the development of a simple on-line language learning course to investigate the effectiveness of production-based instruction (PBI) and comprehension-based instruction (CBI), and a mixture of the two, for absolute beginner learners of the Scottish Gaelic language. 31 participants were given access to one of four variants of the teaching materials and followed a programmed course for 4 days, followed by an immediate post-test on the fifth day and a delayed post-test a week later. Test scores were analysed using descriptive statistics to identify overarching trends, and a closer analysis of error types and individual errors was carried out to look for underlying causes, as well as evidence of partial acquisition. It was found that the PBI group obtained the best results in all test activities, although the difference in scores in receptive tests was marginal. The CBI group showed the poorest performance, although again the difference in receptive tests was marginal.

The generalisability of the findings of this study are limited by the scale of the study, both in terms of the number of participants and in terms of timescale; a longitudinal study would be required to see if the effects noted in this study have an effect on the long-term retention of the language studied.

Acknowledgements

I would like to express my gratitude to all the teaching and administration staff of the TESOL masters programme, in particular my supervisor Dr. Edward Moran, whose advice and support, not to mention patience, was crucial in making this project possible.

To the hundreds of strangers who volunteered to take part in this study, and the dozens of others who helped distribute my call for participants through their social networks, I do not have adequate words to express my gratitude – though not one of my research questions, I believe this study has proven a great deal about the kindness of strangers.

My thanks in particular go to the participants who during and after the study took the time to write to me and thank me for the opportunity to take part in this study – it is genuinely humbling when people who are doing you a huge favour turn round and thank you for it.

Study Materials

The materials used in this study are available at <https://lingua.guru/login> . Note that access to this site is controlled by secure user accounts, and a password is needed for each user. Individual accounts have been configured for 3 markers and will be active during the marking period only, as the materials will be taken off-line after marking is complete (due to the copyright restrictions noted below).

The usernames and passwords are as follows:

|  |  |  |
| --- | --- | --- |
| Marker | Username | Password |
| First marker | marker1 | TESOLdissert2017 |
| Second marker | marker2 | TESOLdissert2017 |
| Third marker | marker3 | TESOLdissert2017 |

Detailed access and usage instructions for markers are included in Appendix K at the end of this document.

Copyright Notice

All graphical images used in module 1 were created by the author/researcher and are copyright 2017 student 2536385. Although not discussed in this paper, a number of additional images were used in module 2. 4 of these were obtained from free sources on the internet – details and credits are provided in Appendix L.

Sound files were created using the synthesised voice *Ceitidh* by CereProc Ltd. under an academic license obtained via CALL Scotland and the University of Edinburgh (<http://www.thescottishvoice.org.uk/licence/>). As this license is for use by students at Scottish universities, the materials will be taken offline when marking is complete.

Glossary

CALL Computer-assisted language learning

CBI Comprehension-based instruction

IQR Interquartile range

PBI Production-based instruction

SCT Socio-cultural theory

SD Standard deviation

SLA Second language acquisition

ZPD Zone of proximal development

Contents

[Chapter 1. Introduction 1](#_Toc490641820)

[1.1. Background 1](#_Toc490641821)

[1.2. Background Theory 1](#_Toc490641822)

[1.3. Statement of the Problem 2](#_Toc490641823)

[1.4. Choice of Language 2](#_Toc490641824)

[1.5. Purpose of the Study 3](#_Toc490641825)

[1.6. Definition of key terms 3](#_Toc490641826)

[1.6.1. Second Language Acquisition 3](#_Toc490641827)

[1.6.2. Computer-Assisted Language Learning 3](#_Toc490641828)

[1.6.3. Productive and Receptive Skills 4](#_Toc490641829)

[1.6.4. Comprehension-Based Instruction and Production-Based Instruction 4](#_Toc490641830)

[1.6.5. Mediation: Self-Regulation and Other-Regulation 4](#_Toc490641831)

[1.7. Summary of Content and Structure 4](#_Toc490641832)

[Chapter 2. Literature Review 5](#_Toc490641833)

[2.1. Introduction 5](#_Toc490641834)

[2.2. Theories of Learning 5](#_Toc490641835)

[2.2.1. Behaviourism and Cognitivism 5](#_Toc490641836)

[2.2.2. Socio-Cultural Theory and Dual Coding Theory 6](#_Toc490641837)

[2.3. The Field of Second Language Acquisition 7](#_Toc490641838)

[2.3.1. Comprehensible Input and Comprehensible Output 7](#_Toc490641839)

[2.3.2. Focus-on-Form 8](#_Toc490641840)

[2.3.3. CBI and PBI 9](#_Toc490641841)

[2.3.4. Summary 11](#_Toc490641842)

[2.4. Computer-Assisted Language Learning 11](#_Toc490641843)

[2.4.1. The Tutorial Role of CALL in Controlled Practice 11](#_Toc490641844)

[2.4.2. CALL and the Internet as a Research Tool 11](#_Toc490641845)

[2.5. Summary 12](#_Toc490641846)

[2.6. Research questions, hypothesis and rationale 13](#_Toc490641847)

[Chapter 3. Materials Design 15](#_Toc490641848)

[3.1. Introduction 15](#_Toc490641849)

[3.2. Software Platform 15](#_Toc490641850)

[3.2.1. Server 15](#_Toc490641851)

[3.2.2. Client 15](#_Toc490641852)

[3.3. Procedural Generation of Content 16](#_Toc490641853)

[3.4. Learning Content 17](#_Toc490641854)

[3.4.1. Visual Design Principles 17](#_Toc490641855)

[3.4.2. Task Presentation and Feedback 19](#_Toc490641856)

[3.4.3. Module Organisation 22](#_Toc490641857)

[Chapter 4. Methods and Procedures 26](#_Toc490641858)

[4.1. Introduction 26](#_Toc490641859)

[4.2. Methods 26](#_Toc490641860)

[4.2.1. Experimental Design 26](#_Toc490641861)

[4.2.2. Descriptive Statistics 27](#_Toc490641862)

[4.3. Recruitment of Participants 28](#_Toc490641863)

[4.4. Study Design 28](#_Toc490641864)

[4.4.1. Target Language Items 29](#_Toc490641865)

[4.5. Data Gathering and Participants’ Use of the System 31](#_Toc490641866)

[4.5.1. Pre-test 31](#_Toc490641867)

[4.5.2. Assigning Participants to Treatment Groups 31](#_Toc490641868)

[4.5.3. Study Content 31](#_Toc490641869)

[4.5.4. Testing Strategy 32](#_Toc490641870)

[4.6. Analysis of the Data 32](#_Toc490641871)

[4.6.1. The Python Programming Language 33](#_Toc490641872)

[4.6.2. Receptive Post-tests 33](#_Toc490641873)

[4.6.3. Productive post-tests 34](#_Toc490641874)

[4.7. Validity and Reliability 35](#_Toc490641875)

[4.7.1. Validity 36](#_Toc490641876)

[4.7.2. Reliability 37](#_Toc490641877)

[4.7.3. Mitigating Dropout Rates 37](#_Toc490641878)

[4.8. Ethical Considerations 37](#_Toc490641879)

[4.9. Conclusion 38](#_Toc490641880)

[Chapter 5. Results and Discussion 39](#_Toc490641881)

[5.1. Introduction 39](#_Toc490641882)

[5.2. Description of the Study Sample 39](#_Toc490641883)

[5.3. Receptive Skills: 40](#_Toc490641884)

[5.3.1. Analysis of Scores 41](#_Toc490641885)

[5.3.2. Analysis of Errors 42](#_Toc490641886)

[5.3.3. Discussion: RQ1 42](#_Toc490641887)

[5.3.4. Discussion: RQ3 43](#_Toc490641888)

[5.4. Productive Skills 43](#_Toc490641889)

[5.4.1. Analysis of Scores 43](#_Toc490641890)

[5.4.2. Analysis of Errors 46](#_Toc490641891)

[5.4.3. Discussion: RQ2 49](#_Toc490641892)

[5.4.4. Discussion: RQ4 50](#_Toc490641893)

[5.5. Summary 50](#_Toc490641894)

[Chapter 6. Conclusion 51](#_Toc490641895)

[6.1. Summary of Findings 51](#_Toc490641896)

[6.1.1. RQ1 51](#_Toc490641897)

[6.1.2. RQ2 51](#_Toc490641898)

[6.1.3. RQs 3 and 4 51](#_Toc490641899)

[6.2. Pedagogical Implications 51](#_Toc490641900)

[6.3. Limitations and Delimitations 52](#_Toc490641901)

[6.4. Future Research 53](#_Toc490641902)

[6.5. Conclusion 54](#_Toc490641903)

[References 55](#_Toc490641904)

[Appendix A. Glossary of Gaelic used in this study 62](#_Toc490641905)

[Appendix B. Sign-up survey 63](#_Toc490641906)

[B.1. Page 1 63](#_Toc490641907)

[B.2. Page 2 – Background 63](#_Toc490641908)

[B.3. Page 3 – Contact details 64](#_Toc490641909)

[B.4. Page 4 – Thank you for your participation. 64](#_Toc490641910)

[Appendix C. Study participant demographics 65](#_Toc490641911)

[C.1. Age, gender and language background 65](#_Toc490641912)

[C.1.1. Genuine beginners who completed module 1 65](#_Toc490641913)

[Appendix D. Emails sent during the study 67](#_Toc490641914)

[D.1. Welcome email/login details 67](#_Toc490641915)

[D.2. Invitation to complete immediate post-test 67](#_Toc490641916)

[D.3. Invitation to complete delayed post-test 68](#_Toc490641917)

[Appendix E. Summary of participants’ scores by group. 69](#_Toc490641918)

[E.1. Section I: Productive (before receptive) 70](#_Toc490641919)

[E.1.1. Overall results (minor typographical correcting) 70](#_Toc490641920)

[E.1.2. Scores disregarding spelling errors 70](#_Toc490641921)

[E.2. Section II: Receptive 71](#_Toc490641922)

[E.3. Section III: Productive 72](#_Toc490641923)

[E.3.1. Base scores (minor typographical errors corrected) 72](#_Toc490641924)

[E.3.2. Scores, disregarding spelling errors 72](#_Toc490641925)

[Appendix F. User improvement/decline: 73](#_Toc490641926)

[Appendix G. Individual participants’ scores 76](#_Toc490641927)

[Appendix H. Individual participants’ errors (receptive) 79](#_Toc490641928)

[H.1. Immediate post-test 80](#_Toc490641929)

[H.2. Delayed post-test 81](#_Toc490641930)

[Appendix I. Individual participants’ errors (productive) 83](#_Toc490641931)

[I.1. Errors of noun-adjective word order 85](#_Toc490641932)

[I.2. Detail of errors coded Z (other) 86](#_Toc490641933)

[I.2.1. Immediate post-test 86](#_Toc490641934)

[I.2.2. Delayed post-test 88](#_Toc490641935)

[Appendix J. Teaching and test materials 97](#_Toc490641936)

[J.1. Pretest 97](#_Toc490641937)

[J.1.1. Content 97](#_Toc490641938)

[J.2. Module 1 99](#_Toc490641939)

[J.2.1. Learning materials 99](#_Toc490641940)

[J.3. Post-tests 103](#_Toc490641941)

[J.3.1. Part I: productive skills 103](#_Toc490641942)

[J.3.2. Part II: receptive skills 104](#_Toc490641943)

[J.3.3. Part III: productive skills 104](#_Toc490641944)

[Appendix K. Access instructions 105](#_Toc490641945)

[Appendix L. Copyright information 108](#_Toc490641946)

**Word count: 16494**

List of tables

[Table 4‑1: treatments by group and module 28](#_Toc490639527)

[Table 5‑1: Group sizes and treatments (module 1) 39](#_Toc490639528)

[Table 5‑2: distribution of age ranges across treatment groups in final sample. 40](#_Toc490639529)

[Table 5‑3: gender distribution across treatment groups 40](#_Toc490639530)

[Table E‑1: Scores by group, immediate post-test, Part I (productive skills) 70](#_Toc490639531)

[Table E‑2: Scores by group, delayed post-test, Part I (productive skills) 70](#_Toc490639532)

[Table E‑3: Scores by group, immediate post-test, Part I (productive skills) (Spelling errors discounted) 70](#_Toc490639533)

[Table E‑4: Scores by group, delayed post-test, Part I (productive skills) (Spelling errors discounted) 70](#_Toc490639534)

[Table E‑5: Scores by group, immediate post-test, Part II (receptive skills) 71](#_Toc490639535)

[Table E‑6: Scores by group, delayed post-test, Part II (receptive skills) 71](#_Toc490639536)

[Table E‑7: Scores by group, immediate post-test, Part III (productive skills) 72](#_Toc490639537)

[Table E‑8: Scores by group, delayed post-test, Part III (productive skills) 72](#_Toc490639538)

[Table E‑9: Scores by group, immediate post-test, Part III (productive skills) (spelling errors discounted) 72](#_Toc490639539)

[Table E‑10: Scores by group, immediate post-test, Part III (receptive skills) (spelling errors discounted) 72](#_Toc490639540)

[Table F‑11: Part I: Productive post-tests: relative performance by group. 74](#_Toc490639541)

List of figures

[Figure 3‑1: Comparison of images representing noun phrases and bare nouns 18](#_Toc490639450)

[Figure 3‑2: An example of a multiple choice question 19](#_Toc490639451)

[Figure 3‑3: Interaction sequence (including error in response) for productive question type. 20](#_Toc490639452)

[Figure 3‑4: Interaction sequence (including error in response) for receptive question type. 22](#_Toc490639453)

[Figure 3‑5: Example of staged introduction and practice. 23](#_Toc490639454)

[Figure 3‑6: Order of presentation of grammatical contrasts (module 1, unit 2) 24](#_Toc490639455)

[Figure 3‑7: Decreasing scaffolding within a unit of CBI 25](#_Toc490639456)

[Figure 4‑1: Annotated example of boxplot generated using Matplotlib 33](#_Toc490639457)

[Figure 5‑1: Post-test scores for receptive skills test (immediate and delayed) 41](#_Toc490639458)

[Figure 5‑2: Comparison of immediate and delayed post-tests of productive skills for P and R treatments 45](#_Toc490639459)

[Figure 5‑3: Comparison of post-test scores with spelling errors disregarded 46](#_Toc490639460)

[Figure 5‑4: A visual mapping of productive error types by group. 47](#_Toc490639461)

# Introduction

This chapter sets out the background for the present study, first in terms of the researcher’s personal motivation, and then in terms of the academic research underpinning the study. It then proceeds to discuss the specific problem the present study seeks to address: the lack of research in this area relating to the ab initio learner; and then describe the purpose of the study. A brief justification of the language used in the study is then presented, after which key terms relevant to the present study are defined. Finally, a brief outline of the structure and content of this document is given.

## Background

As a graduate in computer science and a keen learner of languages, I have used a great number of different computer tools in my own studies, and evaluated a number of different packages for use by my students. Many of the beginners’ resources I have worked with have focused on practice of short, isolated sentences, but while some (Rosetta Stone 2017a) focus primarily on teaching by testing learner comprehension of the target language, others (TeachMe! 2017) teach primarily through forced early production and others still (Duolingo 2017) use a mixture of the two. Such materials are becoming more and more part of the mainstream of language learning, with Babble and Duolingo announcing subscriber bases of over 1 million (Speigel Online 2016) and over 110 million in 2016 (TechInAsia.com 2016) respectively, and Rosetta Stone claiming to be used by over 22000 educational institutions worldwide (Rosetta Stone 2017a).

## Background Theory

The effects on acquisition of productive and receptive practice has long been an area of disagreement within the second language acquisition community. Perhaps the most famous instance of this is the opposition between Stephen Krashen’s (1977) *input hypothesis*, which suggests that language is only acquired through understanding messages, and Merrill Swain’s (1985) *output hypothesis*, which suggests instead that output of language is necessary to allow learners to evaluate and modify their theories on the form of the target language.

Shintani et al. (2013) more recently carried out a review of literature relating to acquisition of grammar which identified a general tendency for subjects taught receptively to perform better in receptive tasks in post-tests and similarly a tendency for productively-taught subjects to perform better in productive tasks, but found no evidence that suggested that either method alone was superior overall.

## Statement of the Problem

The focus of previous research into the relative effectiveness of comprehension-based and production-based instruction (see section 1.6.4) has typically investigated its effects on intermediate and advanced-level learners. The applicability of this to the ab initio learner is not clear – previous studies (Shintani et al. 2013) have taken learners’ previous study and used their existing knowledge of the language’s vocabulary and grammar to allow the isolation of a narrow range of linguistic variables. The absolute beginner, however, will be forced to attend to multiple lexical and grammatical variables simultaneously, and the increased demands on learner attention are likely to have a large effect on the learning process.

Furthermore, commercial language software is not generally open to adaptation by teachers or researchers; style of teaching intervention, language content and order of teaching are all fixed, making it impossible to research the effectiveness of the teaching methods without other variables confounding the results. Investigations of such software are restricted to measuring the overall effectiveness against an external benchmark, such as the WebCAPE test used for assessing college entrants in the USA, and are often commissioned by the publishers of the software themselves (e.g. Vesselinov 2009; Vesselinov and Grego 2012). Commercially available software is therefore insufficient to investigate teaching effects, and it is necessary to produce research tools specifically designed to isolate the variables under study.

## Choice of Language

The present study uses the teaching of Scottish Gaelic (henceforth “Gaelic”) to research general principles of language learning for beginners. It is increasingly difficult to find true beginners of more widely-taught languages such as English, especially when conducting research online: 51.4% of all websites are available in English whereas less than 0.1% of all sites offer Gaelic versions (w3techs.com 2017) which means most internet users are likely to have had considerable passive exposure to English, but not to Gaelic. Robin (2013) notes that computer-based resources are rare and hard to find for most of what he terms *less commonly taught languages*, and this is certainly the case for Gaelic. The European Commission recently published a list of the best free language learning websites (EPALE 2016), and none of the sites listed offer any resources for learning Gaelic; popular paid-for sites Rosetta Stone (2017b) and Babbel.com (2017) similarly offer no courses in Gaelic. Learners seeking this type of material are therefore likely to readily volunteer to take part in a study offering them.

## Purpose of the Study

The present study aims to investigate and compare the effects of exclusively production-based and exclusively comprehension-based instruction and mixed approaches on the simultaneous acquisition of grammar and vocabulary for absolute beginners of Gaelic using computer-based learning tools.

## Definition of key terms

### Second Language Acquisition

Ortega defines second language acquisition (SLA) as:

“the scholarly field of enquiry that investigates the human capacity to learn languages other than the first, during late childhood, adolescence or adulthood, and once the first language or languages have been acquired” (Ortega 2009, pp.1-2)

This definition is necessarily expansive, given the breadth of competing theories within SLA: Ortega includes both formal instruction and naturalistic learning in this definition, and explicitly rejects Krashen’s (1985) acquisition-learning hypothesis which distinguishes between acquisition as a natural internal process and learning as the conscious study and conscious recall of memorised rules (Ortega 2009, pp.136-138). “[M]ost researchers in the field do not sustain any principled distinction between the two terms” (Mitchell et al. 2013, p.2). The present study therefore makes no such distinction and uses the terms *acquisition* and *learning* interchangeably.

### Computer-Assisted Language Learning

Computer-assisted language learning (CALL) has been defined as:

“learning language in any context with, through and around computer technologies” (Egbert 2005, p.4)

and as:

“any process in which a learner uses a computer and, as a result, improves his or her language” (Beatty 2010, p.7)

These are very broad definitions, recognising the wide potential of technology.

The materials in this study represent only a fraction of this potential, and focuses on “the special nature of multimedia” (Beatty 2010, p.158) to provide interactive materials incorporate images and audio and providing supportive feedback.

### Productive and Receptive Skills

Skills in language are commonly split into *productive* and *receptive* skills (Scrivener 2011, p.26). Productive skills are those of speaking and writing, where the learner outputs the target language; receptive skills are those of listening and reading, where the learner is required to comprehend the target language (*ibid.*).

The present study restricts itself to the written mode; in this paper, “productive” refers exclusively to writing and “receptive” to reading.

### Comprehension-Based Instruction and Production-Based Instruction

Shintani and Ellis (2010) make a distinction between *comprehension-based instruction* (CBI), where receptive-skills tasks are used to encourage acquisition, and *production-based instruction* (PBI) where productive tasks are used for the same purpose. The word “instruction” is used here to indicate a directed process where the teacher of course designer has specific targets for acquisition; VanPatten (1996) argues that this is necessary to encourage noticing of grammatical features; and whether or not this is the case, such structured learning is easily observable, measurable and replicable.

### Mediation: Self-Regulation and Other-Regulation

Sociocultural theory (see section 2.2.2) uses the term *mediation* to describe how tools, concepts and symbols can be manipulated to regulate mental activity, hence learning (Lantolf 2011, p.24; Mitchell et al. 2013, p.221). Where the manipulation is carried out by the learner, this is termed *self-regulation*, and where the manipulation is carried out by another person, this is termed *other-regulation* (Mitchell et al. 2013, p.222).

## Summary of Content and Structure

This paper presents a brief review of the literature relevant to the present study (Chapter 2) followed by a description of the software-based materials used in the research (Chapter 3). After this, the methods of research are described, including both the experiment and analysis (Chapter 4), followed by presentation and analysis of results (Chapter 5). Finally, the paper concludes with a summary of the results and limitations, alongside suggestions and recommendations for future research (Chapter 6).

# Literature Review

## Introduction

This chapter presents an overview of previous research that is relevant to the present study, before presenting the research questions to be addressed in this study.

This starts with an overview of several overarching theories of learning, then proceeds to discuss theories specific to SLA. A discussion then follows of the literature relating to CALL, followed by a brief summary and a presentation of the research questions addressed in the present study.

## Theories of Learning

Prior to the modern era, psychological models were predominantly built on the idea that human thought was symbolic in nature, with schools such as *Associationism* placing all thought and reasoning as being built on associations between different symbols that model the physical world; this was later challenged by the Wurzburg School that promoted the idea of *imageless thought* based on internal states (French and Colman 1995, p.59). These theories all assumed that thought could be consciously observed by the thinker, but this was challenged by later theories that sought to replace introspection with objectively verifiable external observation (*ibid*, p.60).

Two of these movements, Behaviorism and Cognitivism, are described below, after which the discussion turns to two theories, Dual Coding Theory and Socio-Cultural Theory, that propose that while lower-order thought is not symbolic in nature, higher-order human reasoning is.

### Behaviourism and Cognitivism

In the early twentieth century, J.B. Watson sought to challenge psychology’s reliance on introspection, and replace it with “a purely objective experimental branch of natural science” (Watson 1913) which he termed *Behaviorism*; Behaviorists argued that all thought was merely a complex form of behaviour, and subjects’ response to stimulus was seen as objectively verifiable data (French and Colman 1995, pp.59-60). Learning under this framework was taken to be a process of associating perceived stimulus with the most appropriate response by *reinforcement*, where the learner receives some kind of reward or positive feedback after performing the desired behaviour; this became known as *stimulus-response (SR) theory* (French and Colman 1995, pp.59-60). Applying these theories to education, Skinner (1968) suggested that teaching should occur as a series of small steps, and that feedback should be immediate in all stages of learning.

Later researchers rejected the notion that thought arose simply as a complex behaviour in response to stimulus. This school was termed *Cognitivism* as unlike Behaviorism it sought to investigate internal mental processes. Cognitivists rejected behaviourist principles as overly simplistic: Bruner suggested that SR theory was only valid in infant learning and in situations where “the psychologist’s ingenious experimental apparatus … prevents anything from happening but an overt response” (Bruner 1966, p.18); and Ausubel criticised his contemporary Gagné for creating a learning model centred on SR theory which he described as “effective with infrahumans and rote learning of nonsense materials,” and “not supportable as an approach for meaningful learning” (Ausubel et al. 1978, p.362).

Cognitivism criticises behaviourist pedagogy for encouraging rote learning, and suggests instead that learning must be meaningful to be internalised. However, opinions differ between Cognitivists as to how this may be achieved. Ausubel suggested that *reception learning*, learning where “the entire content of what is to be learned is presented to the learner in final form” (Ausubel et al. 1978, p.24), may be meaningful if presented correctly, whereas Bruner (1966) proposed that only *discovery learning*, where learners discover target rules through investigation was truly meaningful; but Ausubel (*op. cit.*) countered that a discovery activity could be contrived where learning remained rote. Ausubel also stated that for reception to be meaningful, deliberate practice must be engaged in, with feedback that is not merely positive or negative, as in SR theory, but that provides “cognitive or *informational* effects” in order to “increase the subject’s confidence in the correctness of what is understood” and facilitate the students’ own selective focus on their weaknesses (*op. cit.,* p.423).

Bruner suggested that teaching involved a process of *scaffolding*, which he defined as: “the steps taken to reduce the degrees of freedom in carrying out some task so that the child can concentrate on the difficult skill she is in the process of acquiring” (Bruner 1978, p. 19); thus the teacher must optimise the learning environment and order of teaching to reduce the likelihood of incorrect generalisation.

### Socio-Cultural Theory and Dual Coding Theory

Socio-cultural theory (SCT) draws on the work of Vygotsky.

A core part of Vygotsky’s theory draws on the ability of a teacher to guide a student by drawing their attention to salient features, which inspired Bruner’s notion of scaffolding (Kyriacou 2009, p.30), and perhaps understandably many texts for teachers tend to focus on his definition of the Zone of Proximal Development (ZPD) as:

“The distance between the actual development level as determined through problem solving under adult guidance or in collaboration with more capable peers.” (Vygotsky 1978, p.86; cited in Capel et al. 2013, p.57)

However, while this notion of peer assistance as *other-regulation* is crucial to his theories, an equally important component of SCT is that of *self-regulation*, where higher-order symbolic thought is used by sophisticated learners to further their own learning (Mitchell et al. 2013 p.222).

Paivio (1971) supports this view, and expands on it with *dual coding theory* (DCT), which proposes that the brain maintains two distinct, interconnected symbolic systems: *verbal* or linguistic, and *non-verbal*, based on sensory images. This gives rise to two types of associations between symbols: *referential associations,* links between entities in different systems (e.g. associating the word “car” with the mental concept describing cars); and *associative associations* between items within the same system (e.g. associating English “book” with French “livre”; or the visual image of a Bunsen burner with the smell of gas) (Clark and Paivio 1991, pp.151-153).

## The Field of Second Language Acquisition

SLA emerged as an academic field as part of the cognitivist movement in the late 1960s (Atkinson 2011, p.11) but quickly expanded to cover other theories, leading to a “social turn” beginning in the mid-90s as a reaction against cognitivism (Ortega 2011, p167). This turn has provided “theoretical alternatives that the existing cognitive theories could not help us unpack prior to the social turn” (Ortega 2011, p167) and thus is complementary to a cognitive approach. A key part of this social turn was the application of SCT to SLA, termed *SCT-L2*, as per Lantolf (2011), who notes that fully successful completion of tasks gives no insight into mediation strategies, which instead can only be observed when learners have not fully acquired target language items – thus research tasks should be well beyond the subjects’ current level, but within their ZPD (*ibid.*, pp.26-27).

### Comprehensible Input and Comprehensible Output

Krashen’s *input hypothesis* proposes that “We acquire language when we understand messages, when we understand what people tell us and when we understand what we read” (Krashen 1982) and speaks of *comprehensible input* as input that is appropriate to a learner’s level and can therefore be learned from. Krashen suggests that the only role for explicitly taught language is as a “monitor” to allow the learner to self-correct output, but that this will not lead to successful acquisition (*ibid.*). However, Naiman et al. (1978, pp.14-15) identified certain strategies common to successful language learners, and among those were “realization of language as a system” and “monitoring of L2 performance”; while this study was based on self-reported strategies, it seems unlikely that there would be such a consistent mismatch between learners’ strategies and the theory were it true.

Swain countered the input hypothesis with her *output hypothesis* which asserts that production is a necessary part of language acquisition, in that it gives the learner the opportunity to try out their theories of the new language and use feedback from communication partners to confirm or disprove the theory (Swain 1985). In her observations of English speaking students on a French immersion programme, she noted that “aspects of […] second language proficiency cannot be totally accounted for on the basis of input received” (*op. cit.*, p236) and states that “the phenomenon of individuals who can understand a language and yet can only produce limited utterances in it” (*op. cit*., p.249) shows that understanding of meaning alone does not lead to acquisition of language form.

Swain’s hypothesis is more in line with SCT, as the process by which such output occurs is one of symbolic self-regulation, allowing the learner to operate in their ZPD rather than only being able to employ fully-acquired language. Krashen’s notion of a monitor, however, disregards self-regulation, restricting the scope for mediation to a very weak form, with other-regulation in the form of roughly-tuned input as the only path to development.

### Focus-on-Form

Swain argues that:

“comprehensible input is crucial to grammatical acquisition, *not* because the focus is on meaning, *or* because a two-way exchange is occurring, but because by being understood […] it permits the learner to focus on form” (Swain 1985, p.248)

The student must therefore be “pushed” (*op. cit*., p.249) to produce grammatically correct language that goes beyond the need to merely be understood, as grammatically incorrect forms still permit full expression of meaning. She also argues that failure to acquire language in target-language medium classrooms occurs because “content teaching focuses on meaning” whereas “what second language learners need is to focus on form-meaning relationships” (Swain 1988, p.81). Schmidt (1990) presents a similar argument from a psychological perspective, suggesting that conscious thought and conscious noticing are necessary for acquisition of language patterns. These findings are congruent with the self-reported learner strategies noted in Naiman et al.’s (1978) study.

### CBI and PBI

Various models have been developed and terms coined to discuss the difference between directed teaching relying primarily on input or on output.

VanPatten compares the two approaches in a reception-learning setting, with explicit teaching of language followed by controlled output termed “*traditional instruction*” and explicit teaching focused directly on assisting in later comprehension activities termed “*processing instruction*” (VanPatten and Oikkenon 1996; Morgan-Short and Bowden 2006). Morgan-Short and Bowden (*op. cit.*) highlight the importance of meaningfulness in learning tasks through their use of the term *meaningful output-based instruction*. Shintani and Ellis (2010) make a distinction between *comprehension-based instruction* (CBI), where receptive-skills practice is used to encourage acquisition, and *production-based instruction* (PBI) where productive-skill practice is used for the same purpose. A comprehensive review of relevant literature (Shintani et al. 2013) found that overall there was evidence that CBI led to notable gains in receptive skills and PBI led to notable gains in productive skills, but could not conclude that either was better overall.

However, Shintani’s (2012) during study of 15 6-year-old Japanese children learning 36 vocabulary items and the plural -s suffix, she notes that despite the intervention being CBI, subject were observed engaging in spontaneous production of language. Similarly, Morgan-Short and Bowden (2006, p.42), in their study of 45 subject on a first-semester Spanish course at a US university, caution that scheduled review material in their teaching “might have served as incidental input” to the MOBI (or CBI) treatment group. This raises the question of how much of the observed effects is due to the teaching intervention, and how much to this student-student interaction affected the learning outcomes, and whether CBI or other input-based instruction would be less effective if used in a situation that affords no opportunities for student-student interaction.

Erlam’s (2003) study focused on the teaching of additional meanings of known French language forms to 66 high-school students of approximately 14 years old, and employed explicit explanation of language points followed by either structured-input activities or output-based practice base. Unlike many of the other studies included in Shintani et al.’s (2013) review, Erlam highlights that “the output group performed better than the structured-input group on both comprehension tests” (Erlam 2003, p.577) although her conclusions are weaker than this, stating simply that “there is no greater advantage for structured-input instruction over meaning-oriented, output based instruction” and that “the meaning-oriented nature of instruction may play a key role in SLA” (*ibid*., p.579).

Izumi (2002) and Leeser (2008) both investigated the effects of output tasks in addition to input tasks. Izumi’s (2002) study involved teaching relative clauses to 61 intermediate level ESL students in two US universities, and found that members of the group with output were more successful in noticing features in the input tasks, as well as noticing differences between their output and the target forms. In Leeser’s (2008) study of 47 intermediate Spanish learners acquiring past tense forms, students engaging in output in addition to input reported more noticing of vocabulary than input-only students, and although they reported no additional noticing of past forms, they showed better use of them in their written production. Both studies also found that a mixed treatment resulted in better comprehension in input-based post-tests than CBI alone (Izumi 2002; Leeser 2008).

Of all the studies examined by Shintani et al. (2013), Nagata’s (1998) study of 14 second-semester students of Japanese in a US university, made the strongest claim for the value of output, concluding that “given the same grammatical instruction, output-focused practice is more effective than input-focused practice” in teaching Japanese honorifics (*ibid.*, p.33).

Only two studies included in Shintani et al.’s (2013, supporting material) review were identified as involving beginners: Lee and Benati (2007) and Shintani and Ellis (2010). However, Lee and Benati described their subjects as "enrolled in a second semester intermediate course" (Lee and Benati 2007, p.107), and Shintani and Ellis's study looked at young learners of age 6-7 and who already had 4 months of prior study of English (Shintani and Ellis 2010, p.618); therefore the applicability of these findings to absolute beginners is not yet proven.

Research comparing processing instruction with *structured input* – i.e. input that has been designed to encourage learning without explicit teaching – has suggested that the role of explicit instruction in input-based learning may not be significant in developing either receptive or productive skills (VanPatten and Oikkenon 1996); it should be noted, however, that this finding is confounded by notable differences in the productive scores in pretest between groups. However, Ellis (2005) argues that implicit and explicit methods offer different and complementary benefits. He advocates for a “balanced learning curriculum that provides opportunities for meaning-focused output, form-focused learning, and fluency development” (Ellis 2005, p.34), but notes that the balance of activities will vary depending on the learners’ stage.

### Summary

All of the research into PBI and CBI has demonstrated acquisition resulting from pushed output, as per Swain (1985), contrary to Krashen’s (1982) input hypothesis, as well as acquisition from input. Disagreement revolves around which method results in better acquisition, and while there is no agreement on this, those who studied a mixture of CBI and PBI found it to be superior to CBI alone.

Additionally, while here is no general consensus on the value of grammatical explanations in language instruction, there remains ample evidence that acquisition can occur both with and without such explanations. However, all the research examined presented particular language points that subjects were expected to acquire.

## Computer-Assisted Language Learning

### The Tutorial Role of CALL in Controlled Practice

Levy (1995,1997; cited in Hubbard and Siskin 2004) proposed a dichotomy between CALL in *tutorial* and *tool* roles; with the key defining difference being that a tutorial CALL application would include some assessment or evaluation of student proficiency. Hubbard and Siskin (*op. cit.* p.456-457) consider that these tutor roles are “parallel functions” of language learning software, with most software performing both roles, and that Levy’s definition excludes functions that might otherwise be considered the role of a teacher, such as presentation of rules, simply because there is no evaluation carried out. Instead they suggest assessing the tutorial role through the “*teaching presence* of the designer” (*ibid.* p.457).

Jarvis and Szymczyk (2010) also discuss the notion of tutorial roles in CALL, and note that it is often associated with behaviourism and characterised as “drill and kill” (p.33); however, they note that this use of CALL replaces established equivalent practices on pen and paper, while improving on them by offering immediate feedback. Indeed, feedback is not the sole preserve of Behaviorism, and is central to ideas such as scaffolding and SCT-L2. As long as the feedback has what Ausubel termed as informational effect, such criticism is unfounded.

### CALL and the Internet as a Research Tool

Much has been written in psychology literature about the use of the internet in conducting research (e.g. Hewson 2003; Kraut et al. 2004), as well as literature on education (e.g. Cohen et al. 2013, pp.276-287,331-333). It is perhaps surprising, then, that there is less discussion within the CALL literature itself for this application of CALL techniques, with Beatty (2010, p.190) devoting less than a page to the topic of conducting research with CALL.

Reips (2002a) highlights dropout rates as a particular issue in online research, and high dropout rates are also a noted feature of free online courses (Hone and El Said 2016; Khalil and Ebner 2014). As the current research was carried out online in the form of a free course, steps were required to mitigate the potential effects of high drop-out. Reips (2002a) discusses several ways to do this: he presents figures demonstrating that asking for private information early on increases retention (*ibid*., p.242) and proposes a number of techniques specifically for the internet (*ibid*.), among which:

* *Warm up*: participants use the system for a period of time before the study begins – this will ensure that early dropouts occur before the true study,
* *High hurdle*: a large survey or other complex task at the beginning of the study will encourage earlier drop-out of participants likely not to complete the study.  
  (based on Reips 2002a, p.243)

## Summary

While there is much debate within the SLA community over the relative merits of productive and receptive skills practice in acquiring language, there is a broad consensus in the literature reviews that for acquisition to occur, learners must be engaged in activities which encourage and reward attention to meaning. The balance of opinion and evidence in the literature indicates that learners must be directed to notice form, although again opinions on this differ.

The notion of *scaffolding* accords well with Hubbard and Siskin’s idea of *teaching presence* in CALL, as the designer’s presence will be noticed and felt most clearly when it actively aids the learner in acquiring language through careful introduction and pacing of learning material.

Through the lens of SCT, Swain’s *pushed output* is a clear form of symbolic mediation: first the learner will attempt to self-regulate in order to produce a correct utterance, and if unsuccessful may be given the opportunity to use feedback from a communication partner or tutor.

CALL is the ideal tool for investigating CBI as it can be used to eliminate the confounding variable of spontaneous production observed by Shintani (2012) in her classroom study.

## Research questions, hypothesis and rationale

This research aims to investigate the relative effectiveness of reception-based and production-based practice on the uptake and retention on new language by ab initio learners of Gaelic using a web-based CALL platform.

Investigations centre on four research questions, and data is presented to answer each of these in Chapter 5.

The literature on this topic widely varying views on the effectiveness of receptive-based and productive-based interventions, but these theories have been tested largely with intermediate and advanced learners, and with a narrow range of language under examination. The following questions seek to determine whether the findings reported in the literature hold for ab initio learners learning grammar and vocabulary in an integrated manner.

RQ1: How does an early focus on CBI in CALL instruction affect learner performance in post-tests measuring receptive skills, compared to a PBI focus?

Much of the literature suggests that a receptive focus will lead to a higher performance in the task by the R group, but in this study, it is expected that this will not hold, as the number of linguistic variables to be mastered by beginners is higher than those in the restricted range of language points tested in previous studies.

Furthermore, it is expected that as subjects’ attention will be drawn to the lexical content of the sentences, a high number of errors will involve failure to correctly identify gender markers, and that this will be particularly pronounced among the R group.

RQ2: How does an early focus on PBI in CALL instruction affect learner performance in post-tests measuring productive skills, compared to a CBI focus?

It is expected that the group receiving PBI will achieve better scores on productive post-tests than the group receiving CBI.

It is predicted that due to the complexity of Gaelic spelling rules, the receptive group are likely to produce a high number of erroneous wordforms; however, this is not likely to be the only error encountered. Thus scores will be compared with spelling errors treated as mistakes and with spelling errors disregarded.

Further, it is expected that the R group will make a particularly high number of errors relating to noun-adjective agreement, as this aspect of form is never tested in the learning materials.

What effect does a mixture of CBI and PBI have on learner performance in post-tests measuring receptive skills compared to CBI or PBI alone?

Leeser (2008) and Izumi (2002) both found mixed input and output led to better receptive skills than input alone, and the hypothesis in the study is that this finding will hold true here and that a mixed approach will prove superior to CBI. However, the studies examined did not compare a mixed approach with PBI only; the hypothesis in this study is that a mixed approach will offer no advantage over PBI.

What effect does a mixture of CBI and PBI have on learner performance in post-tests measuring productive skills compared to CBI or PBI alone?

Mixed instruction is expected to result in better production during post-tests that CBI only, as per both Leeser (2008) and Izumi (2002); however, it is not expected that the mixed approach will prove superior to PBI in this regard, as the complexity of Gaelic orthography and the reduced volume of practice for the mixed groups is likely to result in lower scores.

This effect (if detected) may be reduced when spelling errors are discounted, and it is likely that the mixed treatments will perform at least as well as the productive group in terms of syntax, but it is likely that mixed treatment group will have greater difficulty with inflectional morphology in adjective agreements.

# Materials Design

## Introduction

This section describes the structure of the materials used in the study. (Instructions for accessing the materials can be found in Appendix K.)

First, a brief overview is provided of the technology underlying the materials, then a description is given of the question types constructed and how the technology is used to facilitate this.

## Software Platform

“Data collection techniques on the Internet can be polarized into *server-*side and *client-*side processing” (Reips 2002b, p246): client-side processing is carried out by software on the subjects’ devices, and server-side processing is carried out centrally, on a remote server. While server-side processing reduces the effect of technical differences between subjects’ own devices, client-side processing reduces the effect of network speeds on the study.

In early stages of the pilot for this study, it was noted that the server responded slowly, and client-side processing was therefore chosen in order to maintain subjects’ engagement with the material.

### Server

The website for the study is hosted on a Linux server located in the UK.

The virtual learning environment (VLE) *Moodle* is installed on the server and used to manage participants’ logins and access to individual sections of the learning materials and tests. It presents the webpages containing the tasks to the participants and logs their interactions with the system and collates basic reports. No data is changed or analysed on the server.

### Client

The client software in this study consists of the user’s browser and a Javascript web app executing within it.

The participant’s first interaction with the system is via the Moodle interface.

However, the software used for the teaching and testing itself was written by the researcher in HTML and Javascript and runs completely within the participant’s web browser. Verification of user responses and feedback is carried out, then results are communicated to the VLE via the SCORM 1.2 (see section 3.2.2.2) applications programming interface.

During piloting, the webpage was found to be compatible with all recent versions of *Firefox*, *Google Chrome* and *Safari*, as well as the *Edge* browser in Windows 10, but incompatible with all versions of *Internet Explorer.*

The client employs a number of technologies:

#### Javascript

Javascript is a programming language that executes code within a web browser. It was designed specifically to allow dynamic webpages that changed to user input, and for the validation of data in a browser before submitting to a server – for example, checking that valid credit card details have been entered before attempting to purchase goods (the-art-of-web 2017).

Javascript is used here to check user responses client-side and provide appropriate feedback to participants.

#### SCORM 1.2

The *shareable content object reference model* (SCORM) specification is a standard for the sharing learning data between various e-learning platforms that describes a common framework for reporting information such as the completion status of tasks and objectives, as well as learner responses to individual tasks (SCORM.com 2009).

#### HTML 5

*Hypertext markup language* (HTML) is the language used to define the content of the vast majority of webpages worldwide (Mozilla 2017). The latest version of HTML, HTML5, was designed to facilitate the integration of interactive elements with text, and has inbuilt support for multimedia elements (w3schools 2017).

The materials here presented make use of the standard’s audio file handling capabilities to provide a model of pronunciation to participants.

## Procedural Generation of Content

The material presented in this study was produced using a technique called procedural generation.

“Procedural generation is the process of creating content using an algorithm. This in itself has no element of randomness. If the functions, expressions, algorithms, and inputs that are used to generate the content remain the same, then you'll always get the same results” (Green 2016, ch.1).

Procedural generation offers a number of benefits to the CALL materials designer:

* Larger activity sets can be created
* Cost of production is reduced
* Variety can be increased
* Material can be revisited by learners more enjoyably

(adapted from Green 2016, ch.1)

Procedural generation was used to create materials that appeared to be ordered randomly to users; however, the order was the same for all participants within and across treatment groups.

## Learning Content

The following description of the design of the learning content has been split into three sections. The first describes how the visuals used in the study were designed to enhance the learning process. The second section describes the task types used in the study. The third section details how questions are arranged to encourage focus-on-form.

### Visual Design Principles

Within computer interaction design, there is a great deal of debate over whether abstract representations of objects are more or less effective than realistic representations in icons and interfaces. The main argument for realism is that it makes users more comfortable by “tap[ping] into people’s understanding of the physical world” whereas abstract interfaces are more efficient to use but can be “off-putting to the newcomer” (Preece et al. 2015, p.181).

The same disagreement is evident in language teaching resources, both electronic and in print, with many books using photographs to provide realistic imagery, and others using abstract images or cartoons. Many resources even use a mixture of the two styles.

The design for the present study presents material as abstract images and is informed by the principles of Gestalt psychology detailed below.

#### Gestalt and Visual Elements

Cohen et al. (2013, p.284) suggest that applying three principles of Gestalt psychology to page design increases retention rates in online surveys: *proximity* (physical grouping of items), *similarity* (psychological grouping by similar appearance) and *prägnanz* (simplicity, regularity and symmetry of graphical elements, resulting in quicker, easier perception). These principles have been applied here in the hope that they will have similar positive effects on users’ engagement with learning tasks.

**Proximity**: Multiple choice image options were presented in rectangular area that allows options to be quickly compared to each other.

**Similarity**: When presenting coloured items, images vary only in colour, so the noun is represented by the same form. The image representing the noun without an adjective is different (see Figure 3‑1), thus avoiding the situation where an uncoloured item is confused with a white item.

Using realistic images would have broken this principle, and differences in the images beyond the clothes’ colours may have drawn attention away from the target concept.

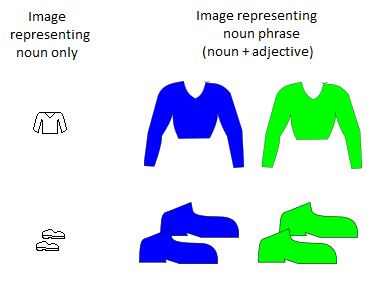


Figure ‑: Comparison of images representing noun phrases and bare nouns

Prägnanz: Images for module 1 were almost completely symmetrical. Flatshading simplifies the image, foregrounding colours as meaningful and making them less ambiguous (for example, a graded red might have been mistaken for pink).

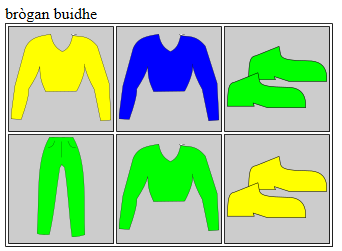


Figure ‑: An example of a multiple choice question

Note in the above example (Figure 3‑2) how the arrangement of buttons ensures *proximity*; how the repeated visual elements provide *similarity*; and how the flat colouring, the near symmetry of the trousers and jumper, and the two identical shoes in each pair gives *prägnanz*.

### Task Presentation and Feedback

Two types of task are used in this study, one involving language production and the other involving reception. During the teaching units, these provide feedback to the subject, who reattempts the task until completing it successfully. The tasks are identical in form in the testing units, but subjects are only given one attempt at each and receive no feedback.

The material is presented in such a way as to promote referential associations (Paivio, 1971) between an image and the target language; native language translations are not provided.

During learning units, an audio file containing the correct answer is played after the correct answer is given. This is omitted in test units.

#### Productive Task

An example of an interaction with the productive task type is shown in Figure 3‑3. This shows a learner making an error, receiving correction and subsequently entering the correct answer.

The learner is presented with an image to describe by typing a word, phrase or sentence. Gaelic features long vowels, marked with a grave accent (in the module, this occurs in the word *brògan* only); as many participants will not have a reliable method of entering diacritics, on-screen buttons are provided that insert accented vowels into the text area.

After entering an answer, the participant clicks “Check” and receives feedback. In test units, this button is instead marked “Submit”. If the answer is incorrect, the correct answer is revealed to the participant, who must then retype it in the box in order to proceed.

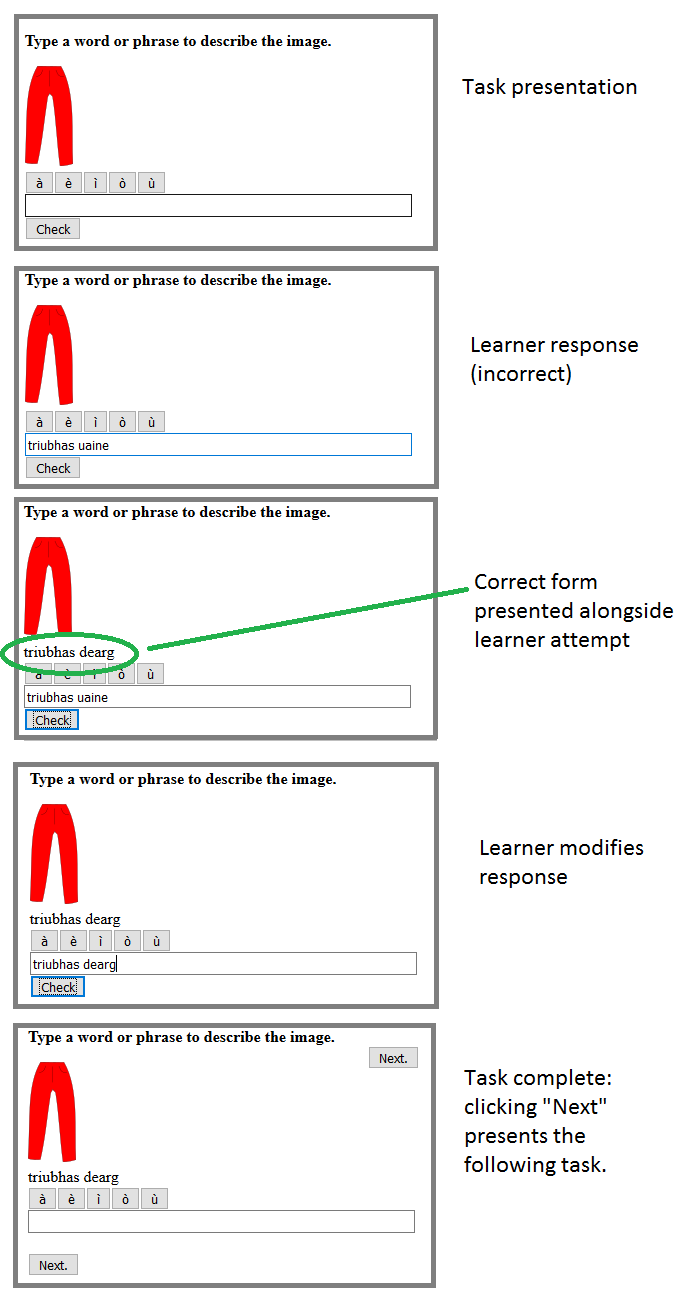


Figure ‑: Interaction sequence (including error in response) for productive question type.

#### Receptive Task

An example of an interaction with the receptive task type is shown in Figure 3‑4, which again includes a deliberate error.

The learner is presented with a word, phrase or sentence in Gaelic and a set of buttons showing pictures. When the learner clicks on an image during a learning unit, feedback is displayed either indicating a right or wrong answer, and if wrong, the participant reattempts the task until giving the correct answer.

In test units, only the first answer is accepted, and the subject proceeds directly to the next question.

In the first two units, the images unambiguously represent a single word or a noun phrase consisting of a noun and an adjective (as in the example) and feedback given on error includes a description of the image selected. However, in the third and fourth units of the module, all four colours and items of clothing are shown in all images, but the prompt only refers to one or two of them. Here, learners are only given feedback indicating a correct or incorrect response. More sophisticated error-specific feedback is possible here, but has not been used as equivalent feedback for the productive tasks was not possible in the time available for this study. This means that the informational content Ausubel et al. (1978) call for in feedback is not present, reducing later feedback to Behavioristic positive/negative reinforcement.

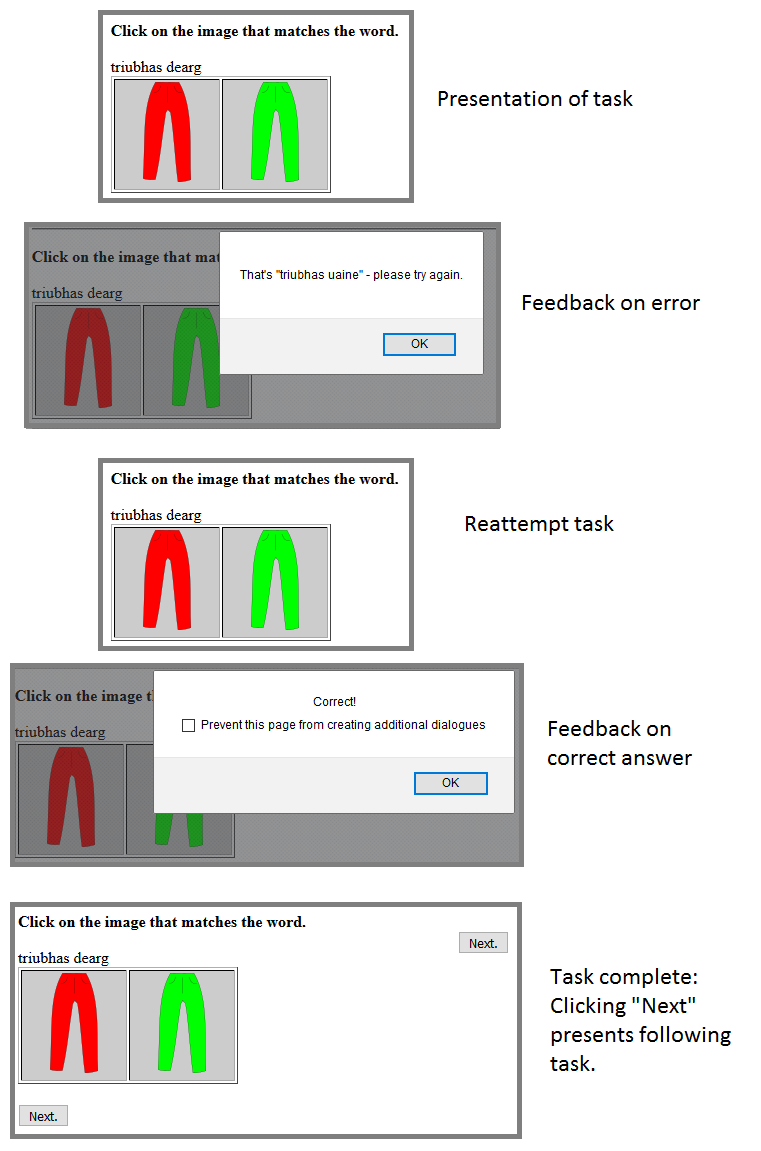


Figure ‑: Interaction sequence (including error in response) for receptive question type.

### Module Organisation

Each module comprises 4 learning units and 2 post-tests. Within each module, task complexity is gradually increased by the staged introduction of language, starting with individual words, combining nouns and adjectives into noun phrases, then finally inserting these noun phrases into pattern sentences. This scaffolds the learner, directing their attention to the language point to be acquired, thus encouraging focus-on-form.

Each unit within a module consists of a combination of presentation of new language points and practice of the presented material.

New vocabulary is presented in pairs, and then practice of both examples is given; another pair is then presented, and both examples are again practiced before all four are practiced together (as shown in Figure 3‑5). Where 8 items are presented, this sequence is repeated for items 5-8, after which items 1-8 are all practised together. Ordering of practice questions varies between units, but within in each unit is the same for all subjects and all treatment groups, to eliminate ordering effects as a confounding variable (Figure 3‑5 shows productive type tasks, but the sequence is identical with receptive tasks).

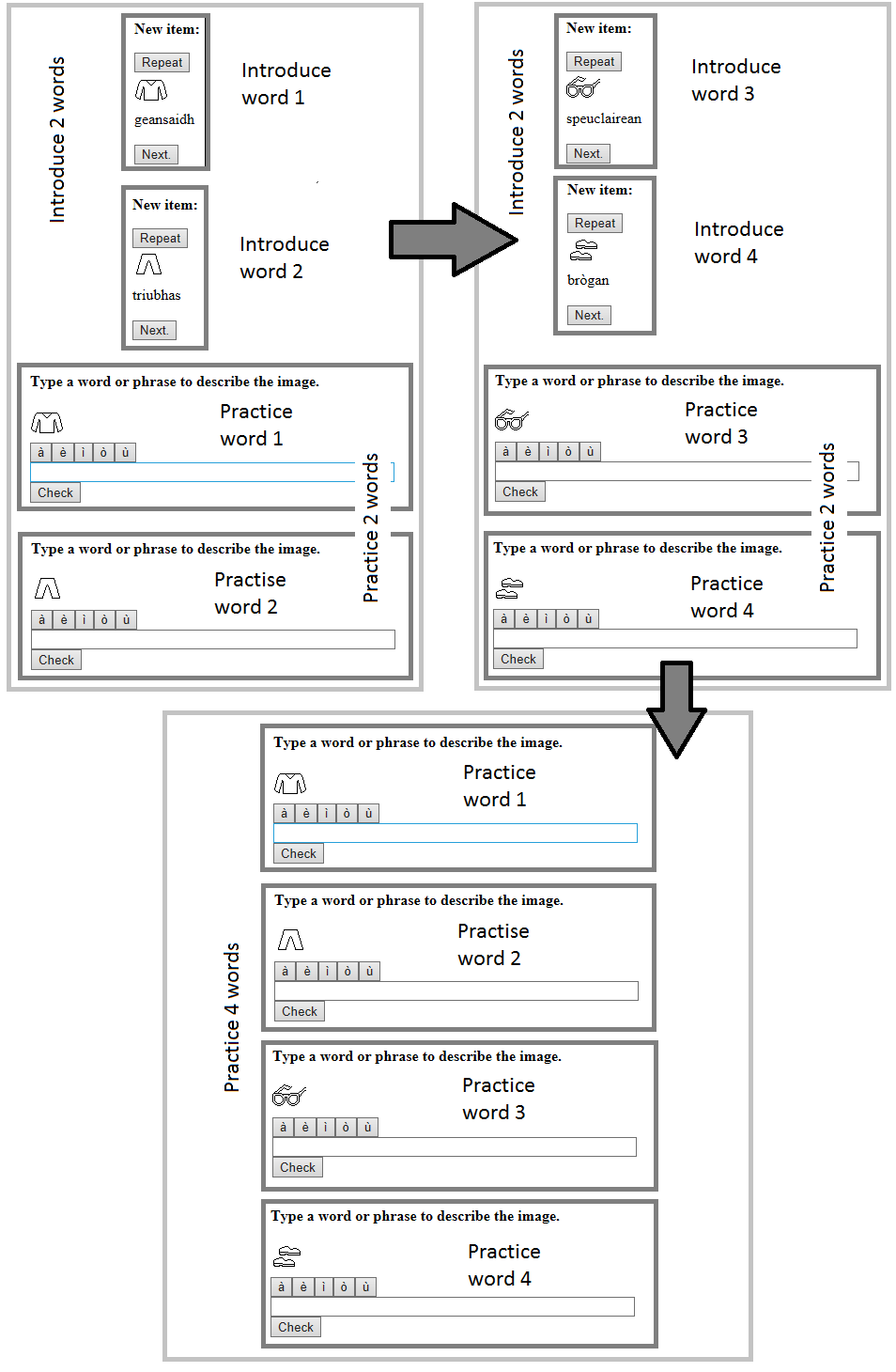


Figure ‑: Example of staged introduction and practice. Note that ordering here is consecutive, but in other sections this is not the case.

Grammatical concepts are presented similarly; however instead of each item presented being considered a discrete item, items are presented in pairs to contrast a part of the rule; contrasts are then highlighted between pairs until the full rule has been presented. Figure 3‑6 shows the items presented to demonstrate the rule for adjective inflection for number; practice follows the same pattern as described above, but has been omitted for brevity. Participants are then left to acquire the application of the declension pattern to the additional adjectives *gorm* and *buidhe* through a combination of intuition and feedback from the system.

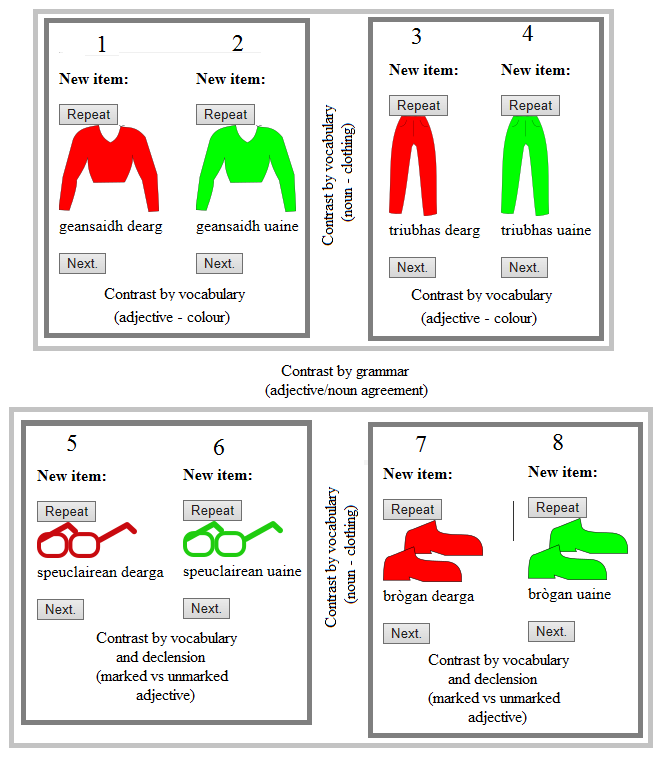


Figure ‑: Order of presentation of grammatical contrasts (module 1, unit 2)

Once all new language in a unit has been introduced, additional practice of all items is then given using the same style of question. For participants receiving a mixture of PBI and CBI, the tasks switch between exercise types after 50% of the tasks have been presented.

#### Additional Scaffolding in CBI

The receptive tasks presented for CBI are designed to be highly scaffolded initially, and gradually reduced scaffolding and increase task difficulty as each session proceeds. Figure 3‑7 shows 3 tasks from the final teaching unit of the model demonstrating the three ways scaffolding is reduced:

* early tasks provide arrows indicating the items under examination, but these are removed in later tasks;
* the number of options presented as answers increases;
* the number of differences between options reduces.

Participants are thus required to pay increasing attention to all elements of the question as the task progresses.

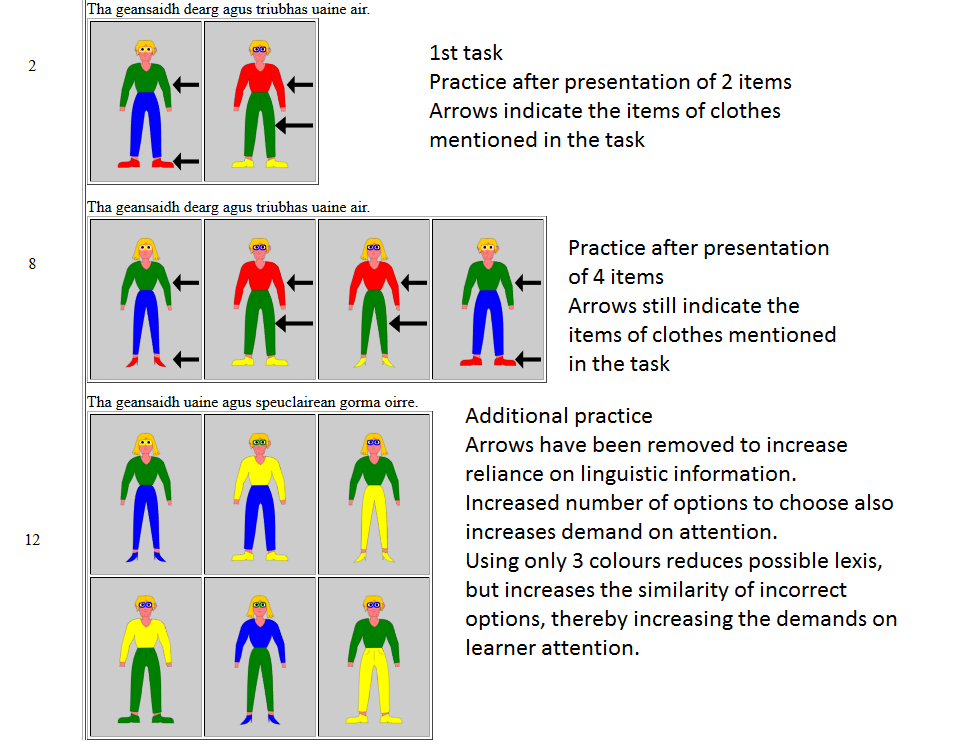


Figure ‑: Decreasing scaffolding within a unit of CBI

# Methods and Procedures

## Introduction

This chapter details the methods used in the present study. First, the overall design of the study is presented, including an overview of the target language items under examination. After this, the data-gathering phase is described, including how the participants interact with the system. This is followed by a description of how the data obtained was then analysed. After this, the concepts of validity and reliability are discussed; as high dropout rates were anticipated and present a threat to both validity and reliability, this section also details steps that were taken to mitigate the effects. Finally, the ethical considerations of this study are presented.

## Methods

This section details the methodology followed in the present study in terms of experimental design and the statistical analysis undertaken.

### Experimental Design

The study was undertaken as an *experiment* (Cohen et al. 2013, p.316), with subjects assigned to multiple treatment groups randomly to assure equivalence of groups; however, due to low completion rates, the final study samples showed notable demographic differences (see section 5.2), thus the study may be considered *quasi-experimental* (*ibid*., p.323). The design was a variant on what Cohen et al. (2013, p.319) term a *post-test two experimental groups design*, except with four experimental groups. Although a pre-test was administered, this is not a true pre-test in that it does not aim to indicate a change between the pretest and post-test (*ibid*.) and is of a different form from the post-test (*op. cit.,* p.493) but instead seeks only to verify that candidates are indeed beginners.

No control group has been used. Three factors make a control group impractical: first, as participants have no knowledge of Gaelic, a control group could confidently be predicted to receive zero marks in all productive tests, and a score no greater than chance (around 2) in the receptive test; secondly, the task types are taught during instruction, so the testing would not be equivalent; and finally, with no engagement in the course, a control group would not be subject to the same dropout as treatment groups, and therefore there would be no equivalence between control and treatment.

The post-test procedure is one of an immediate post-test approximately 24 hours after the learning is complete and a delayed post-test 7 days later.

### Descriptive Statistics

Quantitative data has been analysed using descriptive statistics. Opie (2004a, pp.132-133) states that this type of statistical analysis does not allow conclusions to be drawn that generalise over a wide population, but that it is appropriate to masters-level research given the restrictions on time available. As a consequence, results in this study show tendency that have no statistical significance, and are presented as hypotheses for further study.

#### Central Tendency

“The central tendency of a set of scores is the way in which they tend to cluster round the middle of a set of scores, or where the majority of scores are located.” (Cohen et al. 2013, p.626)

There are three typical measures of central tendency: mean, mode and median. The *mean* is often referred to the average and is calculated by summing figures for all members of a population, the dividing by the size of the population; *median* is the “midpoint score of a range of data” (*ibid*.) where there is an odd number of samples, and where there are an even number of samples it is the average of the two middle samples; *mode* refers to the most frequently occurring result (*ibid.*). The tables in Appendix E show these measures for the treatment groups in the study, alongside a list of individual scores, sorted numerically to highlight distribution patterns.

However, Opie (2004b, p.214) does not consider mode as describing a genuine central tendency, as clustering effects may result in modes that are far from central values; mode has therefore not been examined in this study.

#### Dispersal

The spread of data values around their central tendencies is termed *dispersal* and there are several measures of this: *standard deviation* (SD), *interquartile range* (IQR) and *range* (Cohen et al. 2013, pp.627-630). The SD is a measure of the distance from the mean value; the *interquartile range* (IQR) describes distribution around the median value, including the nearest quarter of datapoints above and below the median – thus it covers half of all data (*ibid.*). The *range* covers all datapoints in a sample, but *outliers* (datapoints that are sufficiently different from all others that they may be considered erroneous) are often excluded from the calculation of range (*ibid.*).

Coe (2002) discusses using a single “pooled” SD for all groups; however, he advises against doing this if samples cannot be guaranteed as representative of the same population, and as there were demographic differences between the participants completing this study (5.2), in this case individual SDs were produced for each sample group.

## Recruitment of Participants

Recruitment of participants was carried out via the internet, with adverts placed on a number of Facebook groups directing interested parties to the online consent survey (see Appendix B). Five pieces of information were requested: age, gender, first language, background as a learner of Gaelic (if any) and background studying other languages. Age, gender and first language were used primarily to ensure assigned treatment groups were roughly equivalent demographically. Respondents indicating a history of Gaelic study were excluded from the study, as were those with previous exposure to Irish and Welsh.

## Study Design

The material for this research was designed to be studied over a period of three weeks, with language presented in three modules. The first two only were designed as candidate material for analysis, as the material covered in the third week would not be subject to a post-test.

The study compared 4 treatments. Each group received the same presentation of target language examples as described in section 3.4.2 and were differentiated by type of practice: productive only (P); receptive only (R) (to address RQs1 and 2); productive-then-receptive (PR); and receptive-then-productive (RP) (to address RQs 3 and 4). In both PR and RP, exactly half of the learning tasks were productive and half receptive.

Subjects received different treatments each week, as shown in Table 4‑1 below. Note that due to limitations of time, no delayed post-test was performed for Module 3, which instead served two purposes: firstly, to maintain subjects’ engagement and ensure they returned to the site for the final delayed post-test; secondly, to ensure equivalence between the tests of modules 1 and 2, it was important that subjects studied unrelated language between the immediate and delayed post-tests.

|  |  |  |  |
| --- | --- | --- | --- |
| Study group | Treatment | | |
| Module 1 | Module 2 | Module 3 |
| a | P | R | RP |
| b | PR | RP | R |
| c | R | P | PR |
| d | RP | PR | P |

Table ‑: treatments by group and module

This design built redundancy into the study, as the data could then be used in several ways:

1. Ideally, both modules 1 and 2 would be used, allowing a comparison both between-subjects (within modules) and within-subjects (between modules), adding to the reliability of results and mitigating for the low statistical reliability of findings.
2. Data from module 2 could be used as the sole data for a between-subjects study. Module 1 would then be regarded as a warm-up as per Reips (2002a, p.243), allowing problems related to familiarisation with the material to be eliminated as a confounding variable in the study.
3. Finally, data from module 1 could be used as the sole data for a between-subjects study if insufficient participants completed module 2 to provide reliable results.

Crucially, the study was designed such that this decision could be made once the data gathering was complete.

During the study, drop-out rates were relatively high, and completion numbers for the second module were low. Furthermore, an unidentified technical bug resulted in a loss of data for the final two questions in the delayed post-test for several users. For these reasons, it was decided to use option 3 – data from module 1 only. In the rest of this chapter, the term “the module” will therefore be used to refer to module 1. However, some reference is made to module 2 in Chapter 5 (Results and Discussion) where it is relevant to subjects’ errors in the delayed post-test for module 1.

### Target Language Items

The target of the module was the comprehension and production of sentences such as in the following example:

Tha **geansaidh gorm** agus **brògan gorma air**.  
Be\_pres jumper blue\_S and shoes blue\_P on-him.  
He’s wearing a blue jumper and blue shoes/He’s got a blue jumper and blue shoes on.

A glossary and grammar of the language covered is available in Appendix A.

Although superficially simple, this sentence covers a number of complex linguistic variables:

* Noun-adjective word order.  
  The ordering of nouns and adjectives within a noun phrase is the opposite of English, with the adjective following the noun it qualifies.
* Monosyllabic Gaelic adjectives are marked for number, but polysyllabic adjectives are not.   
  e.g. gorm (s), gorm**a** (p); *but* buidhe (s), buidhe (p)
* Sentence word order.  
  Gaelic is a VSO language, so all sentences start with a verb, in this case “Tha”, the present indicative of “be”. This sentence is atypical in that it uses an idiomatic construction where what English would consider the object (the clothing) is treated as the subject, and its relation to the wearer is encoded in an adverbial at the end, making it appear as VOS to the beginner, which is markedly different from English’s SVO word-order.  
  Note that the language taught in the second module uses the basic VSO pattern, which, given the lack of explanation, presents a potential source of confusion for study participants.

This complexity was intended so that participants would be operating within their ZPD as per DeKeyser (1998), with the expectation that this would result in errors that would allow a more detailed analysis of the progress of participants’ acquisition of the language.

#### Module 2

As stated above, although module 2 is not analysed in this paper, the language presented is relevant in that it relates to certain errors made by a number of users in the post-test. Target sentences were of the form:

Tha **tè Ghearmailteach** a’ fuireach anns **a’ Ghearmailt**.  
Be\_pres woman German\_SingFem [progressive marker] living in [the] Germany.  
There’s a German woman living in Germany./A German woman is living in Germany.

As well as the use of the VSO order stated above, this introduces the progressive/continuous aspect in the present tense. This is a potential source of confusion as the semantics of the first module would commonly be expressed using the progressive present.

## Data Gathering and Participants’ Use of the System

### Pre-test

Respondents to the initial signup survey who indicated that they were absolute beginners were provided with a login to access the platform, where a brief pre-test was presented to verify that they were genuine beginners. As noted in Section 4.2.1, this was intended to confirm their status as beginners and thus not a true experimental pre-test. It consisted of general learner phrases, two intermediate sentences, and 7 colours.

Any candidate showing understanding of the intermediate sentences were excluded, as were any students successfully answering one question or partially answering any two questions.

Automatic marking for these questions was impossible, so all respondents self-identifying as beginners were initially allowed to proceed. Analysis of the pre-tests was carried out at the end of the course, and only on participants who completed both post-tests. This reduced the number of tests requiring manual marking from 99 to 31.

### Assigning Participants to Treatment Groups

Due to the high number of respondents, it was found that random assignment of subjects to groups resulted in a representative spread of age, gender and language background across treatments, so further sophistication in devising groups was not attempted.

### Study Content

Participants were given access to four lessons covering the language described in Appendix A, which they were asked to complete on consecutive days if possible.

The first lesson presented and practised the concrete lexical vocabulary individually (nouns: clothing; adjective: colour).

The second lesson presented and practised the combination of nouns and adjectives, demonstrating the noun-adjective word order and the pattern for declining adjectives in plural.

The third lesson presented and practised the sentence structure for saying he or she is wearing something, using the combined nouns and adjectives from the previous lesson.

Finally, the noun phrases in the above sentence structure were replaced with compound noun-phrases joined with *agus* (“and”) consisting of two items of clothing with colours.

### Testing Strategy

After participants completed the fourth lesson in the week, the post-test for the material was made available to them, and they were emailed and advised to complete the test within the next day or two. Participants were also requested to do no revision or study of the target language for the module in the following week.

In the next week, further material was presented covering different language points. This was intended to be the main material for the study; however, the completion rate for this second module was considered too low to merit analysis, with one group consisting of only 3 participants.

A week after completing the immediate post-test, participants were invited to complete a delayed post-test.

The post-test was split into three sections:

1. 8 productive tasks. Subjects should not have seen the target language earlier that day.
2. 16 receptive tasks.
3. A further 8 productive tasks.

Because these three sections covered the same language points, the difference between parts I and III provide a measure of partial acquisition – subjects who have partially acquired target language points are more likely to mediate and self-correct their output based on the receptive input provided in Part II.

## Analysis of the Data

Data was extracted from Moodle and analysed using ad hoc scripts programmed using Python.

As the data was derived from language test answers, the main analysis carried out was quantitative; however, test data can also be analysed qualitatively (Brown 2014, p.37), and this is done here to determine the nature of participants’ errors. This gives a form of *thick data*: data that describes “participants’ intentions, strategies and agency” (Cohen et al. 2013, p.540).

### The Python Programming Language

The Python programming language includes various sophisticated functions for handling scientific data and is ideally suited to the ad hoc analysis and visualisation of datasets of various forms (Ghent University 2017).

Python was chosen over software packages such as SPSS as Python can be run on a server as part of a website. It is intended that if this research proceeds to a full study, the tools used for analysis and marking will be integrated with the web platform, allowing reports to be run directly on the server.

Two code libraries that were used extensively were statistics, which offers functions to calculate the descriptive statistical measures discussed in section 4.2.2, and matplotlib, to produce the visualisations described in the next section.

#### Boxplots

Graphical representations were created using the Matplotlib software library for Python (Hunter 2017). These plots show the mean, median, IQR and range of the data along with any outliers. Matplotlib defines any datapoint lying outwith the IQR by more than 1.5 times the size of the IQR as an outlier and excludes it from the data range calculation (Matplotlib.org 2017). An annotated example is given below (Figure 4‑1).

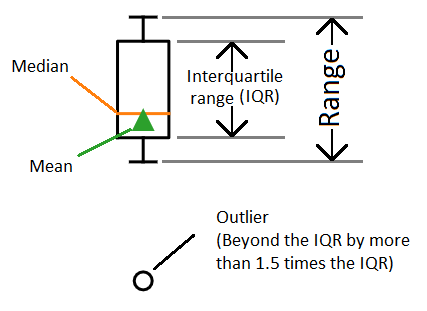


Figure ‑: Annotated example of boxplot generated using Matplotlib

### Receptive Post-tests

Correct and incorrect answers per user were counted and analysed in section 5.3.1.

To determine the nature of errors made in each of the immediate and delayed post-test, incorrect answers were identified and manually tagged into three categories:

**V**: error in lexical vocabulary (wrong clothes and/or colour)

**G**: error in functional/grammatical items (incorrect gender)

**V+G**: both types of error occurring in a single answer.

Counts of these categories were then compiled for every user. The data extracted can be seen in Appendix H.

A brief qualitative analysis of patterns found in these errors is presented in section 5.3.2.

### Productive post-tests

Errors in productive output were tagged with the following categories of errors:

**T**: minor typographical errors

**S**: spelling errors that didn’t obscure the intended word

**O**: other word used in place of target word

**U**: unrecognisable wordform  
(neither a clear spelling error nor recognisable as another word)

**E**: use of an English word in place of an unknown Gaelic form

**A**: incorrect agreement of noun and adjective

**#**: word order

**M**: missing word

**Z**: any other error

**X**: unclassified  
(This category was used as an intermediate designation to flag unclear cases for further investigation. No errors remained tagged as such by the end of the study.)

“Designing the coding scheme is a complex and important issue, but not one to be faced after data have been collected” (Gorard 2001, p.55). The initial set of categories devised in response to errors in the pilot study did not include the E category, as this error had not occurred with the initial pilot sample, which was added during data analysis when it was observed that several subjects had used English forms as placeholders when they could not remember the appropriate Gaelic word.

Typographical errors typically involved adjacent keys on a keyboard (e.g. truibhas→teiubhas; speuclairean→apeuclairean); if there was any doubt, other occurrences of the word were checked in the participants’ output. One user misspelt “brògan” as “brògam”, but as this occurred several times, this was classed as S. Additionally, several users typed “the” in place of “tha”. Personal experience has shown that this is a common error even for accomplished users of Gaelic, so this was classed as T after verifying that the users in question had successfully used the correct form in other parts of the test.

One user consistently wrote “brock” instead of “dearg”. As this is close to the singular form (“bròg”) of the word for shoes (“brògan”), this was classified as O.

Where an error in noun-adjective agreement was accompanied by another error, such as a spelling error or incorrect word, both errors were tagged.

From the tagged data, two scores were extracted for each user in each productive skills section of the test: in the first, minor typographical errors were eliminated and answers with only these minor errors were considered correct; in the second, all spelling errors were disregarded, to produce a measure of the grammatical and lexical errors isolated from orthographic issues. These were extracted as a list of results per group, presented alongside median and mean with standard deviation; the data is included here as Appendix E, and boxplots are included alongside an analysis of the descriptive statistics in section 5.4.1.

In order to look qualitatively at the nature of the errors, counts of the occurrences of different classes of errors were extracted by user, and by section of post-test, sorted by group, as presented in Appendix A. Figures cannot be directly compared between columns (see notes in Appendix A), only within them. This data is analysed in section 5.4.2, and visual overview is presented (Figure 5‑4) that indicates the presence or absence of at least one occurrence of the errors per user. The analysis of these error types is presented in section 5.4.2.

## Validity and Reliability

*Validity* “is concerned with the integrity of the conclusions that are generated from a piece of research” (Bryman 2016, p.41)

*Reliability* “indicates the extent to which our measurement instruments and procedures produce consistent results in a given population in different circumstances” (Dörnyei 2007, p50).

McDonough and McDonough (1997, p.64) assert that validity is more important than reliability where a choice must be made, as a repeatable experiment that does not validly replicate real-world settings provides few practically applicable findings.

### Validity

One manner of looking at validity is to consider the distinction between *internal* and *external* validity. Internal validity “seeks to demonstrate that the explanation of a particular event, issue or set of data which a piece of research provides can actually be sustained by the data” (Cohen et al. 2013, p.183) and “relates mainly to the issue of causality” (Bryman 2016, p.41) while external validity “refers to the degree to which the results can be generalized to the wider population, cases, settings, times or situations” (*ibid.,* p.186) and “is concerned with […] whether the results of a study are repeatable” (Bryman 2016, p.41).

A key factor offering internal validity is that the immediate and delayed post-tests are completely equivalent in their grammatical coverage and ordering, but these similarities are rendered opaque to participants by the alternating of pairs of vocabulary within grammatically-equivalent sets; thus, scores are almost entirely comparable, but participants gain no advantage in the delayed post-test from their memory of the immediate test (Scaife 2004, p.67).

“Ecological validity is the extent to which research findings would generalize to settings typical of everyday life” (Wegener and Blankenship 2007) and is a type of external validity. Reips (2002b, p.247) suggests that internet research such as the present study typically has high ecological validity as participants operate in familiar settings of their own choosing, eliminating any effects caused by the unfamiliar setting of laboratory-based experiments; although this has the consequence that “the researcher has no control over the experimental setting” (Cohen et al. 2013, p.332) which must be accounted for. In the present study, this ecological validity is enhanced as the setting is likely to be identical to the setting where self-learners would access similar CALL resources. However, this clearly reduces the applicability of the findings to classroom practice, where the environment is potentially more similar to a laboratory.

Triangulation refers to “the use of two or more methods of data collection in the study of some aspect of human behaviour” (Cohen et al. 2013, p195) and can increase the validity of findings. Although only one data collection method is used, this study employs a form of *methodological triangulation* by examining the data in two ways. First, correct and incorrect responses are analysed numerically to rate the success in learning of the subjects, after which individuals’ errors are analysed closely to try to identify underlying trends that identify evidence of partial acquisition and the problems underlying subjects’ errors.

### Reliability

Internet-based research poses particular challenges to the reliability of findings: as the exact circumstances and environment in which the participants engage with the study are beyond the researcher’s control, and often invisible to the researcher. To maintain reliability, all language items and concepts under investigation were tested multiple times in the post-tests; this meant that the impact of minor errors and slips on scores was diminished by averaging, and such slips could be identified by cross-referencing across questions.

### Mitigating Dropout Rates

Dropout rates in this study were expected to be high, potentially compromising the validity and reliability of findings, so several steps were taken to reduce the impact of dropouts.

* Canvassing for participants was targeted at a number of groups who were expected to have high motivation to complete the course.
* The course for this study was designed as independent modules such that data from either of the first two weeks could be used even for participants who did not complete the subsequent week(s). Alternatively, if findings for the first week were confounded by dropouts, these could be disregarded and used as a warm-up for the second week, as per Reips (2002a).
* Recruitment to the study was held open after the start date, with the aim that should any groups suffer high dropout rates, additional subjects could be placed in the study group to ensure sufficient data was available at the end of the study for each group. *(It should be noted that while certain groups did experience higher dropout rates than others, this occurred too late for any such corrections to take place.)*

## Ethical Considerations

Dörnyei (2007, pp.63-64) describes educational research as a form of social research relating to closely to people’s lives, and states that research ethics must be rigorous to protect subjects’ privacy and rights.

The current study was carried out with the approval of the general university ethics panel (GUEP) of the University of Stirling, and conforming to their standards. All participants were informed of the purpose and nature of the study at the outset, and participants were informed of their right to withdraw from the study at any time. Participants were also informed that information would be used anonymously, and accounts were set up using pseudonyms.

All data was stored on servers based in the UK and subject to UK data protection legislation. Servers were protected with industry standard HTTPS secure logins.

## Conclusion

This chapter has described the methods employed in the present study, how data was gathered and analysed, and how this affects the validity and reliability of the conclusions drawn, as well as ethical considerations taken in the study.

# Results and Discussion

## Introduction

This chapter presents the findings of the research. First, information about the final study sample is presented, followed by analysis and discussion of subjects’ performance in post-tests. The chapter closes with a summary of the study’s findings.

The discussion of the post-test comprises two sections, detailing receptive skills and productive skills, each of which addresses two research questions. These sections are subdivided into a quantitative analysis of participants’ test scores followed by a qualitative analysis of the errors arising in participants’ responses, after which the findings are discussed in respect to the research questions relevant to the section.

## Description of the Study Sample

Due to dropout rates, as well as the elimination of several users (see section 4.5.1), the final samples for the groups completing module 1 (including both post-tests) were smaller than anticipated, as shown in Table 5‑1 below.

|  |  |  |
| --- | --- | --- |
| Group | Treatment | Sample size |
| a | P | 7 |
| b | PR | 7 |
| c | R | 6 |
| d | RP | 11 |

Table ‑: Group sizes and treatments (module 1)

Although the groups initially had a representative distribution of demographics, this balance was lost as participant numbers reduced. Table 5‑2 shows the spread of age ranges across groups, as well as average values, and it can be clearly seen that participants in group A, receiving treatment P, are typically younger than in any other group (mean age 30, median 23), and that participants in group C, receiving treatment R, are notably older than any other group (mean age 47, median 50.5).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Groups | a:P | b:PR | c:R | d:PR |
| Ages | 22 | 20 | 28 | 25 |
|  | 23 | 27 | 40 | 26 |
|  | 23 | 31 | 47 | 28 |
|  | 23 | 32 | 54 | 37 |
|  | 29 | 44 | 55 | 39 |
|  | 40 | 54 | 59 | 40 |
|  | 50 | 60 |  | 45 |
|  |  |  |  | 46 |
|  |  |  |  | 53 |
|  |  |  |  | 54 |
|  |  |  |  | 58 |
| Mean | 30 | 38 | 47 | 41 |
| Median | 23 | 32 | 50.5 | 40 |

Table ‑: distribution of age ranges across treatment groups in final sample.

Furthermore, groups differed significantly in their levels of prior language learning experience. All participants in groups A, B and D reported some experience of learning, while 2 in group C reported no prior experience. Group A included two childhood bilinguals, and none were present in any other group. One subject in group C was a non-native speaker of English, as well as 5 of the 11 in group D.

Gender-balance also differed significantly between groups, as shown in Table 5‑3. Group A consisted solely of females, and groups B and C were predominantly female, while group D had an equal number of males and females (disregarding undeclared gender).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Group | a:P | b:PR | c:R | d:RP |
| Male participants | 0 | 1 | 1 | 5 |
| Female participants | 7 | 6 | 5 | 5 |
| Gender undeclared | 0 | 0 | 0 | 1 |

Table ‑: gender distribution across treatment groups

Participants’ survey responses are included in Appendix C.

## Receptive Skills:

The present study seeks to investigate receptive skills by answering the following questions:

RQ1: How does an early focus on CBI in CALL instruction affect learner performance in post-tests measuring receptive skills, compared to a PBI focus?

RQ2: What effect does a mixture of CBI and PBI have on learner performance in post-tests measuring receptive skills compared to CBI or PBI alone?

Part II of the post-tests, described previously in section 4.6.2, was used to address these questions.

### Analysis of Scores

Part II of both post-tests, discussed here, is scored from 0-16.

The scores obtained by each treatment group during the receptive sections (Part II) of the immediate and delayed post-tests is summarised visually in Figure 5‑1, which shows that the differences between groups were minimal.

|  |
| --- |
| Figure ‑: Post-test scores for receptive skills test (immediate and delayed). Summary of figures can be found in Table E‑5 and Table E‑6 in Appendix E |

The mean scores for both the P and R treatment groups were 15.14 (± 1.215) and 14.17 (±0.9833) in the immediate post-test and 16 (± 0) and 14.5 (± 1.517) in the delayed post-test respectively, meaning that the P group exhibited marginally better performance. It is worth noting that in the immediate post-test, over half of the P group obtained full marks, whereas all of the R group made at least one error and half made at least two errors (see Table E‑5); more notable still is that in the delayed post-test, the P treatment group all received full marks, while only 2/6 in the R treatment group did (see Table E‑6 in Appendix D).

The PR and RP groups performed very similarly to each other in both receptive tests, obtaining mean scores of 15.29 (±0.756) and 15.45(±0.934) respectively in the immediate post-test, and 15 (±1.414) and 15.45 (±0.82). RP and PR were therefore very similar to each other, and although they were very close to the P treatment in the immediate post-test, they performed marginally worse in the delayed test.

A final observation is that no members of the P treatment group showed any decline in performance in receptive skills between the immediate and delayed post-test, with all scoring equally or higher, whereas all other treatments saw 2 or 3 subjects scoring lower in the delayed test than the immediate (figures are included in Appendix F).

### Analysis of Errors

The figures presented above show that three members of the P treatment group committed errors in the immediate receptive post-test. As noted in Chapter 4, subjects in the P treatment received no practice in the receptive task prior to this test, and to verify whether this affected their scores, the distribution of incorrect errors was examined. Subject a2f gave incorrect answers to the 7th and 11th questions of this section and subject a7f provided an incorrect answer to the 3rd question. One subject, a6f, committed an error in the first question, but also committed errors in the 14th and 16th (final) task. This suggests that unfamiliarity with task type has had little effect, if any, on the results for the P treatment group.

In looking at specific errors by users in the post-tests (see Appendix H), it appears as though errors in lexical items were on the whole more common than failure to recognise gender-marked items. In the immediate post-test, within the R group 8 incorrect answers involved errors on lexical vocabulary, but only 5 involved gender identifiers. RP showed 5 errors in the immediate including lexical mistakes, and only 3 including gender marking mistakes. In the delayed post-test, all of PR’s errors involved lexical errors, and only 1 involved a problem with gender marking; 8 of R’s errors were in lexical vocabulary and 1 in gender marking; and the RP group showed an even split of 3 of each.

### Discussion: RQ1

The study finds no evidence to support the applicability to beginners of Shintani et al.’s (2013) finding that receptive instruction results in a better performance in receptive in post-tests; indeed, the data supports Nagata’s (1998) finding that productive practice leads to better gains in receptive skills than receptive practice.

Notably, the P group were the only one whose performance on this task improved over the period between the immediate and delayed post-test. However, it should be remembered that the group receiving the P treatment in module 1 were subjected to the R treatment in module 2, so practiced this task type more frequently than other groups in the week between the immediate and delayed post-test, which may account for this effect.

The initial hypothesis that receptively-taught subjects would pay more attention to lexical vocabulary that the productive group appears is also unproven; indeed, the R group made fewer errors of gender than of lexical vocabulary. The numbers are too small to lead to any solid conclusions, the fact that this pattern is repeated in the RP group suggests that this finding bears further investigation.

It seems likely that the large amount of redundancy in the language presented allowed participants to complete receptive tasks with little need to attend to form, thus not creating a strong connection between form and meaning (Swain 1985) or analysing the language as a system as per Naiman et al. (1978).

### Discussion: RQ3

The study finds no evidence to suggest that a mixed approach is of any additional benefit to learners in a teaching intervention of this sort. However, if the trends shown were replicated on a larger-scale study, this would indicate that the mixed approach is superior to the receptive-only approach, as per Izumi (2002) and Leeser (2008), but marginally inferior to the productive-only approach.

## Productive Skills

The present study seeks to investigate productive skills by answering the following questions:

RQ2: How does an early focus on PBI in CALL instruction affect learner performance in post-tests measuring productive skills, compared to a CBI focus?

RQ4: What effect does a mixture of CBI and PBI have on learner performance in post-tests measuring productive skills compared to CBI or PBI alone?

Parts I and III of the post-tests, described in section 4.6.3, was designed to address these questions. Receptive skills were tested in Part II; a comparison of Parts I and III of each test therefore allows observation of partial acquisition, as participants are able to use the input from the receptive tasks to assist in mediating their output.

### Analysis of Scores

Parts I and III of the post-tests, discussed here, are each scored from 0-8.

When scores were initially analysed, as shown in Figure 5‑2, it became immediately clear that the R treatment group performed worse than the P group throughout. In Part I of the immediate post-test, P obtained a mean of 6±2.89 and R obtained 0; in Part III, the mean for P was 7.43±0.79 and for R 1.33±3.26; in the delayed post-test P achieved 3.57±3.75 and R achieved 0.17±0.41 in Part I; and P achieved 7.14±2.27 and R achieved 1.83±3.12 in Part III.

While the P group showed a marked ability to mediate their output in the second set of productive tasks in both the immediate and delayed post-tests, the R group showed little improvement in performance. Several showed slightly improved scores on Part III of the delayed post-test.

The PR and RP groups showed performance comparable to the P treatment on both Parts I and III of the immediate post-test, but with a broader spread of results. Both groups showed a marked improvement in Part III, indicating an ability to mediate their output based on the input presented in Part II of the test. P and R treatment groups scored higher on average in Part III of the delayed post-test than in the immediate post-test, suggesting increase acquisition, while PR and RP scored lower, indicating decline.

In Part III of the delayed post-test, the P treatment group had a clear advantage over all other groups.

|  |
| --- |
|  |

Figure ‑: Comparison of immediate and delayed post-tests of productive skills for P and R treatments, covering both Part I (before the receptive test) and Part III (after the receptive test). (Full figures in Appendix E.)

Discounting spelling errors notably improved the results for both P and R treatment groups across all tests, as can be seen by comparing the results above to those in Figure 5‑3. The marked change in results between Part I and Part III shows that members of these groups were able to successfully employ intake from the receptive tasks to mediate for all types of errors.

Also of note were the results of Part I in the immediate post-test: as 4/6 of the R group got at least one answer correct where spelling was corrected for, it is clear immediately that most understood what was asked of them by the task despite having no previous exposure to the specific task type. Thus while this may have resulted in lower performance, it does not necessarily render the findings invalid.

With respect to the PR and RP treatment groups, while scores in Part I of the immediate post-test were notably higher than before spelling errors were discounted, their respective Part III scores were similar before and after accounting for spelling errors, suggesting that participants’ attention was focused on spelling errors in Part II, diverting attention from other types of error.

Both PR and RP scored consistently higher the R group. Their scores were comparable to the P group initially, but P showed better relative performance progressively as the study advanced.

|  |
| --- |
| Figure ‑: Comparison of post-test scores with spelling errors disregarded Summary figures can be found in Table E‑3, Table E‑4, Table E‑7 and Table E‑8 in Appendix E |

### Analysis of Errors

As described in section 4.6.3, errors were categorised into 8 types. Figure 5‑4 shows a visual indication of the occurrence of these types of errors in the output of participants in the two productive sections of both the immediate and delayed post-tests. (Full details including error counts are included in Appendix A.)

The image shows at a glance that the P treatment committed a narrower range of error types in all test sections than any other treatment; that within this range, fewer individuals committed these errors; and that members of the R treatment group committed a wider range of error types than any of the other groups.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group | | Immediate | | | | | | | | | | | | | | | | |  |  | Delayed | | | | | | | | | | | | | | | | |
| Part I | | | | | | | |  | Part III | | | | | | | |  | Part I | | | | | | | |  | Part III | | | | | | | |
| s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z |  | s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z |
| a:P | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  |  |  |
| b:PR | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  |  |  |
| c:R | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  |  |  |
| d:RP | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Figure ‑: A visual mapping of productive error types by group.

Key:

s: spelling

u: unrecognisable wordform

o: other word

e: English word

m: missing

#: incorrect order

a: noun/adjective agreement

z: all other errors

(Full details: section 4.6.2)

Red: at least one error of this type

Green: no error of this type.

Each row represents a single user.

Notably, none of the P group made any errors on adjective agreement in either section of the delayed post-test. By contrast, all but one of the R treatment made at least one error of agreement in part I and again in part II. Fewer in each of PR and RP made errors with agreement than the R group, but more than P.

Participants in the P treatment only made errors of spelling and omission of words in Part I of the delayed post-test, suggesting that they had forgotten only vocabulary items in the intervening week, and the results for Part III show that even these errors were successfully mediated for, suggesting that the participants in the P group were close to full acquisition of all target language items. Furthermore, all participants in the P group appear to have fully acquired the rule for declension of adjective for number, with few errors in the immediate post-test and none in the delayed post-test; none of the other groups managed to use this rule with 100% accuracy in any of the post-tests.

Examination of the answers supplied by members of the R group show that 4/6 made clear attempts to produce sentences in the first section of the immediate post-test, while 2/6 did not. In the final section of the delayed post-test, only 1/6 did not attempt to produce complete sentences. It should be assumed that this user did not understand the intended output from the task. Correcting for this gives a mean score of 2.2±3.35 (or 3.8±3.03 with spelling errors excluded), which remains notably inferior to the average for the P treatment and the mixed treatments discussed in the previous section.

The ordering of nouns and adjectives should be easily determinable and noticeable from input. In the Part I of the immediate post-test, only two participants made errors in this ordering – b2f made a single error (treatment PR), but got the order correct on every other attempt, whereas c5f (treatment R) produced the incorrect order on every attempt she made. C5f then successfully mediated her output in Part III based on the input provided in Part II; no participants made any errors in Part III. Only two errors of this type occurred in Part I of the delayed post-test, from c3f and again from c5f, but in both cases this was a single isolated incident accompanied by several correct attempts, suggesting that the rule was partially acquired, and only a small degree of mediation was required in order to avoid inadvertent slips into the English order.

The errors classed as “other” took various forms; a summary of all such errors is including in section I.2 of Appendix I. Of these errors, one type is of particular interest: in the delayed post-test, several participants made errors showing signs of interference from the material studied between the immediate and delayed post-tests.

Three participants (b2f, d1f and d2m) all consistently attempted to apply the pattern “tha fear/tè …” from module 2 in place of the pattern “tha … air/oirre” from module 1. For example:

Target form: Tha speuclairean dearga agus brògan buidhe air.  
 Lit: *Is spectacles red\_p and shoes yellow on-him.*Given form (b2f,q2): Tha fear speuclairean dearg agus brògan buidhe.  
 Lit: *Is man spectacles red\_s and shoes yellow.*

Subject c4f made this type of error twice, and in other responses made no attempt to mark for person, indicating that she was unsure; subject d1f made this type of error in her first answer, but then applied the correct pattern in all subsequent questions, showing that she had not fully acquired the pattern, but was able to self-regulate in the absence of any input. All of these participants were successfully able to use the input in Part II to mediate their output, avoiding replicating the error; although subject c4f was only able to mediate partially – in Part III her answers used the correct vocabulary (air/oirre), but she continued to use the word order taught in module 2.

A particularly interesting example is subject d2m, who offered several responses of the following form:

Expected response: Tha brògan buidhe agus speuclairean dearga air.  
Given form (q2): Tha fear an ? speclairean dearg agus brogan buidhe

The spacing here implies that the participant believed a word was missing. It may well be that he was trying to recreate the present progressive construction from module 2 (e.g. *Tha fear a’ fuireach… -* a man is living…) and is indicating uncertainty about the Gaelic for “wearing”, even thought this was not part of the idiomatic expression taught.

The interference errors described above all show attempts by participants to engage with meaning, but there are errors also involving interference of semantically and grammatically unrelated forms, such as b4f’s confusion of *agus* (“and”, module 1) with *anns an/anns a’* (“in the”, module 2) to give *anna/ana*; and c1m’s repeated use of *an* (“the”, module 2) in place of tha (“is”, both modules), which is particularly noteworthy as he failed to notice this in the input in Part II and consequently consistently committed the same error in all responses in Part III.

It is noteworthy that no errors of interference were committed by participants in the P treatment group.

### Discussion: RQ2

The data presented quite clearly shows that the P treatment group produced more correct language that the R group, just as predicted by Shintani, Li and Ellis (2013). Close examination of individual users’ errors suggested that participants in the R treatment may have suffered initially due to the unfamiliarity of the productive task; however, discounting the subject from the R group who showed no attempt to produce the whole sentence still left the P group with a clear advantage.

Analysis of error types shows that the P group had acquired more of the target language items, and that while the R group were generally able to mediate their production with feedback, their grasp of the rules still required much more conscious mediation. In particular, P showed much stronger command of the syntactic and inflectional rules presented than the R group. This evidence of mediation shows that participants were monitoring performance, as per reported strategies in Naiman et al.’s (1987) survey.

Although the group receiving treatment P committed no errors due to interference from the second module, these results are difficult to draw conclusions from. In the second module, this group received treatment R, and as this treatment has been demonstrated here to be weaker, it is not possible to determine whether the lack of interference is due to superior learning of module 1’s language or inferior learning of module 2’s language. While this still implies the superiority of the P treatment over R, without further investigation into how the two treatments interact, this remains unconfirmed.

### Discussion: RQ4

The data presented above indicates that both PR and RP gained a notable advantage over the R group in the production of language, echoing Izumi (2002) and Leeser’s (2008) findings that a combination of input and output practice resulted in better output than input alone. However, they showed no advantage over the P group, instead obtaining marginally lower scores on average.

The differences that were found were much more pronounced where spelling errors were included, initially suggesting that the main advantage of a mixed approach over a receptive-only approach is in terms of acquiring accurate spelling, but closer analysis of errors indicated that participants in the R group committed a much wider range of errors than either the PR or RP treatment groups. This suggests that the PR and RP groups may be closer to acquiring the target forms that the members of the R group. Comparing the range of errors of the mixed groups to that of the P group also suggests that the difference in performance is greater than the statistical analysis indicates.

The data therefore indicates that a mixed approach is inferior to PBI, but superior to CBI, so finds no grounds to recommend a mixed approach when learning with materials of this type.

## Summary

The data presentation and analysis above shows that in this study, participants receiving productive-based instruction outperformed those receiving receptive-based instruction in all measures, and that participants receiving a mixture of CBI and PBI outperformed those receiving CBI only, but underperformed those receiving PBI only, suggesting that PBI is the most appropriate treatment for absolute beginners studying with this type of material.

# Conclusion

This section first summarises the findings of the previous chapter, then presents a brief description of the implications this has for teachers and professional practitioners. The limitations of the present study are then presented, followed by some recommendations for how a full-scale study might be developed to further investigate and confirm or refute the findings of this pilot.

## Summary of Findings

The data examined above suggests that at the earliest stages of language learning, a productive-based method of instruction provides better acquisition of all elements of language than comprehension-based instruction, leading to better performance in both productive and receptive post-tests.

The findings are presented here in terms of the research questions they correspond to.

### RQ1

With regards to the claim that CBI results in improved performance in receptive post-tests, this study found no evidence to suggest that this holds true for ab initio learners. Indeed, the data tended to suggest that PBI-instructed participants performed marginally better, and that CBI-instructed participants had some difficulty distinguishing between lexical items.

### RQ2

The study found strong evidence to support the hypothesis that participants taught using PBI performed better on productive tasks in post-tests, indicating a better acquisition of productive skills. Furthermore, participants receiving productive treatment showed signs of an ability to mediate for almost all errors after seeing similar examples of the target language.

### RQs 3 and 4

Trends suggested that these mixed methods were superior to CBI alone and inferior to PBI alone.

## Pedagogical Implications

The findings of the study, though tentative have quite notable implications in teaching, particularly in the field of internet language learning.

The difficulty CBI-instructed participants had in acquiring the language points under examination call into question the usefulness of receptive skills practice at an early stage, as the variety of language that can be presented to learners at this level means it is difficult to design receptive tasks that involve genuine focus-on-form. This is particularly true in a CALL setting: a face-to-face teacher can manipulate familiar elements of their physical environment to add meaningful variety to their lesson, whereas all materials within a CALL environment must either be formally introduced or conform directly to the norms and conventions of learners’ online experiences.

## Limitations and Delimitations

The present study was delimited by the time available, and as such, no measure has been taken of the long-term effects of the intervention; it may be that a longitudinal study would find that the initial advantages shown by the P group were negated or indeed overturned by the other treatment groups after an extended period study or an extended period of rest.

The size of samples was also a limiting factor, meaning no conclusions could be drawn with statistical significance. In particular, the difference between PBI and the mixed treatments were typically marginal, and a larger sample size would be required to determine if this described a genuine effect, or was caused by the sample size in the present study.

It could be argued that the CBI intervention did not necessitate meaningful attention to all elements of the target sentence, but this is perhaps the biggest weakness of CBI for absolute beginners – some features of language are by nature redundant, and the reduced depth and breadth of language presented to the ab initio learner means increases that redundancy; thus it is only in pushed output, much like Swain’s (1985) Output Hypothesis suggests, that gives all components of the language the required salience to be acquired.

Furthermore, a decision was made to make treatment equivalent in terms of the questions asked, whereas the majority of studies attempt to make treatments equivalent instead by time-on-task, even though this results in a different number of questions being presented. No measure of time-on-task was available for any group, and it is strongly suspected that had this data been available, it would show that the P treatment group would be shown to have spent longest during learning exercises, followed by the PR and RP groups, with the R group spending least time. Thus the observed differences in post-test performance between the treatment groups may be explainable as a result of time-on-task, and not as a direct result of the type of practice received.

Several additional confounding variables make drawing firm conclusions from the present data difficult. First, there is a notable difference in demographics between groups, and there is an apparent negative correlation between average age and performance; furthermore, the language background of groups differed greatly, and the group with highest performance was the only one containing child-hood bilinguals, while the one with lowest performance was the only one where any participants had no prior experience of language learning whatsoever.

While there is evidence of all groups using mediation strategies to use intake provided in receptive tasks to improve their own production, the present study cannot determine whether the most successful learners had fully acquired the language structures presented, or merely memorised them by a behavioristic rote method.

A major confounding influence on the results is the presence of interference from the second module of the course in the post-tests for all treatment groups other than P. As each group received a different treatment in the second week, it is impossible to disentangle the effects of weaknesses in acquisition of the material in module 1 from those of the strength of acquisition of the material from module 2. However, it should be noted that as the groups receiving R (the weakest treatment) and P treatment (the strongest) swapped, the two effects would appear to be mutually reinforcing rather than counteracting each other. This interference the important issue of *meaningfulness*, which Morgan-Short and Bowden (2006) consider particularly important when comparing the effects of comprehension and production in acquisition. With the participants whose post-test performance was affected by interference we have evidence of deep engagement with meaning; however, as no-one in the P group made any such error, there is insufficient evidence to be sure that their improved performance of the test was due to a better internalisation of the target structures, rather than behavioristic memorisation.

However, the study provides a firm basis on which to build a future large scale study, and has provided valuable insight into how to organise this, and the instruments required for measuring this, as described in the next section.

## Future Research

Several of the trends identified in this study show potential for future investigation, and it is the author’s hope to extend this into a PhD-level study. Such a study would aim to repeat the findings with another cohort of absolute beginners, potentially in several languages, and would include a long-term component to verify whether the effects seen at the early stage result in lasting effects on acquisition.

With improved automatic marking of participant responses, the error classification developed in the qualitative analysis of the present study could be reimplemented as a quantitative measure to provide a deeper analysis of which types of language data are most easily and quickly acquired. Integrating this into the automatic marking system will also enable more sophisticated informational feedback to scaffold learning and promote noticing.

Doing a similar study on a longer timescale would also allow an examination of whether language learned had been memorised by rote or acquisition of the language system; later modules examining the same language rules in different contexts or in combination with different language points are likely to be made measurably easier if the system has been acquired.

Prior to commencing this study, little thought had been given to interference between structures in the target language. In this study, this was caused by language being presented between taking the immediate and delayed post-tests. This interference appears to highlight gaps in partial acquisition by making it harder for the subject to self-regulate, and increases reliance on additional input for mediation. To improve the validity of measures of interference in a future study, any language teaching or presentation given between the immediate and delayed post-tests would be identical for all treatment groups; thus any differences in the interference experienced between groups could be correlated unambiguously with the variable under study.

Such a study should also take into consideration previous research findings suggesting neither PBI nor CBI is superior for intermediate learners and seek to identify if there is a point in the learning process where PBI ceases to show clear signs of superiority over CBI.

## Conclusion

Within the material in the present study, good evidence was found for the superiority of the PBI approach for absolute beginners, and data further indicated the superiority of pure PBI over a blend of PBI and CBI.

However, the present study was of too small a scale, both in terms of participant numbers and the length of the study, to draw firm conclusions, and it is hoped that this research can be extended into a longer-term study.

References

Atkinson, D. (2011) Introduction. In: D. Atkinson ed. *Alternative approaches to second language acquisition.* Abingdon: Routledge, pp.1-23.

Ausubel, D.P., Novak, J.D. and Hanesian, H. (1978) *Educational psychology: a cognitive view, 2nd ed.* New York: Holt, Rinehart and Winston.

Babbel.com (2017) Learn Spanish, French or other languages online – Babbel. Available at: <http://uk.babbel.com> [Accessed: 2nd August 2017]

Beatty, K (2010) *Teaching and Researching Computer-Assisted Language Learning*, *2nd ed*. Harlow: Pearson Education.

Brown, J.D. (2014) *Mixed methods research for TESOL.* Edinburgh: Edinburgh University Press.

Bruner, J.S. (1966) *Towards a Theory of Instruction.* Cambridge, Mass.: Harvard University Press.

Bruner, J.S. (1978) The role of dialogue in language acquisition. In: A. Sinclair, R., J. Jarvelle, and W. J.M. Levelt eds. The Child's Concept of Language. New York: Springer-Verlag.

Bryman, A. (2016) *Social research methods, 5th ed.* Oxford: Oxford university press.

Capel, S., Leask, M. and Turner, T (2013) *Learning to teach in the secondary school: a companion to school experience, 6th ed*. Abingdon: Routledge.

Clark, J.M. and Paivio, A. (1991) Dual coding theory and education. *Educational psychology review*, 3 (3), pp.149-210.

Cohen, L., Manion, L. and Morrison, K. (2013) *Research methods in education, 7th ed*. Abingdon: Routledge.

Coe, R. (2002) It's the Effect Size, Stupid: What effect size is and why it is important. Annual *Conference of the British Educational Research Association.* University of Exeter, 12th-14th September 2002. Available at: https://www.leeds.ac.uk/educol/documents/00002182.htm [Accessed: 27th July 2017]

DeKeyser, R.M. (1998) Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty and J. Williams eds. *Focus on Form in Classroom Second Language Acquisition.* Cambridge: Cambridge University Press, pp.42-63.

Dörnyei, Z. (2007) *Research methods in applied linguistics*. Oxford: Oxford University Press.

Duolingo (2017) *Duolingo language courses - Duolingo.* Available at: <https://www.duolingo.com/courses/all> [Accessed: 20th July 2017]

Egbert, J.L. (2005) Conducting research on CALL. In: J.L. Egbert, G.M. Petrie eds. *CALL Research Perspectives*. New York: Taylor and Francis, pp.3-8.

Ellis, N.C. (2005) At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in second language acquisition*, 27(2), pp.305-352.

EPALE (2016) *6 of the best free language learning sites.* Available at: <https://ec.europa.eu/epale/en/blog/6-best-free-language-learning-sites> [Accessed: 2nd August 2017]

Erlam, R. (2003) Evaluating the relative effectiveness of structured-input and output-based instruction in foreign language learning. *Studies in second language acquisition,* 25 (4), pp.559-582.

French, C.C. and Colman, A.M. (1995) *Cognitive Psychology*. London: Longman.

Ghent University (2017) *Data manipulation, analysis and visualisation in Python*. Available at: [http://www.ugent.be/doctoralschools/en/doctoraltraining/courses/ specialistcourses/bse/data-manipulation-visualisation-python.htm](http://www.ugent.be/doctoralschools/en/doctoraltraining/courses/specialistcourses/bse/data-manipulation-visualisation-python.htm) . [Accessed: 5th August 2017]

Gorard, S. (2001) *Quantitative methods in educational research: the role of numbers made easy.* London: Bloomsbury.

Green, D. (2016) *Procedural content generation for C++ game development.* [eBook] Birmingham: Packt Publishing.

Hewson, C. (2003) Conducting research on the internet. *PSYCHOLOGIST-LEICESTER-*, 16(6), pp.290-293.

Hone, K.S. and El Said, G.R. (2016) Exploring the factors affecting MOOC retention: A survey study. *Computers & Education, 98,* pp. 157-168.

Hubbard, P. and Siskin, C.B. (2004)Another look at tutorial CALL*. ReCALL*, 16 (2), pp.448-461.

Hunter, J.D. (2007) Matplotlib: A 2D graphics environment. *Computing In Science & Engineering,* 9 (3), pp.90-95.

Izumi, S. (2002) Output, input enhancement, and the noticing hypothesis. *Studies in second language acquisition,* 24 (4)*,* pp.541-577.

Jarvis, H. and Szymczyk, M. (2010)Student views on learning grammar with web-and book-based materials. *ELT journal,* 64 (1)*,* pp.32-44*.*

Khalil, H. and Ebner, M. (2014) MOOCs Completion Rates and Possible Methods to Improve Retention - A Literature Review. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2014*. Chesapeake, VA: AACE, pp. 1236-1244.

Krashen, S. (1977) Some issues relating to the monitor model. In H. Brown, C. Yorio, and R. Crymes, eds. *Teaching and learning English as a Second Language: Trends in Research and Practice: On TESOL '77: Selected Papers from the Eleventh Annual Convention of Teachers of English to Speakers of Other Languages,* Miami, Florida, April 26 – May 1, 1977. Washington, DC: TESOL, pp. 144–158.

Krashen, S. (1982) *Principles and practice in second language acquisition.* Available at: <http://www.sdkrashen.com/content/books/principles_and_practice.pdf> . [Accessed: 10th August 2017]

Krashen, S.D. (1985) *The Input Hypothesis: Issues and Implications.* New York: Longman

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J. and Couper, M. (2004) Psychological research online: report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *American psychologist*, 59 (2), p.105.

Kyriacou, C. (2009) *Effective teaching in schools: Theory and practice*, 3rd ed. Cheltenham: Nelson Thornes.

Lantolf, J.P. (2011) The sociocultural approach to second language acquisition: Sociocultural theory, second language acquisition, and artificial L2 development. In: D. Atkinson ed. *Alternative approaches to second language acquisition.* Abingdon: Routledge.

Lee, J.F. and Benati, A. (2007) Comparing modes of delivering processing instruction and meaning-based output instruction on Italian and French subjunctive. In: J.F. Lee and A. Benati eds. *Delivering processing instruction in classrooms and in virtual contexts: Research and practice.* London: Equinox, pp.99-136.

Leeser, M.J. (2008) Pushed output, noticing, and development of past tense morphology in content-based instruction. *Canadian Modern Language Review,* 65 (2), pp.195-220.

Levy, M. (1995) Integrating CALL: the tutor and the tool. In: A. Gimeno ed. *Technology Enhanced Language Learning: Focus on Integration, Proceedings of EuroCALL ’95.* Valencia: Universidad Politécnica de Valencia, pp.239-248

Levy, M. (1997) *CALL: Context and conceptualization.* Oxford: Clarendon Press.

Matplotlib.org (2017) *matplotlib.pyplot.boxplot*. Available at: <https://matplotlib.org/2.0.2/api/pyplot_api.html#matplotlib.pyplot.boxplot> [Accessed: 4th August 2017]

McDonough, J. and McDonough, S. (1997) *Research methods for English language teachers.* London: Arnold.

Mitchell, R., Myles, F. and Marsden, E. (2013) *Second language learning theories, 3rd ed*. Abingdon: Routledge.

Morgan-Short, K. and Bowden, H.W. (2006) Processing instruction and meaningful output-based instruction: effects on second language development. *Studies in Second Language Acquisition*, 28 (1), pp.31-65.

Mozilla (2017) *HTML | MDN*. Available at: <https://developer.mozilla.org/en-US/docs/Web/HTML> [Accessed: 12th August 2017]

Nagata, N. (1998) Input vs. output practice in educational software for second language acquisition. *Language Learning & Technology,* 1 (2), pp.23-40.

Naiman, N. ed. (1978) *The good language learner (Vol. 4)*. Bristol: Multilingual Matters.

Opie, C. (2004a) Presenting Data. In: C. Opie ed. *Doing educational research.* London: SAGE Publications, pp.130-178.

Opie, C. (2004b) Quantitative Analysis: A Glossary of Terms. In: C. Opie ed *Doing educational research.* London: SAGE Publications, pp.119-299.

Ortega, L. (2009) *Understanding second language acquisition.* Abingdon: Routledge.

Ortega, L. (2011) SLA after the social turn: Where cognitivism and its alternatives stand. In: D. Atkinson ed. *Alternative approaches to second language acquisition.* Abingdon: Routledge.

Paivio, A. (1971) *Imagery and Verbal Processes.* New York: Holt, Rinehart and Winston, Inc.

Preece, J., Sharp, H. and Rogers, Y. (2015) *Interaction design: beyond human-computer interaction, 4th ed*. Chichester: John Wiley & Sons.

Reips, U.-D. (2002a) Internet-based psychological experimenting: five dos and donts. *Social Science Computer Review*, 20 (3), pp.241-249.

Reips, U.-D. (2002b) Standards for internet-based experimenting. *Experimental Psychology*, 49 (4), pp.243-256.

Robin, R.M. (2013) CALL and less commonly taught languages. In: M. Thomas, H. Reinders and Warschauer, M eds. *Contemporary computer-assisted language learning*. London: Bloomsbury, pp.303-322.

Rosetta Stone (2017a) *Rosetta Stone* [homepage] Available at: <http://www.rosettastone.co.uk/> [Accessed: 27th July 2017]

Rosetta Stone (2017b) *Rosetta Stone® Language Courses | Choose Your New Language Course!* Available at: <http://www.rosettastone.co.uk/languages> [Accessed: 20th July 2017]

Scaife, J. (2004) Reliability, Validity and Credibility. In: Opie, C. (ed) *Doing educational research.* London: SAGE Publications, pp.58-72.

Schmidt, R.W. (1990) The role of consciousness in second language learning. *Applied linguistics,* 11 (2), pp.129-158.

SCORM.com (2009) *Run-time reference.* Available at: <https://scorm.com/scorm-explained/technical-scorm/run-time/run-time-reference/> [Accessed: 30th July 2017]

Scrivener, J. (2011) *Learning Teaching: The Essential Guide to English Language Teaching, 3rd ed*. Oxford: MacMillan.

Shintani, N. (2012) Input-based tasks and the acquisition of vocabulary and grammar: A process-product study. *Language Teaching Research*, 16 (2), pp.253-279.

Shintani, N. and Ellis, R. (2010) The incidental acquisition of English plural -s by Japanese children in comprehension-based and production-based lessons. *Studies in Second Language Acquisition*, 32 (4), pp.607-637.

Shintani, N., Li, S. and Ellis, R. (2013) Comprehension‐Based Versus Production‐Based Grammar Instruction: A Meta‐Analysis of Comparative Studies. *Language Learning,* 63 (2), pp.296-329.

Skinner, B.F. (1968) *The Technology of Teaching.* New York: Appleton-Century-Cross.

Speigel Online (2016) *Sprachlern-App Babbel hat eine Million zahlende Nutzer*. Available at: <http://www.spiegel.de/netzwelt/apps/sprachlern-app-babbel-hat-eine-million-zahlende-nutzer-a-1077801.html> [Accessed: 28th July 2017]  
[English version via Google Translate: <https://translate.google.co.uk/translate?hl=en&sl=de&u=http://www.spiegel.de/netzwelt/apps/sprachlern-app-babbel-hat-eine-million-zahlende-nutzer-a-1077801.html&prev=search> ]

Swain, M. (1985) Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In: S. Gass and C. Madden eds. *Input in second language acquisition*. Rowley, MA: Newbury House*,* pp.235-253.

Swain, M. (1988) Manipulating and complementing content teaching to maximize second language learning. *TESL Canada Journal,* 6 (1)*,* pp.68-83.

TeachMe! (2017) *Learn Gaelic quickly.* Available at: <https://www.linguashop.com/learn-gaelic> [Accessed: 14th August 2017]

TechInAsia.com (2016) *How this language learning app got 110 million users without spending a dollar on marketing*. Available at: <https://www.techinasia.com/how-duolingo-got-110-million-users>. [Accessed: 27th July 2017]

The-art-of-web (2017) *JavaScript: Form Validation: Credit Card numbers.* Available at: <http://www.the-art-of-web.com/javascript/validate/3/> [Accessed: 30th July 2017]

VanPatten, B. (1996) *Input processing and grammar instruction: Theory and research*. Norwood, NJ: Ablex.

VanPatten, B. and Oikkenon, S. (1996) Explanation versus structured input in processing instruction. *Studies in Second Language Acquisition*, 18 (4), pp.495-510.

Vesselinov, R. (2009) *Measuring the effectiveness of Rosetta Stone.* Available at: http://comparelanguageapps.com/documentation/EffectivenessStudy2008\_RS.pdf

Vesselinov, R. and Grego, J. (2012) *Duolingo Effectiveness Study: Final report.* Available at: <http://comparelanguageapps.com/documentation/DuolingoReport_Final.pdf>

Vygotsky, L.S. (1978) *Mind and Society: The Development of Higher Psychological Processes.* Cambridge, MA: Harvard University Press.

w3schools(2017) *HTML5 Introduction.* Available at: <https://www.w3schools.com/html/html5_intro.asp> [Accessed: 12th August 2017]

w3techs.com (2017) *Usage Statistics and Market Share of Content Languages for Websites, July 2017*. Available at: <https://w3techs.com/technologies/overview/content_language/all> [Accessed: 20th July 2017]

Watson, J.B. (1913) Psychology as the Behaviorist views it. *Psychological Review,* 20*,* pp.158-177. Available at: <http://psychclassics.yorku.ca/Watson/views.htm>. [Accessed: 7th August 2017]

Wegener, D.T. and Blankenship, K.L. (2007) Ecological validity. In: R.F. Baumeister and K.D. Vohs eds. *Encyclopedia of Social Psychology*. Thousand Oaks: SAGE Publications. Available at: <http://sk.sagepub.com/reference/socialpsychology/n167.xml> [Accessed: 28th July 2017]

1. Glossary of Gaelic used in this study

Colours

|  |  |  |
| --- | --- | --- |
| **English** | **Gaelic singular** | **Gaelic plural  (where different)** |
| Blue | Gorm | Gorma |
| Green | Uaine |  |
| Red | Dearg | Dearga |
| Yellow | Buidhe |  |

Items of clothing:

|  |  |  |
| --- | --- | --- |
| **English** | **Gaelic** | **Grammatical number** |
| Jumper | Geansaidh | Singular |
| Trousers | Triubhas | Singular |
| Glasses | Speuclairean | Plural |
| Shoes | Brògan | Plural |

Target phrase

|  |  |  |
| --- | --- | --- |
| **English meaning** | **Gaelic** | **Hyperliteral gloss** |
| He’s got a <XX> on/He’s wearing a <XX>. | Tha <XX> air. | Is <XX> on-him. |
| She’s got a <XX> on/She’s wearing a <XX>. | Tha <XX> oirre. | Is <XX> on-her. |
| He’s got a <XX> and a <YY> on/He’s wearing a <XX> and a <YY>. | Tha <XX> agus <YY> air. | Is <XX> and <YY> on-him. |
| She’s got a <XX> and a <YY> on/She’s wearing a <XX> and a <YY>. | Tha <XX> agus <YY> oirre. | Is <XX> and <YY> on-her. |

1. Sign-up survey

Below is a complete copy of the sign-up survey, incorporating the study consent form and the collection of demographic data and language history.

Note that questions marked \* are required – the survey cannot be submitted until these have been answered.

Gender has deliberately been left as optional to account for people who may feel uncomfortable giving that information or who do not identify with gender.

Consent for interviewing has also been left as optional, as it was felt that some potential participants may prefer to leave it blank than to outright refuse, and there was a concern that forcing respondents to do so might result in higher abandonment.

* 1. Page 1

1. Thank you for participating in this study. My name is Niall Tracey, and I am a masters student at Stirling University. I am investigating the effects of practice on language learners. Participants in this study will access an online language learning website and will be assigned a short programme of learning. This programme is delivered over three weeks and there is a short quiz at the end of each week to monitor your progress. Please note that your interactions with the system will be recorded for research purposes. However, when the data is used in the study, it will be anonymised, and no personally identifiable data will be published. Participation in this study is entirely voluntary and you may withdraw at any time without question. If you choose to participate, a log in and password will be issued by email, headed MSc Research: New user account. If you have any questions about this research, you may contact me at npt00003@students.stir.ac.uk, or contact my research supervisor, Dr Edward Moran, at edward.moran@stir.ac.uk . \*  
   I wish to take part in this study and I consent to my data being stored as described above.
   1. Page 2 – Background
2. I am \_\_\_\_ years old. \*
3. I am \_\_\_\_.  
   male female
4. English is my first/native language. \*  
   yes no
   1. If you answered “no” above, what is your native language?
5. I have previous experience of studying Scottish Gaelic. \*  
   yes no
   1. If you answered yes above, what is your previous experience of Scottish Gaelic?
6. I have previous experience of studying other languages.   
   yes no
   1. If you answered yes above, please list the languages you have studied, along with an estimate of your level. (Beginner, intermediate, advanced, fluent.) Please use a new line for each language.
   2. Page 3 – Contact details
7. My current email address. (This is required to administer login details for the online platform and will be used only for the purposes of this study.)
8. As part of my research, I may wish to hold a brief spoken interview with a small number of participants via Skype. If you are willing to take part in this, please indicate below. Participants who are selected for interview will be contacted towards the end of the study to arrange a suitable time.   
   I would be willing to take part a spoken interview and test.   
   I do not wish to take part in a spoken interview or test.
   1. Page 4 – Thank you for your participation.

Thank you for choosing to take part in this research study. If you have any questions, please feel free to contact me at [researcher’s\_username]@students.stir.ac.uk .

You should receive an email within the next five days providing login details and instructions for the learning platform.

1. Study participant demographics

Note that due to a technical error, numbering for female participants in group C begins at 3. This became apparent during analysis, but as results are not affected, this was left uncorrected to avoid the risk of introducing errors and inconsistencies in the data.

* 1. Age, gender and language background
     1. Genuine beginners who completed module 1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **alias** | **Age** | **Gender** | **L1  (if not English)** | **Previous language study?** | **Details of previous language study** |
| a1f | 23 | female |  | Yes | French, advanced Portuguese, intermediate Spanish, beginner Greek, beginner |
| a2f | 23 | female |  | Yes | French - beginner Spanish - beginner/intermediate |
| a3f | 22 | female | German/English bilingual | Yes | English&German Beginner French, Korean, Japanese |
| a4f | 40 | female |  | Yes | German (intermediate) French (beginner) |
| a5f | 29 | female |  | Yes |  |
| a6f | 50 | female |  | Yes | French (beginner) German (Austrian) (fluent) |
| a7f | 23 | female | German, but added English at 3-4y old | Yes | French, intermediate Latin, advanced English, fluent |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **alias** | **Age** | **Gender** | **L1  (if not English)** | **Previous language study?** | **Details of previous language study** |
| b1f | 54 | female |  | Yes | French : beginner German: intermediate |
| b1m | 27 | male |  | Yes | German - intermediate French - (pre)intermediate Russian - intermediate Portuguese - beginner Slovak - beginner |
| b2f | 31 | female |  | Yes | Intermediate German Beginner Croat |
| b3f | 60 | female |  | Yes | French advanced |
| b4f | 32 | female |  | Yes | GCSE French Basic Spanish & Thai |
| b5f | 44 | female |  | Yes | beginner french |
| b6f | 20 | female |  | Yes | Spanish- Reached intermediate level but haven't studied for ~5 years. Italian- Intermediate. |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **alias** | **Age** | **Gender** | **L1  (if not English)** | **Previous language study?** | **Details of previous language study** |
| c1m | 54 | male |  | No |  |
| c3f | 55 | female |  | Yes | Latin 1yr very beginner German O Grade A French O Grade A French Higher B and lived 3 months in France aged 17 |
| c4f | 47 | female |  | No |  |
| c5f | 59 | female |  | Yes | French at school to o level but that was 43yrs ago!! |
| c6f | 28 | female |  | Yes | French-Intermediate Norwegian- Beginner |
| c7f | 40 | female | German | Yes | English (advanced) Spanish (intermediate) French (beginner, long forgotten) |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **alias** | **Age** | **Gender** | **L1  (if not English)** | **Previous language study?** | **Details of previous language study** |
| d1f | 26 | female | Danish | Yes | English fluent German advanced Spanish intermediate |
| d1m | 53 | male |  | Yes | German, beginner Italian, beginner |
| d1u | 58 | *withheld* | italian | Yes | French advanced English fluent |
| d2f | 46 | female |  | Yes | French O Grade (at school, many moons ago!) |
| d2m | 45 | male |  | Yes | French O grade and beginners Italian class |
| d3f | 28 | female |  | Yes | French (intermediate) German (intermediate) Spanish (intermediate) |
| d3m | 54 | male |  | Yes | Spanish (intermediate) German (beginner) Ancient Greek (intermediate) American Sign Language (beginner) |
| d4f | 37 | female |  | Yes | French, 7years, fluent German, 5 years, beginner (very much lapsed) Latin, 1 year, beginner Italian, 2 months, beginner |
| d4m | 39 | male | French | Yes | Finnish (Advanced) Arabic (Advanced) |
| d5f | 25 | female | Spanish | Yes | English (Advanced) Portuguese (Intermediate) German (Beginner) Arabic (Beginner) |
| d5m | 40 | male | Polish | Yes | English – fluent German – advanced Spanish – advanced French – intermediate Russian – intermediate |

1. Emails sent during the study
   1. Welcome email/login details

The following email was sent to candidates who completed the entrance survey and were selected to take part in the study:

Thanks again for your participation in this study.

*Note that users have reported problems using the exercises from Safari on Mac OS and Internet Explorer on Windows. If the site does not work correctly on your computer, please try installing* [*Firefox*](https://www.mozilla.org/en-GB/firefox/) *or* [*Google Chrome*](https://www.google.com/chrome/)*.*

**The address for the site is:** [**https://lingua.guru/login**](https://lingua.guru/login)

**Your username is:***[username]*

**Your password is:** *[password]*

*Note: your password consists of 1 capital letter, 3 lower-case letters and 4 digits. You will be asked to change this on first login.*

The first item available on the site is titled:

***Progress check: pre-course quiz***

This is to verify your knowledge of Gaelic before starting the course and should be completed before anything else. This can be completed now or directly prior to starting the material that will be available on Monday.

Over the three weeks that the course is running, additional materials will be released regularly – the first of these will become available on Monday morning.

The work for each week will consist of a *module* of 4 lessons, and a *progress check*. Each lesson and the progress check should take approximately 10-20 minutes to complete. Try if possible to do one lesson a day, and to complete the progress check the day after the last lesson for the week.

Note that during the study, each exercise can only be completed once. However, once the research is complete, all users who complete the study will be given access to the full material for repeated use as well as additional bonus lessons covering a variety of topics.

Further instructions are included with the exercises on the site.

If you have any difficulties accessing or using the site, please do not hesitate to get in touch.

Again, thank you for your participation.

Le meas,

Níall Tracey.

* 1. Invitation to complete immediate post-test

The following was sent to participants within one day of completing the fourth and final teaching unit of module 1.

**Subject: *Gaelic: progress check 1***

Hide original message

*Progress check 1*, the end-of-module quiz for module 1, has been made available to you on the web site.

If you have completed module 1, please complete this quiz in the next 24 hours, otherwise please do it a day after completing the module. Please do this **without** the use of notes or other aids, such as online dictionaries.

Also, I would ask that you do no revision of this topic during the coming week, as there will be another progress check in seven days’ time to measure how much you have remembered or forgotten of the language.  
  
Remember that full notes and extra practice for the language covered in this course will be available once my research is complete, so you will have the opportunity to revise this material later. Additional material covering other new language will also be available.

Le meas,

Níall.

* 1. Invitation to complete delayed post-test

The following email was sent to participants 6 days after completing the immediate post-test.

**Subject: *Gaelic: progress check 1b now available.***

Hide original message

Hi,

Progress check 1b is now available. Please complete this quiz as soon as possible.   
  
As with the previous progress check 1, please do this **without** the use of notes or other aids, such as online dictionaries.

Thanks again for your time,

Níall

1. Summary of participants’ scores by group.

This appendix shows the distribution of participants’ scores in the various sections of the post-test.

Throughout this section, the immediate post-test is shown on the left with the delayed post-test on the right for comparison.

Each individual number in the “All Results” row represents a single user score. These scores have been sorted numerically to provide a rapid visual reference to the reader of the distribution of scores.

Data in each “All Results” row has been centred vertically to provide a quick visual confirmation of median values for each column, in addition to the median presented at the bottom of the column.

Within the two productive sections of the post-test, two sets of figures are presented. The first shows the scores with only obviously mistyped words corrected. The second disregards all spelling errors that do not render the word unrecognisable to the marker.

Mean, standard deviation and median values were calculated automatically using the module *statistics* for Python 3.6.

* 1. Section I: Productive (before receptive)
     1. Overall results (minor typographical correcting)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | Immediate post-test; Part I | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | 0 5 6 7 8 8 8 | 1 3 6 7 7 8 8 | 0 0 0 0 0 0 | 0 0 0 4 5 5 6 7 8 8 8 | | Mean | 6 | 5.71 | 0 | 4.64 | | Std deviation | 2.887 | 2.69 | 0.0 | 3.264 | | Median | 7 | 7 | 0.0 | 5 |   Table E‑: Scores by group, immediate post-test, Part I (productive skills) | |  |  |  |  |  | | --- | --- | --- | --- | --- | | Delayed post-test; Part I | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | 0 0 0 3 6 8 8 | 0 0 0 0 0 2 5 | 0 0 0 0 0 1 | 0 0 0 0 2 4 5 5 6 7 7 | | Mean | 3.57 | 1 | 0.17 | 3.27 | | Std deviation | 3.735 | 1.915 | 0.408 | 2.936 | | Median | 3 | 0 | 0.0 | 4 |   Table E‑: Scores by group, delayed post-test, Part I (productive skills) |

* + 1. Scores disregarding spelling errors

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | Immediate post-test; Part I Spelling errors corrected. | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | 5 6 7 7 8 8 8 | 3 7 7 7 8 8 8 | 0 0 0 0 5 5 | 0 5 5 7 7 7 7 8 8 8 8 | | Mean | 7 | 6.86 | 1.67 | 6.36 | | Std deviation | 1.155 | 1.773 | 2.582 | 2.378 | | Median | 7 | 7 | 0.0 | 7 |   Table E‑: Scores by group, immediate post-test, Part I (productive skills) (Spelling errors discounted) | |  |  |  |  |  | | --- | --- | --- | --- | --- | | Delayed post-test; Part I Spelling errors corrected. | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | 0 0 0 7 8 8 8 | 0 0 0 0 0 5 6 | 0 0 0 0 0 5 | 0 0 0 2 5 5 7 7 7 8 8 | | Mean | 4.43 | 1.57 | 0.83 | 4.45 | | Std deviation | 4.158 | 2.699 | 2.041 | 3.328 | | Median | 7 | 0 | 0.0 | 5 |   Table E‑: Scores by group, delayed post-test, Part I (productive skills) (Spelling errors discounted) |

* 1. Section II: Receptive

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | Immediate post-test; Part II | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | 13 14 15 16 16 16 16 | 14 15 15 15 16 16 16 | 13 13 14 15 15 15 | 13 15 15 15 16 16 16 16 16 16 16 | | Mean | 15.14 | 15.29 | 14.17 | 15.45 | | Std deviation | 1.215 | 0.756 | 0.983 | 0.934 | | Median | 16 | 15 | 14.5 | 16 |   Table E‑: Scores by group, immediate post-test, Part II (receptive skills) | |  |  |  |  |  | | --- | --- | --- | --- | --- | | Delayed post-test; Part II | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | 16 16 16 16 16 16 16 | 12 15 15 15 16 16 16 | 12 14 14 15 16 16 | 14 14 15 15 16 16 16 16 16 16 16 | | Mean | 16 | 15 | 14.5 | 15.45 | | Std deviation | 0.0 | 1.414 | 1.517 | 0.82 | | Median | 16 | 15 | 14.5 | 16 |   Table E‑: Scores by group, delayed post-test, Part II (receptive skills) |

* 1. Section III: Productive
     1. Base scores (minor typographical errors corrected)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | Immediate post-test; Part III | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | 6 7 7 8 8 8 8 | 1 5 6 6 7 8 8 | 0 0 0 0 0 8 | 0 1 5 7 7 7 7 7 8 8 8 | | Mean | 7.43 | 5.86 | 1.33 | 5.91 | | Std deviation | 0.787 | 2.41 | 3.266 | 2.809 | | Median | 8 | 6 | 0.0 | 7 |   Table E‑: Scores by group, immediate post-test, Part III (productive skills) | |  |  |  |  |  | | --- | --- | --- | --- | --- | | Delayed post-test; Part III | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | 2 8 8 8 8 8 8 | 0 5 5 6 6 8 8 | 0 0 0 1 2 8 | 0 0 3 4 6 7 7 7 8 8 8 | | Mean | 7.14 | 5.43 | 1.83 | 5.27 | | Std deviation | 2.268 | 2.699 | 3.125 | 3.069 | | Median | 8 | 6 | 0.5 | 7 |   Table E‑: Scores by group, delayed post-test, Part III (productive skills) |

* + 1. Scores, disregarding spelling errors

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | Immediate post-test; Part III Spelling errors corrected. | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | 6 7 8 8 8 8 8 | 5 6 7 8 8 8 8 | 0 1 1 5 7 8 | 2 2 5 7 7 7 7 8 8 8 8 | | Mean | 7.57 | 7.14 | 3.67 | 6.27 | | Std deviation | 0.787 | 1.215 | 3.445 | 2.284 | | Median | 8 | 8 | 3.0 | 7 |   Table E‑: Scores by group, immediate post-test, Part III (productive skills) (spelling errors discounted) | |  |  |  |  |  | | --- | --- | --- | --- | --- | | Delayed post-test; Part III Spelling errors corrected. | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | 7 8 8 8 8 8 8 | 5 5 6 6 8 8 8 | 0 0 2 4 5 8 | 0 4 4 7 7 7 8 8 8 8 8 | | Mean | 7.86 | 6.57 | 3.17 | 6.27 | | Std deviation | 0.378 | 1.397 | 3.125 | 2.573 | | Median | 8 | 6 | 3.0 | 7 |   Table E‑: Scores by group, immediate post-test, Part III (receptive skills) (spelling errors discounted) |

1. User improvement/decline:

This appendix shows the distribution of participants’ improvement or decline by comparing individual users’ scores in the various sections of the post-test and then aggregating by group.

A positive number indicates that the subject obtained a higher score in the delayed post-test, demonstrating improved performance, whereas a negative number indicates a higher score in the immediate post-test, and a decline in performance.

Each individual number in the “All Results” row represents a single user score. These scores have been sorted numerically to provide a rapid visual reference to the reader of the distribution of scores.

Data in each “All Results” row has been centred vertically to provide a quick visual confirmation of median values for each column, in addition to the median presented at the bottom of the column.

Within the two productive sections of the post-test, two sets of figures are presented. The first shows the scores with only obviously mistyped words corrected. The second disregards all spelling errors that do not render the word unrecognisable to the marker. The version considering spelling errors as errors is shown on the left, and the version disregarding spelling errors is shown on the right.

Mean, standard deviation and median values were calculated automatically using the module *statistics* for Python 3.6.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | Difference; Part I | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | -8 -8 -5 -4 0 0 8 | -8 -7 -7 -4 -3 -3 -1 | 0 0 0 0 0 1 | -8 -5 -5 -2 -1 -1 0 0 0 2 5 | | Mean | -2.43 | -4.71 | 0.17 | -1.36 | | Std deviation | 5.65 | 2.63 | 0.41 | 3.59 | | Median | -4 | -4 | 0.0 | -1 |   Table F‑: Part I: Productive post-tests: relative performance by group. | |  |  |  |  |  | | --- | --- | --- | --- | --- | | Difference; Part I Spelling errors corrected. | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | -8 -8 -5 0 1 1 1 | -8 -7 -7 -7 -3 -3 -2 | -5 0 0 0 0 0 | -8 -7 -5 -3 0 0 0 0 0 0 2 | | Mean | -2.57 | -5.29 | -0.83 | -1.91 | | Std deviation | 4.28 | 2.5 | 2.04 | 3.33 | | Median | 0 | -7 | 0.0 | 0 |   Table F-2: Part I: Productive post-tests: relative performance by group. (Spelling errors discounted.) |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | Difference; Part II | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | 0 0 0 0 1 2 3 | -2 -1 0 0 0 0 1 | -1 -1 -1 1 2 2 | -2 -1 -1 0 0 0 0 0 0 1 3 | | Mean | 0.86 | -0.29 | 0.33 | 0 | | Std deviation | 1.21 | 0.95 | 1.51 | 1.26 | | Median | 0 | 0 | 0.0 | 0 |   Table F-3: Part II: Receptive post-tests: relative performance by group. | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | Difference; Part III | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | -5 0 0 0 0 1 2 | -1 -1 -1 -1 0 0 1 | 0 0 0 0 1 2 | -7 -2 -1 -1 -1 0 0 0 0 1 4 | | Mean | -0.29 | -0.43 | 0.5 | -0.64 | | Std deviation | 2.21 | 0.79 | 0.84 | 2.62 | | Median | 0 | -1 | 0.0 | 0 |   Table F-4: Part III: Productive post-tests: relative performance by group. | |  |  |  |  |  | | --- | --- | --- | --- | --- | | Difference; Part III Spelling errors corrected. | | | | | |  | a:P | b:PR | c:R | d:RP | | All results | 0 0 0 0 0 0 2 | -3 -1 -1 0 0 0 1 | -5 -1 -1 0 0 4 | -2 -1 -1 -1 0 0 0 1 1 1 2 | | Mean | 0.29 | -0.57 | -0.5 | 0 | | Std deviation | 0.76 | 1.27 | 2.88 | 1.18 | | Median | 0 | 0 | -0.5 | 0 |   Table F-5: Part III: Productive post-tests: relative performance by group. (Spelling errors discounted.) |

1. Individual participants’ scores

The following tables catalogues individual participants’ scores across the immediate post-test and delayed post-test, and also shows the improvement (or otherwise) of participants’ scores between each section of the immediate and delayed post-test. A positive number indicates a higher score in the delayed post-test (hence improved performance) and a negative number indicates a lower score (hence a decline in performance).

For each productive score, there are two values; as elsewhere these indicate a score with only minor typing errors excluded and a score with all spelling errors disregarded (assuming it is clear to the marker that it represents the correct target form).

The upper number in the column is the base score, the lower is the score with spelling errors disregarded.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group | Subject |  | Immediate post-test | | |  | Delayed post-test | | |  | Difference:  delayed-immediate. | | |
| 1:P | 2:R | 3:P | 1:P | 2:R | 3:P | 1:P | 2:R | 3:P |
| a:P | a1f |  | 7 | 16 | 7 |  | 3 | 16 | 2 |  | -4 | 0 | -5 |
| 7 | 7 | 8 | 7 | 1 | 0 |
| a2f |  | 5 | 14 | 7 |  | 0 | 16 | 8 |  | -5 | 2 | 1 |
| 5 | 8 | 0 | 8 | -5 | 0 |
| a3f |  | 0 | 16 | 8 |  | 8 | 16 | 8 |  | 8 | 0 | 0 |
| 7 | 8 | 8 | 8 | 1 | 0 |
| a4f |  | 8 | 16 | 6 |  | 0 | 16 | 8 |  | -8 | 0 | 2 |
| 8 | 6 | 0 | 8 | -8 | 2 |
| a5f |  | 8 | 16 | 8 |  | 8 | 16 | 8 |  | 0 | 0 | 0 |
| 8 | 8 | 8 | 8 | 0 | 0 |
| a6f |  | 6 | 13 | 8 |  | 6 | 16 | 8 |  | 0 | 3 | 0 |
| 6 | 8 | 7 | 8 | 1 | 0 |
| a7f |  | 8 | 15 | 8 |  | 0 | 16 | 8 |  | -8 | 1 | 0 |
| 8 | 8 | 0 | 8 | -8 | 0 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group | Subject |  | Immediate post-test | | |  | Delayed post-test | | |  | Difference: delayed-immediate | | |
| 1:P | 2:R | 3:P | 1:P | 2:R | 3:P | 1:P | 2:R | 3:P |
| b:PR | b1f |  | 8 | 15 | 6 |  | 5 | 16 | 5 |  | -3 | 1 | -1 |
| 8 | 6 | 6 | 5 | -2 | -1 |
| b1m |  | 8 | 16 | 8 |  | 0 | 16 | 8 |  | -8 | 0 | 0 |
| 8 | 8 | 0 | 8 | -8 | 0 |
| b2f |  | 3 | 15 | 5 |  | 0 | 15 | 6 |  | -3 | 0 | 1 |
| 3 | 5 | 0 | 6 | -3 | 1 |
| b3f |  | 1 | 16 | 1 |  | 0 | 16 | 0 |  | -1 | 0 | -1 |
| 7 | 8 | 0 | 5 | -7 | -3 |
| b4f |  | 7 | 14 | 8 |  | 0 | 12 | 8 |  | -7 | -2 | 0 |
| 7 | 8 | 0 | 8 | -7 | 0 |
| b5f |  | 6 | 15 | 6 |  | 2 | 15 | 5 |  | -4 | 0 | -1 |
| 8 | 8 | 5 | 8 | -3 | 0 |
| b6f |  | 7 | 16 | 7 |  | 0 | 15 | 6 |  | -7 | -1 | -1 |
| 7 | 7 | 0 | 6 | -7 | -1 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group | Subject |  | Immediate post-test | | |  | Delayed post-test | | |  | Difference: delayed-immediate | | |
| 1:P | 2:R | 3:P | 1:P | 2:R | 3:P | 1:P | 2:R | 3:P |
| c:R | c1m |  | 0 | 13 | 0 |  | 0 | 15 | 0 |  | 0 | 2 | 0 |
| 5 | 1 | 0 | 0 | -5 | -1 |
| c3f |  | 0 | 13 | 0 |  | 0 | 12 | 0 |  | 0 | -1 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 |
| c4f |  | 0 | 14 | 0 |  | 0 | 16 | 0 |  | 0 | 2 | 0 |
| 0 | 7 | 0 | 2 | 0 | -5 |
| c5f |  | 0 | 15 | 0 |  | 0 | 16 | 1 |  | 0 | 1 | 1 |
| 0 | 1 | 0 | 5 | 0 | 4 |
| c6f |  | 0 | 15 | 0 |  | 0 | 14 | 2 |  | 0 | -1 | 2 |
| 0 | 5 | 0 | 4 | 0 | -1 |
| c7f |  | 0 | 15 | 8 |  | 1 | 14 | 8 |  | 1 | -1 | 0 |
| 5 | 8 | 5 | 8 | 0 | 0 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group | Subject |  | Immediate post-test | | |  | Delayed post-test | | |  | Difference: delayed-immediate | | |
| 1:P | 2:R | 3:P | 1:P | 2:R | 3:P | 1:P | 2:R | 3:P |
| d:RP | d1f |  | 7 | 16 | 8 |  | 7 | 15 | 8 |  | 0 | -1 | 0 |
| 7 | 8 | 7 | 8 | 0 | 0 |
| d1m |  | 4 | 13 | 0 |  | 4 | 16 | 4 |  | 0 | 3 | 4 |
| 5 | 2 | 5 | 4 | 0 | 2 |
| d1u |  | 0 | 15 | 1 |  | 2 | 14 | 0 |  | 2 | -1 | -1 |
| 0 | 2 | 2 | 0 | 2 | -2 |
| d2f |  | 8 | 16 | 7 |  | 7 | 16 | 8 |  | -1 | 0 | 1 |
| 8 | 7 | 8 | 8 | 0 | 1 |
| d2m |  | 5 | 16 | 7 |  | 0 | 16 | 0 |  | -5 | 0 | -7 |
| 7 | 7 | 0 | 8 | -7 | 1 |
| d3f |  | 0 | 16 | 8 |  | 0 | 16 | 7 |  | 0 | 0 | -1 |
| 7 | 8 | 7 | 7 | 0 | -1 |
| d3m |  | 5 | 16 | 5 |  | 0 | 14 | 3 |  | -5 | -2 | -2 |
| 5 | 5 | 0 | 4 | -5 | -1 |
| d4f |  | 0 | 15 | 7 |  | 5 | 16 | 7 |  | 5 | 1 | 0 |
| 8 | 7 | 5 | 8 | -3 | 1 |
| d4m |  | 6 | 16 | 7 |  | 5 | 16 | 7 |  | -1 | 0 | 0 |
| 7 | 8 | 7 | 7 | 0 | -1 |
| d5f |  | 8 | 15 | 7 |  | 6 | 15 | 6 |  | -2 | 0 | -1 |
| 8 | 7 | 8 | 7 | 0 | 0 |
| d5m |  | 8 | 16 | 8 |  | 0 | 16 | 8 |  | -8 | 0 | 0 |
| 8 | 8 | 0 | 8 | -8 | 0 |

1. Individual participants’ errors (receptive)

The following tables show the types of errors made by individual users in the receptive tasks (Part II) of the post-tests. The tables indicate correct or incorrect with ticks and crosses – following each cross is a code indicating the type of error:

V: lexical vocabulary (error in colour and/or item of clothing)  
 G: error in gender  
 V+G: combination of errors; both lexical vocabulary and gender

For each participant, a total of errors of each of the three types is given at the each of the row.

For ease of reference, participants have been grouped by treatment:

1. P
2. PR
3. R
4. RP
   1. Immediate post-test

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | 10 |  | 11 |  | 12 |  | 13 |  | 14 |  | 15 |  | 16 |  | 17 |  | 18 |  | 19 |  | 20 |  | 21 |  | 22 |  | 23 |  | 24 |  | 25 |  | V | V+G | G |
| a1f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| a2f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 2 | 0 |
| a3f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| a4f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 1 | 0 |
| a5f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| a6f |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✘ | g | 1 | 1 | 1 |
| a7f |  | ✔ |  | ✔ |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 1 | 0 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| b1f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 1 | 0 |
| b1m |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| b2f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 1 | 0 |
| b3f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| b4f |  | ✔ |  | ✔ |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 2 | 0 |
| b5f |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 1 | 0 |
| b6f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| c1m |  | ✔ |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 1 | 1 | 1 |
| c3f |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✘ | g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 2 | 0 | 1 |
| c4f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | 2 | 0 | 0 |
| c5f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 1 | 0 |
| c6f |  | ✔ |  | ✘ | g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 1 |
| c7f |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 1 | 0 | 0 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| d1f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| d1m |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v+g | ✔ |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | 1 | 2 | 0 |
| d1u |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 1 | 0 | 0 |
| d2f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| d2m |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| d3f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| d3m |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| d4f |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 1 | 0 | 0 |
| d4m |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| d5f |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 1 |
| d5m |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |

* 1. Delayed post-test

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 10 |  | 11 |  | 12 |  | 13 |  | 14 |  | 15 |  | 16 |  | 17 |  | 18 |  | 19 |  | 20 |  | 21 |  | 22 |  | 23 |  | 24 |  | 25 |  | v | v+g | g |
| a1f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| a2f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| a3f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| a4f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| a4f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| a6f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| a7f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| b1f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| b1m | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| b2f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v+g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 1 | 0 |
| b3f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| b4f | ✘ | v | ✘ | v | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | 4 | 0 | 0 |
| b5f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 1 | 0 | 0 |
| b6f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 1 | 0 | 0 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| c1m | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | 1 | 0 | 0 |
| c3f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✘ | v | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | 4 | 0 | 0 |
| c4f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| c5f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| c6f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 2 | 0 | 0 |
| c7f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✘ | g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 1 | 0 | 1 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| d1f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 1 | 0 | 0 |
| d1f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | v | ✔ |  | ✔ |  | 2 | 0 | 0 |
| d1m | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| d2f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| d2m | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| d3f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| d3m | ✔ |  | ✔ |  | ✘ | g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 2 |
| d4f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| d4m | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |
| d5f | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✘ | g | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 1 |
| d5m | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | ✔ |  | 0 | 0 | 0 |

2. Individual participants’ errors (productive)

The table[s] below show[s] the number of each type of productive error committed by each user, categorised by group. Individual counts are provided for each section (part I and part II) for each of the immediate and delayed post-tests.

For clarity and ease of reference, where no errors of a particular type have been committed, the number 0 has been omitted, and the cell left blank.

Note that the exact number of opportunities to make each class of error varies, and a direct comparison of numbers is therefore only possible down columns, not along rows. The first letter of each participant alias corresponds to treatment group, receiving the treatments as follows:

1. P
2. PR
3. R
4. RP

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Immediate | | | | | | | | | | | | | | | | Delayed | | | | | | | | | | | | | | | |
| Part I | | | | | | | | Part III | | | | | | | | Part I | | | | | | | | Part III | | | | | | | |
| s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z |
| a1f |  |  |  |  |  |  | 1 |  |  |  | 1 |  |  |  |  |  | 5 |  |  |  |  |  |  |  | 6 |  |  |  |  |  |  | 1 |
| a2f |  |  |  |  | 3 |  |  |  | 1 |  |  |  |  |  |  |  | 1 |  |  |  | 8 |  |  |  |  |  |  |  |  |  |  |  |
| a3f | 7 |  |  |  | 7 |  |  | 1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| a4f |  |  |  |  |  |  |  |  |  |  |  |  | 2 |  | 1 | 1 | 1 |  |  |  | 16 |  |  |  |  |  |  |  |  |  |  |  |
| a5f |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| a6f |  |  |  |  | 2 |  |  |  |  |  |  |  |  |  |  |  | 1 |  |  |  | 1 |  |  |  |  |  |  |  |  |  |  |  |
| a7f |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 8 |  |  |  |  |  |  |  |  |  |  |  |
|  | Immediate | | | | | | | | | | | | | | | | Delayed | | | | | | | | | | | | | | | |
| Part I | | | | | | | | Part III | | | | | | | | Part I | | | | | | | | Part III | | | | | | | |
| s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z |
| b1f |  |  |  |  |  |  |  |  |  |  |  |  | 2 |  |  |  | 1 | 2 |  |  | 1 |  |  |  |  |  |  |  | 2 |  | 2 |  |
| b2f | 2 |  |  |  | 1 | 1 | 6 |  |  |  |  |  |  |  | 3 |  |  |  | 2 |  | 14 | 7 | 6 | 8 |  |  | 2 |  |  |  | 2 |  |
| b3f | 11 |  |  |  | 1 |  |  |  | 9 |  |  |  |  |  |  |  | 12 | 5 |  |  | 23 |  |  |  | 10 | 2 |  |  | 1 |  |  |  |
| b4f |  |  | 1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 5 | 2 | 6 | 2 | 1 |  | 6 |  |  |  |  |  |  |  |  |
| b5f | 2 |  |  |  |  |  |  |  | 2 |  |  |  |  |  |  |  | 6 |  |  |  | 3 |  |  |  | 3 |  |  |  |  |  |  |  |
| b1m |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 8 |  | 6 |  |  |  |  |  |  |  |  |  |  |  |
| b6f |  |  |  |  | 1 |  |  |  |  |  |  |  | 7 |  |  | 1 |  |  |  |  | 7 |  | 2 | 1 |  |  |  |  |  |  | 2 |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Immediate | | | | | | | | | | | | | | | | Delayed | | | | | | | | | | | | | | | |
| Part I | | | | | | | | Part III | | | | | | | | Part I | | | | | | | | Part III | | | | | | | |
| s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z |
| c3f | 17 |  | 4 | 1 | 30 |  | 1 | 3 | 19 |  |  | 6 | 24 |  | 2 |  | 3 | 1 |  |  | 24 | 1 |  | 4 | 19 | 8 |  |  | 16 |  | 2 | 1 |
| c4f | 18 |  |  | 7 | 7 |  | 3 | 1 | 12 |  |  |  |  |  | 1 |  | 16 |  |  |  | 8 | 2 | 3 | 2 | 17 |  |  |  |  | 4 | 2 | 4 |
| c1m | 25 |  |  |  |  |  | 3 |  | 16 |  |  |  |  |  | 3 |  | 16 |  | 7 |  |  |  | 2 | 8 | 19 | 3 | 8 |  |  |  | 2 | 8 |
| c5f | 11 |  |  |  | 33 | 6 | 2 | 2 | 12 |  | 1 |  |  | 1 | 2 | 5 | 7 |  | 7 |  | 3 | 6 | 2 | 8 | 8 |  |  |  |  |  | 2 | 1 |
| c6f | 11 |  | 1 | 1 | 13 |  | 3 |  | 7 |  |  |  | 1 |  | 2 |  | 10 | 4 |  |  | 4 |  | 5 |  | 5 | 1 |  |  | 2 |  | 2 |  |
| c7f | 14 |  | 2 |  | 1 |  |  |  |  |  |  |  |  |  |  |  | 8 |  |  |  |  |  | 3 |  |  |  |  |  |  |  |  |  |
|  | Immediate | | | | | | | | | | | | | | | | Delayed | | | | | | | | | | | | | | | |
| Part I | | | | | | | | Part III | | | | | | | | Part I | | | | | | | | Part III | | | | | | | |
| s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z | s | u | o | e | m | # | a | z |
| d1m | 1 |  |  |  |  |  | 3 |  | 6 |  |  |  |  |  | 6 |  | 1 |  |  |  |  |  | 3 |  |  |  |  |  |  |  | 4 |  |
| d2m | 2 |  |  |  |  |  | 1 |  |  |  |  |  |  |  | 1 |  | 10 |  |  |  | 8 | 8 | 4 | 8 | 13 |  |  |  |  |  |  |  |
| d1f |  |  |  |  | 7 |  |  | 1 |  |  |  |  |  |  |  |  |  |  |  |  | 1 | 1 |  | 1 |  |  |  |  |  |  |  |  |
| d2f |  |  |  |  |  |  |  |  |  |  |  |  | 2 |  |  | 1 | 1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| d3m |  |  |  |  |  |  | 3 |  |  |  | 1 |  |  |  | 2 |  | 13 | 3 |  | 8 |  |  | 5 |  | 2 |  |  |  |  |  | 4 |  |
| d4m | 1 |  |  |  | 1 |  |  |  | 1 |  |  |  |  |  |  |  | 2 |  |  |  | 1 |  |  |  |  |  |  |  | 1 |  |  |  |
| d3f | 10 |  |  |  | 1 |  |  |  |  |  |  |  |  |  |  |  | 8 |  | 1 |  |  |  |  |  |  |  |  |  |  |  |  | 1 |
| d4f | 8 |  |  |  |  |  |  |  |  |  |  |  | 1 |  |  |  | 1 |  |  |  |  |  | 3 |  | 1 |  |  |  |  |  |  |  |
| d5m |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 1 |  | 14 |  |  |  |  |  |  |  |  |  |  |  |
| d5f |  |  |  |  |  |  |  |  |  |  |  |  | 1 |  |  |  | 2 |  |  |  |  |  |  |  | 1 |  |  |  | 1 |  |  |  |
| d1u | 6 |  |  |  | 10 |  |  |  | 4 |  |  |  | 7 |  |  |  | 6 |  |  |  | 6 |  | 1 |  | 2 |  |  |  | 9 |  |  |  |

Table I-1: Productive errors by user, group and test section

* 1. Errors of noun-adjective word order

|  |  |  |  |
| --- | --- | --- | --- |
| Treatment | Subject | Question number | Response |
| PR | b2f | 1 | Tha triubhas uaine agus dearga geansaidh |
| R | c5f | 2 | dearg geansaidh, bhuidh trousan, gorm bhrogan, uaine speuclaren |
| 3 | bhuidh...... gorm geansaidh |
| 4 | bhuidh trousan |
| 5 | dearg trousan, gorm brogan |
| 7 | gorm geansaidh, dearg trousan |
| 8 | gorm trousan, uaine brogan |
|  | (Note that subject c5f made no attempt at question 6, and entered only 1 word for question 1, therefore never employed the correct order in this section of the test.) |

Table I-2: Noun-adjective ordering errors, immediate post-test, Part I

|  |  |  |  |
| --- | --- | --- | --- |
| Treatment | Subject | Question number | Response |
| No errors of adjective order | | | |

Table I-3: Noun-adjective ordering errors, immediate post-test, Part III

|  |  |  |  |
| --- | --- | --- | --- |
| Treatment | Subject | Question number | Response |
| R | c3f | 6 | barracht gearacht (N.B. this subject used “barracht” consistently in place of dearg, so it is clear this is a mistaken form for “geansaidh dearg”) |
| c5f | 3 | a triubhas dearg agusuaine oirre bhrogan |

Table I-4: Noun-adjective ordering errors, delayed post-test, part I

|  |  |  |  |
| --- | --- | --- | --- |
| Treatment | Subject | Question number | Response |
| No errors of adjective order | | | |

Table I-5: Noun adjective ordering errors, delayed post-test, Part III

* 1. Detail of errors coded Z (other)

Note that in the tables below, only one form of the expected answer is included

* + 1. Immediate post-test
       1. Part I

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Treatment group | subject | Q# | Answer expected | Response given | notes |
| a:P | a3f | 6 | Tha speuclairean uaine agus triubhas gorm oirre. |  | Subject attempted other questions with high(??) degree of success – this is likely a problem with the user-computer interaction or a bug in the software. |
| c:R | c3f | 2 | Tha speuclairean uaine agus brògan gorma air. | speachlaren luine truisheas yellow | “yellow” shows subject misread task. |
| 3 | Tha geansaidh gorm agus speuclairean buidhe oirre. | geanseagh gorm truisseagh brock | Appears to have misread task again |
| 5 | Tha triubhas dearg agus brògan gorma air. | truisseagh brock geansseagh luine | Appears to have misread task again. |
| c:R | c4f | 6 | Tha speuclairean uaine agus triubhas gorm oirre. |  | Subject attempted other questions with reasonable degree of success – this is likely a problem with the user-computer interaction or a bug in the software. |
| c:R | c5f | 2 | Tha speuclairean uaine agus brògan gorma air. | dearg geansaidh, bhuidh trousan, gorm bhrogan, uaine speuclaren | Subject clearly misunderstood the task |
| 6 | Tha speuclairean uaine agus triubhas gorm oirre. |  | Subject made minimal attempts at all other tasks. It is not clear whether this blank response is deliberate |
| d:RP | d1f | 5 | Tha triubhas dearg agus brògan gorma air. |  | Subject got all other answers correct in this section. This is empty response is likely either a problem with the user-computer interaction or a bug in the software. |

* + - 1. Part III – qs 27-28

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Treatment group | subject | Q# | Answer expected | Response given | notes |
| a:P | a4f | 34 | Tha brògan gorma agus geansaidh buidhe oirre. | Tha geansaidh buidhe agus triubhas dearg oirre. | Subject appears to have misread question, as all other answers in this section were given correctly. |
| b:PR | b6f | 28 | Tha triubhas gorm agus brògan buidhe oirre. |  | Subject got all other answers correct in this section. This is empty response is likely either a problem with the user-computer interaction or a bug in the software. |
| c:R | c5f | 27 | Tha brògan buidhe agus speuclairean uaine air. | tha speuclairen a uaine agus brogan a bhuidhe air | This intrusive additional word “a” has no clear origin.  However, as the pattern of usage changes between instances, it is reasonable to assume that the subject is consciously unsure of the target pattern. |
| 28 | Tha triubhas gorm agus brògan buidhe oirre. | tha triubhas a gorm agus bhrogan bhuidhe oirre |
| 29 | Tha geansaidh gorm agus speuclairean uaine air. | tha speuclairean a uaine agus geansaidh a gorm air |
| 30 | Tha triubhas gorm agus speuclairean buidhe air. | tha speuclairean a bhuidhe agus triubhas gorm air |
| 31 | Tha geansaidh dearg agus brògan buidhe air. | tha geansaidh a dearg a bhrogan bhuidhe air |
| d:RP | d2f | 31 | Tha geansaidh dearg agus brògan buidhe air. | Tha geansaidh dearg agus triubhas uaine air | The subject got every other answer correct in this section. It appears the subject simply misread the diagram and looked at the wrong item of clothing. |

* + 1. Delayed post-test
       1. Part I

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Treatment group | subject | Q# | Answer expected | Response given | notes |
| b:PR | b2f | 1 | Tha geansaidh buidhe agus triubhas gorm oirre. | Tè | Use of word from second module, of correct gender. Interference. Subject attempted all subsequent questions with moderate success. |
| 2 | Tha brògan buidhe agus speuclairean dearga air. | Tha fear speuclairean dearg agus brògan buidhe | Subject consistnely uses the module 2 structure in place of module 1, employing the term of appropriate gender throughout. Clear evicence of interference. |
| 3 | Tha triubhas dearg agus brògan uaine oirre. | Tha tè triubhas dearga agus brògan uaine |
| 4 | Tha geansaidh uaine agus brògan gorma oirre. | Tha tè geansaidh uaine agus brògan gorm |
| 5 | Tha geansaidh gorm agus speuclairean dearga air. | Tha fear speuclairean dearg agus geansaidh gorm |
| 6 | Tha brògan buidhe agus geansaidh dearg oirre. | Tha tè geansaidh dearga agus brògan uaine |
| 7 | Tha geansaidh gorm agus triubhas dearg oirre. | Tha tè geansaidh gorm agus triubhas dearga |
| 8 | Tha speuclairean buidhe agus geansaidh dearg air. | Tha fear speuclairean uaine agus geansaidh dearg |
| b:PR | b4f | 1 | Tha geansaidh buidhe agus triubhas gorm oirre. | Tha geansaidh uaine anna triubhas blue oirre | The forms *anna*  and *ana* seem to indicate interference from the preposition + article *anns a’* from module two. Here it has been confused with *agus* which has different grammatical and semantic functions. |
| 2 | Tha brògan buidhe agus speuclairean dearga air. | Tha speuclairean uaine ana shoes |
| 5 | Tha geansaidh gorm agus speuclairean dearga air. |  | Subject attempted all questions in Part II and Part II – it is therefore likely that these blank answers are deliberate, with the subject recognising her inability to answer correctly and making no attempt to do so. |
| 6 | Tha brògan buidhe agus geansaidh dearg oirre. |  |
| 7 | Tha geansaidh gorm agus triubhas dearg oirre. |  |
| 8 | Tha speuclairean buidhe agus geansaidh dearg air. |  |
| b:PR | b6f | 3 | Tha triubhas dearg agus brògan uaine oirre. |  | Subject made partially successful attempts at all other questions. This is likely either a software bug of a problem using the interface. |
| c:R | c1m | 1 | Tha geansaidh buidhe agus triubhas gorm oirre. | an gheansaidh bhuinne agus trubhais gorm oirre | The form *an* seems to indicate interference from the definite article *an* from module two. Here it has been confused with *tha* which has different grammatical and semantic functions. |
| 2 | Tha brògan buidhe agus speuclairean dearga air. | an speachlearan dearg agus brogan bhuinne air |
| 3 | Tha triubhas dearg agus brògan uaine oirre. | an trubhas dearg agus brogan uinne oirre |
| 4 | Tha geansaidh uaine agus brògan gorma oirre. | an gheansaidh uinne agus brogan gorma oirre |
| 5 | Tha geansaidh gorm agus speuclairean dearga air. | an spheachlearan dearg agus gheansaidh gorm air |
| 6 | Tha brògan buidhe agus geansaidh dearg oirre. | an gheansaidh dearg agus brogan bhuinne oirre |
| 7 | Tha geansaidh gorm agus triubhas dearg oirre. | an gheansaidh gorm agus trubhais dearg oirre |
| 8 | Tha speuclairean buidhe agus geansaidh dearg air. |  | empty |
| c:R | c3f | 1 | Tha geansaidh buidhe agus triubhas gorm oirre. |  | The user made very limited attempts to answer the remaining 4 questions, answering with only a single colour and item of clothing.  However, all of the target sentences contain vocabulary that was attempted in other tasks. It is unclear why the subject left these answers blank |
| 3 | Tha triubhas dearg agus brògan uaine oirre. |  |
| 7 | Tha geansaidh gorm agus triubhas dearg oirre. |  |
| 8 | Tha speuclairean buidhe agus geansaidh dearg air. |  |
| c:R | c4f | 1 | Tha geansaidh buidhe agus triubhas gorm oirre. | Tha tè geansaich buidhe agus trubhas gorm. | Uses gender-marked language from module 2 here.  No gender-marked language whatsoever in the remaining attempted answers. It would appear the subject was aware this was wrong |
| 5 | Tha geansaidh gorm agus speuclairean dearga air. | Tha fear speuclairean dearj agus geansaidh gorm. |
| c:R | c5f | 1 | Tha geansaidh buidhe agus triubhas gorm oirre. | a geansaidh bhuidhe agus | (see below) |
| 2 | Tha brògan buidhe agus speuclairean dearga air. |  | empty |
| 3 | Tha triubhas dearg agus brògan uaine oirre. | a triubhas dearg agusuaine oirre bhrogan | The form *a* seems to indicate interference from the definite article *a’* from module two. Here it has been confused with *tha* which has different grammatical and semantic functions. |
| 4 | Tha geansaidh uaine agus brògan gorma oirre. | a geansaidh uaine agus oirre bhrogan gorm |
| 5 | Tha geansaidh gorm agus speuclairean dearga air. | a speuclairen dearg agus air geansaidh gorm |
| 6 | Tha brògan buidhe agus geansaidh dearg oirre. | a geansaidh dearg agus oirre bhrogan bhuidhe |
| 7 | Tha geansaidh gorm agus triubhas dearg oirre. | a geansaidh gorm agus oirre triubhas dearg |
| 8 | Tha speuclairean buidhe agus geansaidh dearg air. | a speuclairean bhuidhe agus air geansaidh dearg |
| d:RP | d1f | 1 | Tha geansaidh buidhe agus triubhas gorm oirre. | Tha tè geansaidh buidhe agus triubhas gorm | Subject uses gender marked language from module 2 (interference). However, she correctly answers all subsequent tasks in this section. |
| d:RP | d2m | 1 | Tha geansaidh buidhe agus triubhas gorm oirre. | Tha tè an geansaidh buidhe agus truihas gorma | Subject uses gender-marked language from module 2 consistently.  The intrusive *an* is clearly also from module 2. However it is not clear whether the question marks indicate doubt about whether the “an” is required (it is omitted in answer 3) or whether he is looking for a word meaning “wearing”, attempting to emply the present continuous pattern presented in module 2 to this situation. |
| 2 | Tha brògan buidhe agus speuclairean dearga air. | Tha fear an ? speclairean dearg agus brogan buidhe |
| 3 | Tha triubhas dearg agus brògan uaine oirre. | Tha tè ? truibhas dearg agus brogan uaine |
| 4 | Tha geansaidh uaine agus brògan gorma oirre. | Tha tè an ? geansaidh uaine agus brogan gorma |
| 5 | Tha geansaidh gorm agus speuclairean dearga air. | Tha fear an ? speclairean dearg agus geansaidh gorm |
| 6 | Tha brògan buidhe agus geansaidh dearg oirre. | Tha tè an ? geansaidh dearg agus brogàn buidhe |
| 7 | Tha geansaidh gorm agus triubhas dearg oirre. | Tha tè an ? geansaidh gorm agus truibhas dearga |
| 8 | Tha speuclairean buidhe agus geansaidh dearg air. | Tha fear an ? speclairean buidhe agus geansaidh dearg |

* + - 1. Part III

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Treatment group | subject | Q# | Answer expected | Response given | notes |
| a:P | a1f | 31 | Tha triubhas gorm agus speuclairean uaine air. |  | Subject attempted other questions with only minor errors – this is likely a problem with the user-computer interaction or a bug in the software. |
| c:R | c1m | 27 | Tha speuclairean uaine agus brògan buidhe air. | an speuchlairan uinne agus brogan bhainne air | Subject has repeated the same error as in Part I, failing to notice the correct form during the receptive input in Part II.  As noted above(\*\*\*), this error would appear to be caused by interference from forms presented in module 2. |
| 28 | Tha geansaidh dearg agus speuclairean uaine oirre. | an speuchlairan uinne agus geansaidh dearg oirre |
| 29 | Tha triubhas dearg agus brògan buidhe air. | an truibhas dearg agus brogan bhainne air |
| 30 | Tha geansaidh dearg agus brògan uaine air. | an geasaidh dearg agus brogan uinne air |
| 31 | Tha triubhas gorm agus speuclairean uaine air. | an speuchlairan uinne agus truibhas gorm air |
| 32 | Tha brògan uaine agus geansaidh gorm oirre. | an geansaidh gorm agus brogan uinne oirre |
| 33 | Tha speuclairean dearga agus geansaidh buidhe oirre. | an speuchlairean dearg agus geansaidh bhainne oirre |
| 34 | Tha speuclairean dearga agus triubhas uaine oirre. | an speuchlairan dearg agus truibhas uinne oirre |
| c:R | c3f | 32 | Tha brògan uaine agus geansaidh gorm oirre. | Gearsach gorm agus truisseau luine | Subject appears to have misread arrows in task diagram. |
| c:R | c4f | 27 | Tha speuclairean uaine agus brògan buidhe air. | Tha air brogan buidhe agus speculairen uaine. | The subject here applies the word order from module 2 to the vocabulary from module 1.  The subject self-corrects after this and uses the correct word-order in questions 31-34. |
| 28 | Tha geansaidh dearg agus speuclairean uaine oirre. | Tha oirre geansaidh dearj agus speculairen uaine. |
| 29 | Tha triubhas dearg agus brògan buidhe air. | Tha air brogan buidhe agus trubhas dearj. |
| 30 | Tha geansaidh dearg agus brògan uaine air. | Tha air brogan uaine agus geansaidh dearj. |
| c:R | c5f | 32 | Tha brògan uaine agus geansaidh gorm oirre. |  | Subject makes moderately successful attempts at all other tasks. Probably an interface or software issue. |
| d:RP | d3f | 32 | Tha brògan uaine agus geansaidh gorm oirre. |  | Subject answered all other tasks correctly – this is probably an interface or software issue. |

1. Teaching and test materials

The questions as presented in this appendix have been reformatted to reduce space and to better suit the print medium. Examples of how these are displayed to participants are shown in section 3.4.2. Note that with the exception of the pre-test, examples shown here only show representative samples – the full material can be seen online as described in Appendix K.

Text within square brackets represents an on-screen button (e.g. [ Next ]) and the symbol [\_\_\_\_\_\_\_\_\_\_] represents a text-entry field.

Images (where shown) have been shrunk for reasons of space; thus some detail may not be visible to the reader. However, it should be noted that during the test, this detail would have been visible to any participant using a desktop or laptop computer.

Note that due to the way the software tracks progress, inline instructions and presentation of new language is numbered in the same sequence as the interactive learning tasks.

* 1. Pretest

The pretest is designed to determine whether or not a candidate for participation is genuine beginner or not. It consists of two sections: general phrases and colours.

* + 1. Content

This quiz is designed to verify your level of Gaelic at the start of the course.

This quiz should be completed in one sitting, without the aid of dictionaries or notes. It should take no longer than ten minutes to complete.

Answers should be typed in the box on-screen. If you cannot type accent marks on your keyboard, use the buttons on-screen – these will insert the letter shown at the end of the box.

Further instructions are included within the exercise itself.

0 **Instructions:**

Top of Form

In the first set of questions, you will be presented with a number of sentences or questions. Respond to these as best you can, using appropriate Gaelic where possible.

If you understand the question but do not know how to answer in Gaelic, type your answer in English.

If you do not understand the prompt, leave the box blank and click *submit*.

*Please note that your answers do not need to be truthful - these questions are intended to check your ability in the language only.*  
[ Next ]

Tasks 1-7 are of the following form:

**Type an appropriate response to the Gaelic below:**

Xxxxx xxx xxx.

[à][è][ì][ò][ù]  
[\_\_\_\_\_\_\_\_\_]  
[ Submit ]

 The phrase represented by Xxxxx xxx xxx. above is from the following:

1. Madainn mhath.
2. Tapadh leat!
3. Ciamar a tha thu?
4. Dè an latha a tha e?
5. Dè an t-ainm a th’ ort?
6. Càite bheil Lunnain?
7. Càite bheil thu a’ fuireach?

8 **Instructions**

In the following questions, answer only in Gaelic. If you know the word, type it in Gaelic then click *submit*.

If you know the word but are not certain of the spelling, make your best attempt.

If you do not know the word at all, leave the box blank and click *submit*.

[ Next ]

Tasks 9-15 are of the following form:

**Name the following colour in Gaelic**

[Image: square block of colour]

[à][è][ì][ò][ù]  
[\_\_\_\_\_\_\_\_\_]  
[ Submit ]

The colour and target wordform for each of the tasks are as follows:

1. orange – orainns
2. blue – gorm
3. red – dearg
4. green – uaine
5. black – dubh
6. yellow – buidhe
7. grey – liath
8. white – geal

* 1. Module 1
     1. Learning materials

The tables below shows the questions presented in the teaching units for module 1.

The image prompt for each productive question is the correct answer for the receptive question immediately to its right; similarly the prompt text of the receptive question is the correct answer to the productive question on the right.

* + - 1. Unit 1

|  | ***PBI*** | ***CBI*** |
| --- | --- | --- |
| 0 | **New item:**  [ Repeat ] http://lingua.studio/guru/img/jumper.png geansaidh  [ Next ] | **New item:**  [ Repeat ] http://lingua.studio/guru/img/jumper.png geansaidh  [ Next ] |
| 1 | **New item:**  [ Repeat ] http://lingua.studio/guru/img/trousers.png triubhas  [ Next ] | **New item:**  [ Repeat ] http://lingua.studio/guru/img/trousers.png triubhas  [ Next ] |
| 2 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/jumper.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  geansaidh   |  |  | | --- | --- | | http://lingua.studio/guru/img/trousers.png | http://lingua.studio/guru/img/jumper.png | |
| 3 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/trousers.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  triubhas   |  |  | | --- | --- | | http://lingua.studio/guru/img/trousers.png | http://lingua.studio/guru/img/jumper.png | |
| 4 | **New item:**  [ Repeat ] http://lingua.studio/guru/img/glasses.png speuclairean  [ Next ] | **New item:**  [ Repeat ] http://lingua.studio/guru/img/glasses.png speuclairean  [ Next ] |
| 5 | **New item:**  [ Repeat ] http://lingua.studio/guru/img/shoes.png brògan  [ Next ] | **New item:**  [ Repeat ] http://lingua.studio/guru/img/shoes.png brògan  [ Next ] |
| 6 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/glasses.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  speuclairean   |  |  | | --- | --- | | http://lingua.studio/guru/img/glasses.png | http://lingua.studio/guru/img/shoes.png | |
| 7 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/shoes.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  brògan   |  |  | | --- | --- | | http://lingua.studio/guru/img/glasses.png | http://lingua.studio/guru/img/shoes.png | |
| 8 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/jumper.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  geansaidh   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/shoes.png | http://lingua.studio/guru/img/jumper.png | http://lingua.studio/guru/img/glasses.png | http://lingua.studio/guru/img/trousers.png | |
| 9 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/trousers.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  triubhas   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/shoes.png | http://lingua.studio/guru/img/jumper.png | http://lingua.studio/guru/img/glasses.png | http://lingua.studio/guru/img/trousers.png | |
| 10 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/glasses.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  speuclairean   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/shoes.png | http://lingua.studio/guru/img/jumper.png | http://lingua.studio/guru/img/glasses.png | http://lingua.studio/guru/img/trousers.png | |
| 11 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/shoes.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  brògan   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/shoes.png | http://lingua.studio/guru/img/jumper.png | http://lingua.studio/guru/img/glasses.png | http://lingua.studio/guru/img/trousers.png | |
| 12 | **New item:**  [ Repeat ] http://lingua.studio/guru/img/red.png dearg  [ Next ] | **New item:**  [ Repeat ] http://lingua.studio/guru/img/red.png dearg  [ Next ] |
| 13 | **New item:**  [ Repeat ] http://lingua.studio/guru/img/green.png uaine  [ Next ] | **New item:**  [ Repeat ] http://lingua.studio/guru/img/green.png uaine  [ Next ] |
| 14 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/red.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  dearg   |  |  | | --- | --- | | http://lingua.studio/guru/img/green.png | http://lingua.studio/guru/img/red.png | |
| 15 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/green.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  uaine   |  |  | | --- | --- | | http://lingua.studio/guru/img/green.png | http://lingua.studio/guru/img/red.png | |
| 16 | **New item:**  [ Repeat ] http://lingua.studio/guru/img/blue.png gorm  [ Next ] | **New item:**  [ Repeat ] http://lingua.studio/guru/img/blue.png gorm  [ Next ] |
| 17 | **New item:**  [ Repeat ] http://lingua.studio/guru/img/yellow.png buidhe  [ Next ] | **New item:**  [ Repeat ] http://lingua.studio/guru/img/yellow.png buidhe  [ Next ] |
| 18 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/blue.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  gorm   |  |  | | --- | --- | | http://lingua.studio/guru/img/blue.png | http://lingua.studio/guru/img/yellow.png | |
| 19 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/yellow.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  buidhe   |  |  | | --- | --- | | http://lingua.studio/guru/img/blue.png | http://lingua.studio/guru/img/yellow.png | |
| 20 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/red.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  dearg   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/blue.png | http://lingua.studio/guru/img/green.png | |
| 21 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/green.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  uaine   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/blue.png | http://lingua.studio/guru/img/green.png | |
| 22 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/blue.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  gorm   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/blue.png | http://lingua.studio/guru/img/green.png | |
| 23 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/yellow.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_]  [ Check ] | **Click on the image that matches the word.**  buidhe   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/blue.png | http://lingua.studio/guru/img/green.png | |
| 24 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/yellow.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_] [ Check ] | **Click on the image that matches the word.**  buidhe   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/trousers.png | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/shoes.png | |
| 25 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/shoes.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_] [ Check ] | **Click on the image that matches the word.**  brògan   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/trousers.png | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/shoes.png | |
|  | … | … |
| 40 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/glasses.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_] [ Check ] | **Click on the image that matches the word.**  speuclairean   |  |  |  | | --- | --- | --- | | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/shoes.png | http://lingua.studio/guru/img/glasses.png | | http://lingua.studio/guru/img/trousers.png | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/blue.png | |
| 41 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/blue.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_] [ Check ] | **Click on the image that matches the word.**  gorm   |  |  |  | | --- | --- | --- | | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/shoes.png | http://lingua.studio/guru/img/glasses.png | | http://lingua.studio/guru/img/trousers.png | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/blue.png | |
| 42 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/shoes.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_] [ Check ] | **Click on the image that matches the word.**  brògan   |  |  |  | | --- | --- | --- | | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/shoes.png | http://lingua.studio/guru/img/glasses.png | | http://lingua.studio/guru/img/trousers.png | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/blue.png | |
| 43 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/red.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_] [ Check ] | **Click on the image that matches the word.**  dearg   |  |  |  | | --- | --- | --- | | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/shoes.png | http://lingua.studio/guru/img/glasses.png | | http://lingua.studio/guru/img/trousers.png | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/blue.png | |
|  | **…** | **…** |
| 56 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/green.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_] [ Check ] | **Click on the image that matches the word.**  uaine   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/blue.png | http://lingua.studio/guru/img/shoes.png | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/trousers.png | | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/green.png | http://lingua.studio/guru/img/glasses.png | http://lingua.studio/guru/img/jumper.png | |
| 57 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/yellow.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_] [ Check ] | **Click on the image that matches the word.**  buidhe   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/blue.png | http://lingua.studio/guru/img/shoes.png | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/trousers.png | | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/green.png | http://lingua.studio/guru/img/glasses.png | http://lingua.studio/guru/img/jumper.png | |
| 58 | **Type a word or phrase to describe the image.**  http://lingua.studio/guru/img/glasses.png [à][è][ì][ò][ù] [\_\_\_\_\_\_\_\_\_] [ Check ] | **Click on the image that matches the word.**  speuclairean   |  |  |  |  | | --- | --- | --- | --- | | http://lingua.studio/guru/img/trousers.png | http://lingua.studio/guru/img/yellow.png | http://lingua.studio/guru/img/blue.png | http://lingua.studio/guru/img/glasses.png | | http://lingua.studio/guru/img/red.png | http://lingua.studio/guru/img/jumper.png | http://lingua.studio/guru/img/shoes.png | http://lingua.studio/guru/img/green.png | |
| … | **…** | **…** |

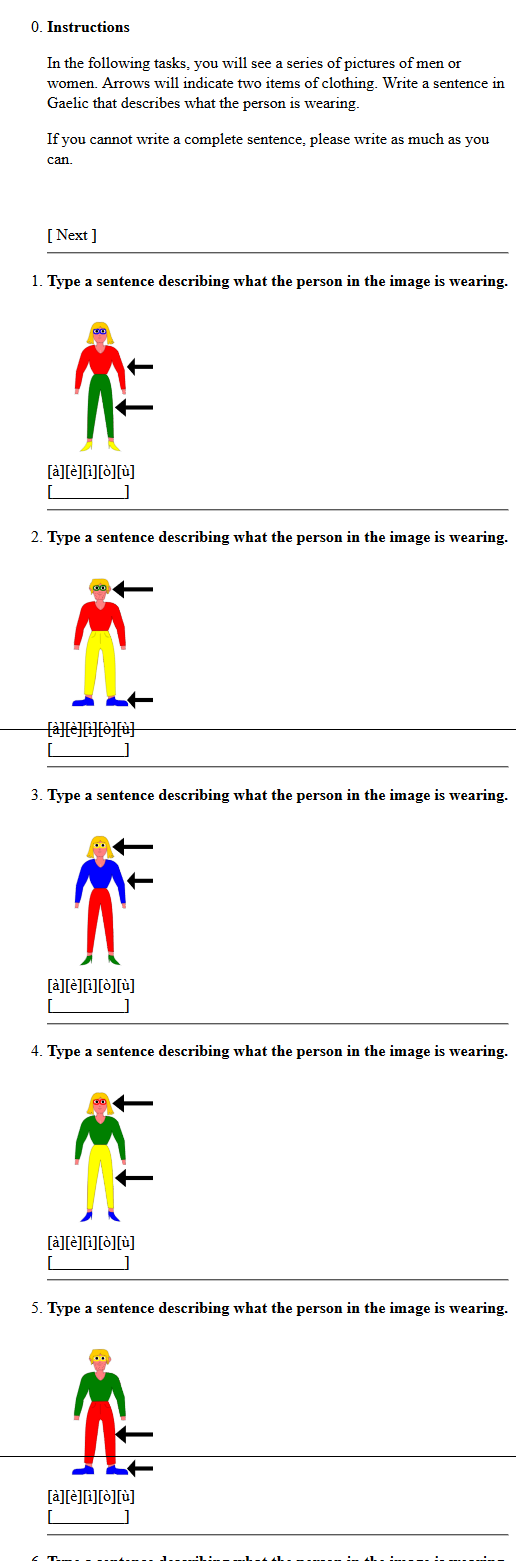
* 1. Post-tests

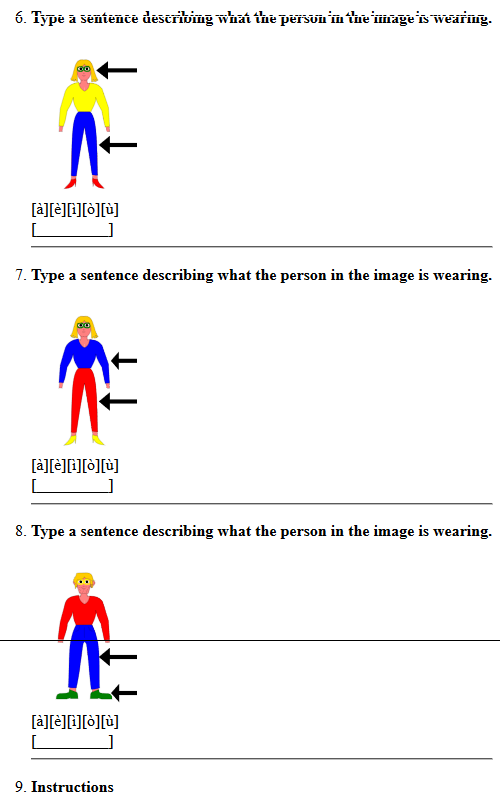
The post-test displayed below is the immediate post-test.

Note that the immediate and delayed post-tests are structurally identical, and differ only by the following alternations of vocabulary that preserve grammatical distinctions (plural nouns are swapped, singular nouns are swapped; adjectives that take plural marking are exchanged with each other, adjective that do not change are swapped with each other).

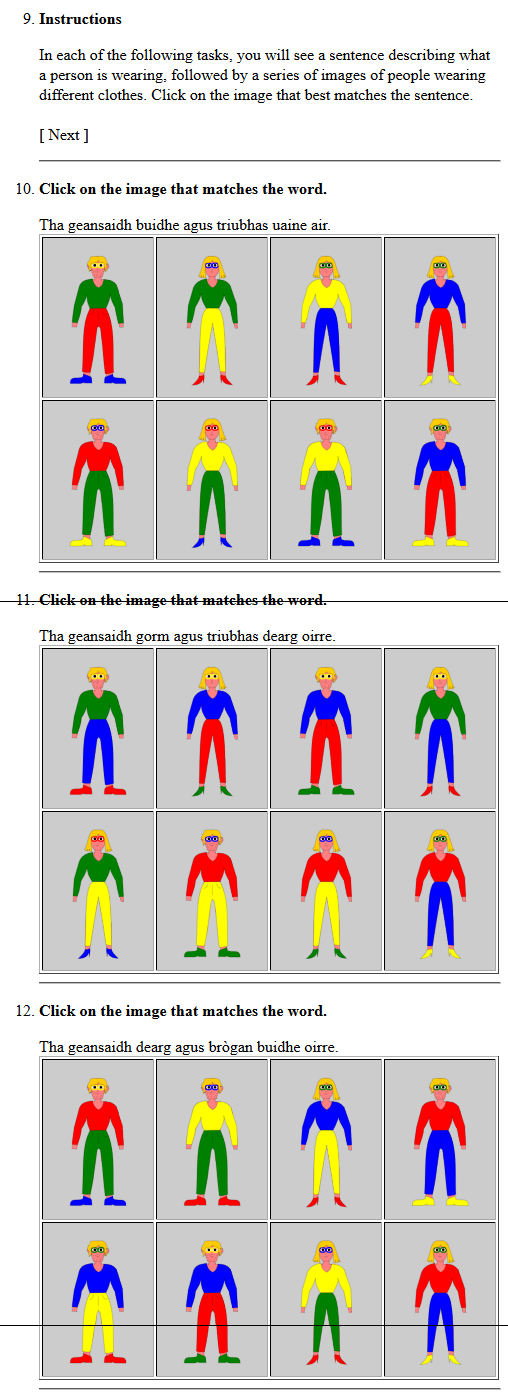
Shoes ↔ spectacles jumper ↔ trouser[s] uaine ↔ buidhe gorm ↔ dearg

* + 1. Part I: productive skills

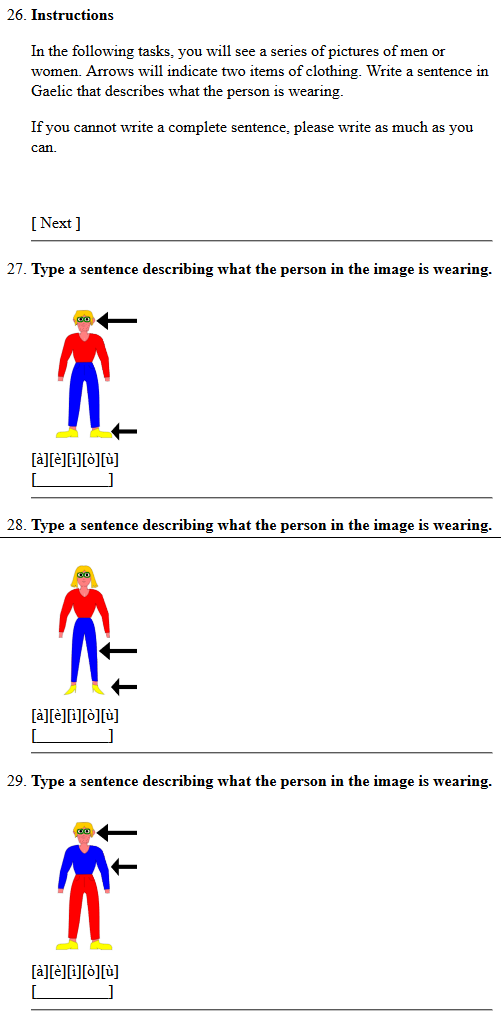




* + 1. Part II: receptive skills



* + 1. Part III: productive skills



1. Copyright information

The majority of images used in this study are the work of the author.

The following images were obtained from internet sources under the Creative Commons Attribution Share-Alike license (CC-BY-SA 3.0 - <https://creativecommons.org/licenses/by-sa/3.0/deed.en>) and are listed with their creators on the website and in this document as required by the terms of the license.

**Image:** <https://commons.wikimedia.org/wiki/Italia#/media/File:Italy_looking_like_the_flag.svg>  
**Creator**: mnemoc

**Image:** <https://commons.wikimedia.org/wiki/File:CARTE_DRAPEAU_FRANCE.svg>  
**Creator:** MG-commonswiki

**Image:** <https://commons.wikimedia.org/wiki/Deutschland#/media/File:Germany_location_map.svg>  
**Creator:** NordNordWest

**Image:** <https://commons.wikimedia.org/wiki/Espa%C3%B1a#/media/File:DO_Valdepe%C3%B1as_location.svg>  
**Creator:** Té y kriptonita

The audio files used in this study were generated using the *Ceitidh* voice by Cereproc ([www.cereproc.com](http://www.cereproc.com)) licensed via CALL Scotland and the University of Edinburgh for academic use (<http://www.thescottishvoice.org.uk/licence/>). In order to comply with the terms of the license, these files will be deleted when marking is complete and a final grade confirmed.

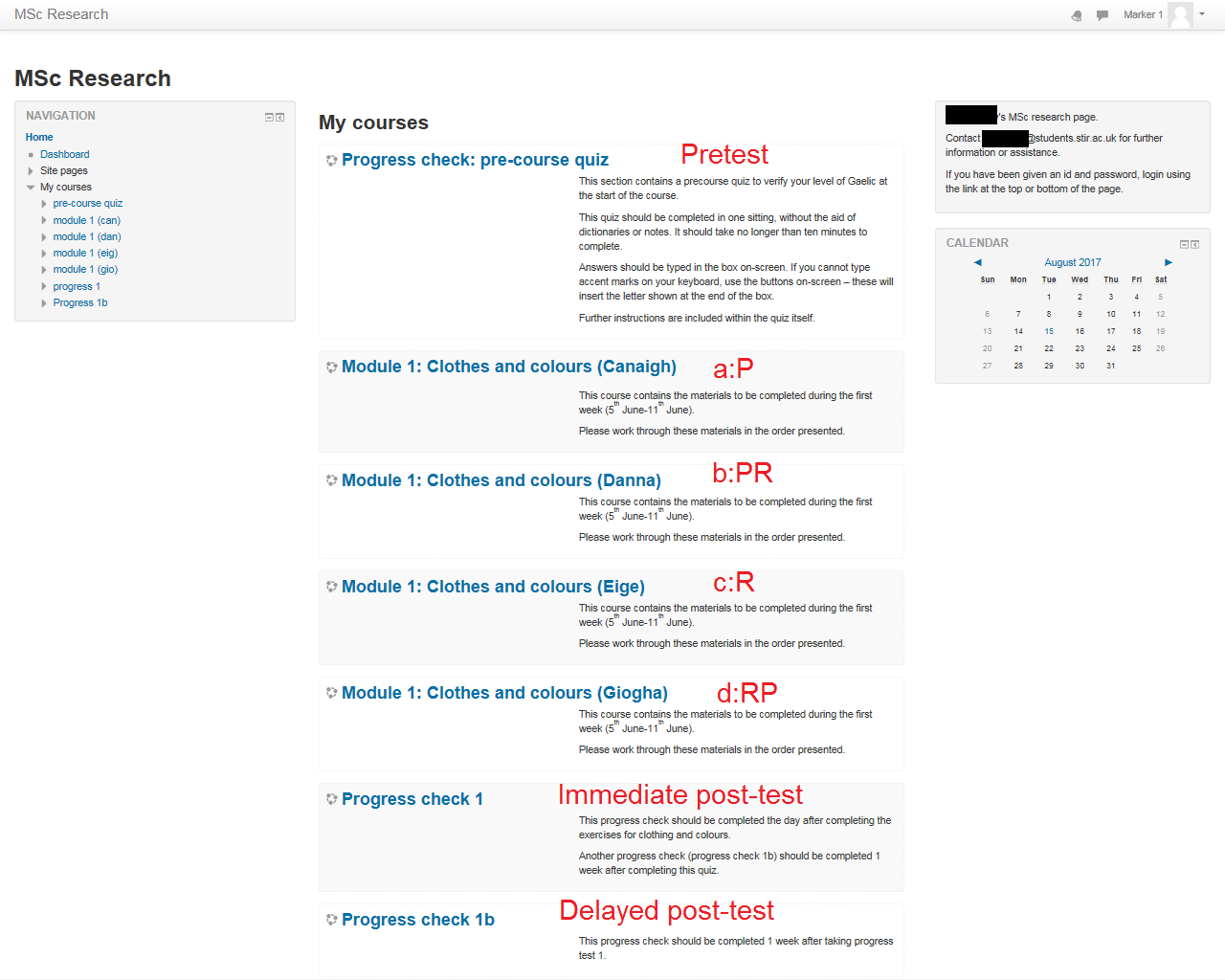
1. Access instructions

The materials used in this study are available at <https://lingua.guru/login/> . Note that access to this site is controlled by secure user accounts, and a password is needed for each user. Individual accounts have been configured for 3 markers and will be active during the marking period only, as the materials will be taken off-line after marking is complete (due to the copyright restrictions noted below).

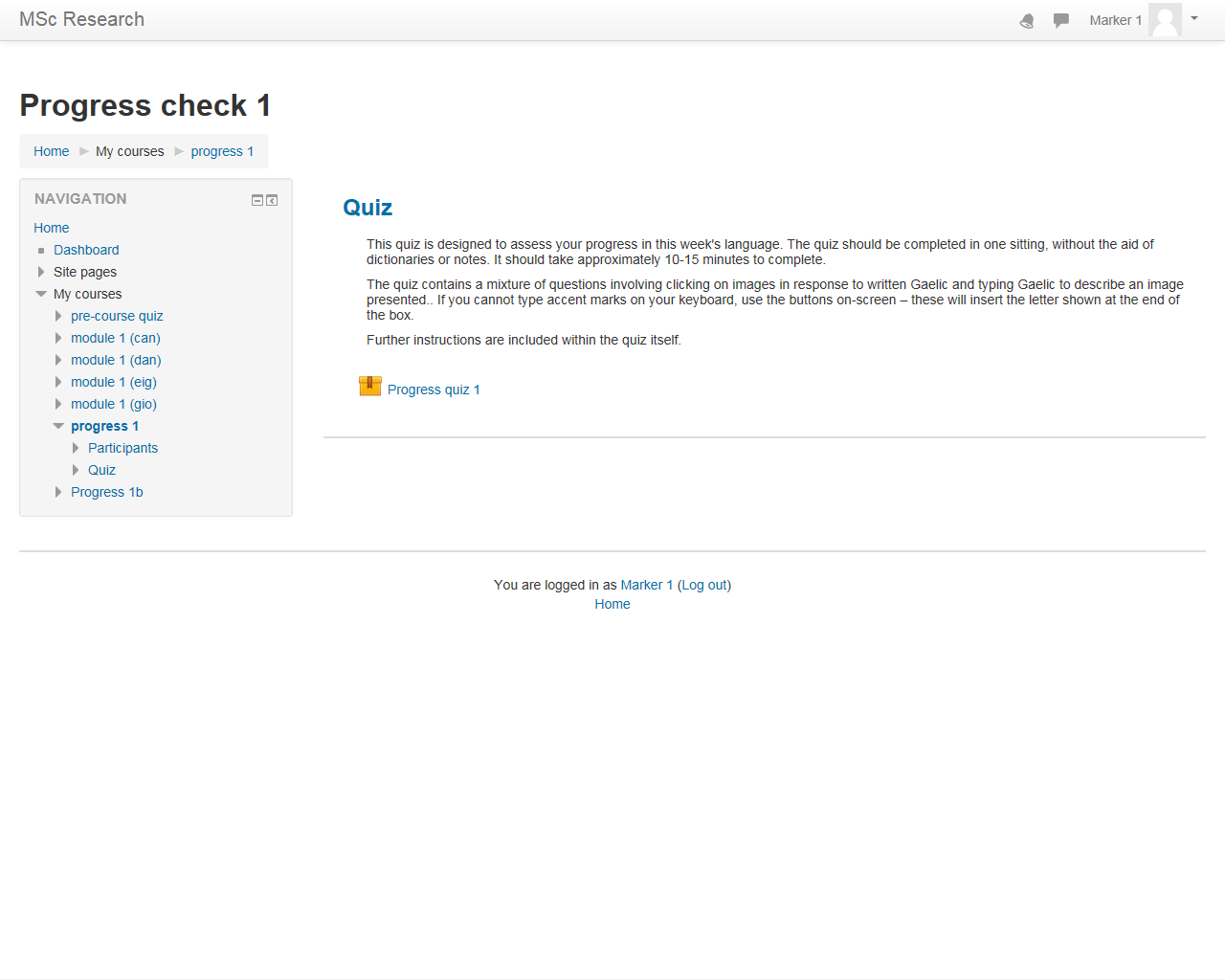
The usernames and passwords are as follows:

|  |  |  |
| --- | --- | --- |
| Marker | Username | Password |
| First marker | marker1 | TESOLdissert2017 |
| Second marker | marker2 | TESOLdissert2017 |
| Third marker | marker3 | TESOLdissert2017 |

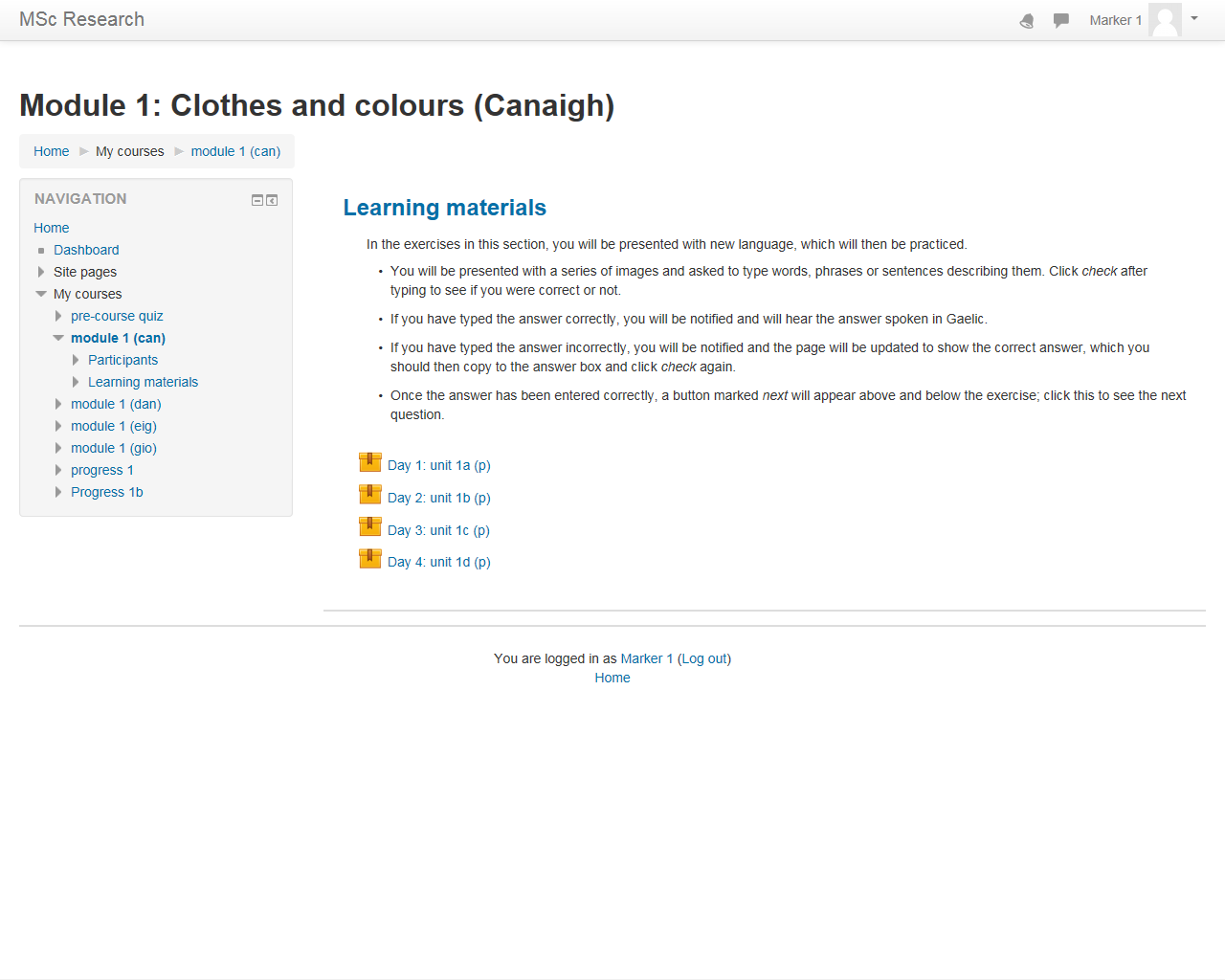
The first screen you will see after login will be as shown overleaf. Annotations have been added in red text to highlight the purpose of each section. Note that study participants were only presented with one of the *Module 1: Clothes and colours* section based on their treatment group. The codenames given were assigned arbitrarily from a list of Gaelic island names in order to prevent any negative associations with the designations of *a, b, c* and *d* as used in this paper.

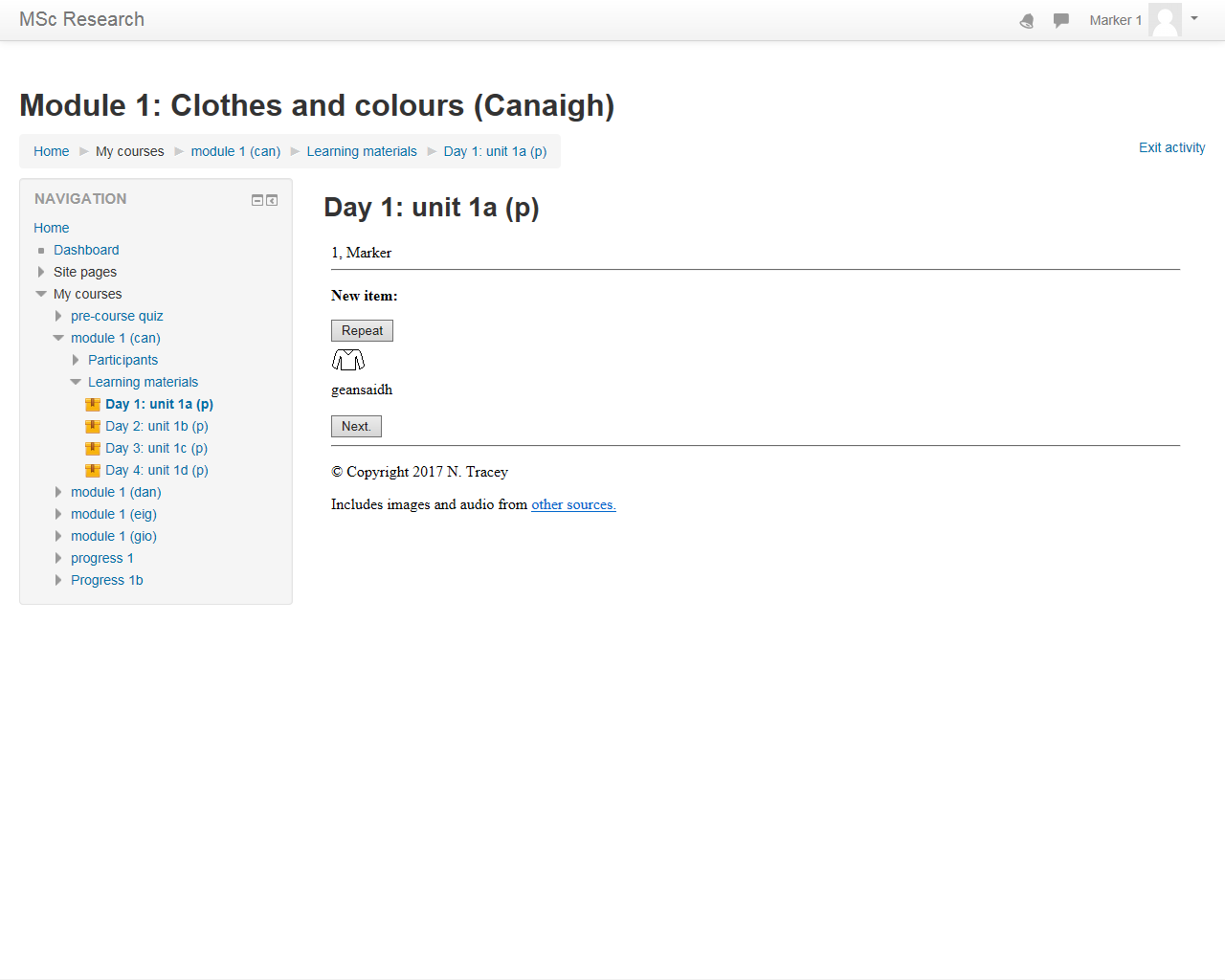


Clicking on one of the *Progress check* (pre- or post-test) sections will bring up a screen with brief task instructions, followed by a link to the test itself:



Clicking on one of the *Module 1* (teaching intervention) sections will bring up a screen with brief instructions, followed by a list of links to the four units within the module, which should be completed in the order presented:



Clicking on links to a test or learning unit will bring up the task in a webpage like the following:  


If you navigate away from an activity without completing it, the next time you access it, you will begin on the last task presented.

All other instructions are included within the web application.