Dadansoddi ar gyfer Polisi
YMCHWIL GYMDEITHASOL Y LLYWODRAETH
GSR
GOVERNMENT SOCIAL RESEARCH
Analysis for Policy

Llywodraeth Cymru
Welsh Government

SOCIAL RESEARCH NUMBER:
31/2018
PUBLICATION DATE:
07/06/2018

Rapid Evidence Assessment: Effective Second Language Teaching Approaches and Methods

Mae'r ddogfen yma hefyd ar gael yn Gymraeg.

This document is also available in Welsh.

 Digital ISBN 978-1-78937-321-9

Rapid Evidence Assessment: Effective Second Language Teaching Approaches and Methods

Authors: Fitzpatrick, T., Morris, S., Clark, T., Mitchell, R., Needs, J., Tanguay, E. and Tovey, B.

Full Research Report: Fitzpatrick, T., Morris, S., Clark, T., Mitchell, R., Needs, J., Tanguay, E. and Tovey, B. (2018). *Rapid Evidence Assessment: Effective Second Language Teaching Approaches and Methods*. Cardiff: Welsh Government, GSR report number 31/2018.

Available at: https://gov.wales/statistics-and-research/assessment-effective-second-language-teaching-approaches-methods/?lang=en

For further information please contact:
Catrin Redknap
Social Research and Information Division
Knowledge and Analytical Services
Welsh Government
Cathays Park
Cardiff
CF10 3NQ
Tel: 0300 025 5720
Email: catrin.redknap@gov.wales

Table of contents

List of Tables

List of Figures

Glossary

AoLE	Area of Learning and Experience
ASR	Automatic Speech Recognition
BES	Best Evidence Synthesis
CALL	Computer Assisted Language Learning
CBI	Content Based Instruction
CEFR	Common European Framework of Reference for Languages
CLIL	Content and Language Integrated Learning
CLT	Communicative Language Teaching
DDL	Data Driven Learning
DELTA	Diploma for Teaching English to Speakers of Other Languages
DGBL	Digital Game Based Learning
EFL	English as a Foreign Language
EI	Elicited Imitation
ELL	Early Language Learning
EPPI	Evidence for Policy and Practice Information
ERIC	Educational Resources Information Centre
ESS	English as a School Subject
FFI	Form Focused Instruction
FonF	Focus on Form (attention to a vocabulary item during a communicative activity)
FonFs	Focus on Forms - learning focus is the vocabulary items
GJT	Grammaticality Judgement Test
GSR	Government Social Research
GTM	Grammar Translation Method
ICT	Information and Communication Technology
IELTS	International English Language Testing System
ILD	Individual Learner Difference

KS	Key Stage
L1	First language
L2	Second language
LEA	Local Education Authority
LLBA	Linguistics and Language Behaviour Abstracts
NLW	National Library of Wales
NNS	Non-native speaker
PICO	Population, Intervention, Comparison, Outcomes
PPP	Presentation, Practice and Production
REA	Rapid Evidence Assessment
SI	Strategy Instruction
SRSD	Self-Regulated Strategy Development
TBL	Task-Based Learning
WG	Welsh Government

1. Introduction

1.1 The purpose of the Rapid Evidence Assessment (REA) reported in this document is to inform the Welsh Government's planning and delivery of Welsh language provision for learners aged 3-16 years, as it undertakes reform of curriculum and assessment arrangements in Wales. The REA was commissioned in the light of two significant national policy drivers: the development of a new school curriculum for Wales, initiated by the *Successful Futures* report (Donaldson, 2015), and *Cymraeg 2050: a million Welsh speakers* (Welsh Government, 2017). Together, these necessitate a significant review of language education in Wales, and create potential for a paradigm shift in approaches to language teaching and learning. This context provides a rare opportunity, from a research perspective, to focus on the effectiveness of language education, in conditions that will tolerate, even encourage, a transformational change to teaching approaches and methods. The REA reported here, comprising an assessment of high quality, relevant research yielded by a principled and extensive literature search, represents a state-of-the-art research platform for policy development. While its relevance to Wales is explicit, this constitutes a significant contribution to teaching policy and practice in all non-dominant target language contexts.

Aims and rationale for the Rapid Evidence Assessment (REA)

1.2 In 2017 Welsh Government (WG) identified a key area in which a rapid evidence assessment (REA - 'the project') should be undertaken: the effectiveness of approaches and methods in second language teaching. The aim of the project was therefore to assess research on language teaching practices and interventions used in international and national contexts parallel or comparable to Wales, in order to inform WG planning for delivery of Welsh language provision for learners aged 3-16. The project will contribute to WG's awareness and appreciation of approaches/methods that are likely to give learners effective functional skills in the Welsh language by the age of 16. The specific research question addressed in this report is:

What teaching approaches and methods are effective in developing young learners' second language competence, according to high quality empirical evidence?

Policy background and development

1.3 There are three principal WG policies which provide the context of the project:

- *Cymraeg 2050 – A million Welsh speakers*.[1] One of the sea-changes required to enable WG to achieve its goal is to transform how learners are taught Welsh, with the aim that by 2050 at least 70 per cent report that they can speak Welsh by the time they leave school. This will require an increase in the number of primary and secondary teachers who can teach Welsh as a subject. In secondary schools this will be from 500 in 2017 to 900 by 2031 and to 1,200 by 2050. Improving how Welsh is taught in English-medium schools is noted as a priority.
- *A Curriculum for Wales – A Curriculum for Life*.[2] Six Areas of Learning and Experience (AoLEs) are identified in the new curriculum for Wales, with progression within each AoLE based on continua of learning. The Welsh language (a compulsory element of the curriculum up to age 16 for all learners) falls within the AoLE for Languages, Literacy and Communication. The emphasis on learning Welsh is defined as developing the language as a means of communication with progress towards "transactional competence" (Donaldson, 2015 p. 60) by age 16.
- *Education in Wales: Our National Mission*.[3] The development of a transformational approach to the learning, teaching and assessment of Welsh is a stated aim of this action plan covering the period up until 2021. The aim is to ensure that all learners will be able to use Welsh when they leave school. One of the priorities for professional learning here is to teach Welsh as a living language and to ensure that greater numbers are able to work effectively in Welsh-medium and/or bilingual settings. The vision to enable all learners to develop their Welsh language skills and use Welsh confidently in everyday life is reiterated in *Welsh in Education: Action Plan 2017-2021*.[4] Particularly relevant to this REA and to enabling WG to realise its vision for children and young people are the development of a new curriculum to inspire learners to learn and use the language, and the need to enhance motivation through accurate measurement of progress, together with ensuring that teaching and learning are based on teachers' understanding of the most effective language acquisition pedagogy.

[1] http://gov.wales/docs/dcells/publications/170711-welsh-language-strategy-eng.pdf
[2] http://gov.wales/docs/dcells/publications/151021-a-curriculum-for-wales-a-curriculum-for-life-en.pdf
[3] http://gov.wales/docs/dcells/publications/170926-education-in-wales-en.pdf
[4] http://gov.wales/docs/dcells/publications/171212-welsh-in-education-action-plan-2017-21-en.pdf

Project objectives

1.4 The aim of this REA is to identify, assess the quality of, scrutinise and report research relating to effective second language teaching approaches and methods for use with 3-16 year olds, taking a principled, systematic and transparent approach to literature search, screening and synthesis. The main objectives of the project are:

- **Objective A** – to conduct a search and scrutiny of published research within the parameters described below (**Section 2. Methodology** of this report);
- **Objective B** – based on what is found through Objective A, to produce a state of the art synthesis of the most relevant and highest quality research, organised thematically (**Section 3. Findings**);
- **Objective C** – to assess how the research findings map onto Welsh language pedagogy (3-16 year olds), particularly in the context of Education in Wales: our national mission (Welsh Government 2017) (**Summary of findings applied to the Welsh context**);
- **Objective D** – to indicate related research areas which, while outside the scope of this report, will be key to achieving the policy aims outlined in 1.2 and 1.3 above (**Section 4. Notes on excluded literature and future research imperatives**).

1.5 The project parameters require that research will only be regarded as relevant and suitable for inclusion when it:

- is directly or indirectly relevant to language learners aged 3-16 years old;
- is directly or indirectly relevant to the context of teaching Welsh in Wales (for example, it has relevance to the teaching of non-dominant target languages);
- focuses on “approach” and/or “method”; research on theoretical models, or teaching techniques/activities are only included if they are relevant to an approach or method;
- addresses deliberate, within class, teaching of second languages that are human, written/spoken, languages.

Investigative framework – definition and interpretation of terms

1.6 The framework for the investigation was based on parameters using PICO (Richardson, Wilson, Nishikawa, & Hayward, 1995). While originally designed for clinical contexts, the Population > Intervention > Comparison > Outcomes

framework, and the overarching principle of evidence-based practice, makes it highly appropriate for the definition of terms in our current investigative framework:

- **Population:** learners of Welsh as a second language aged 3-16. 'Language learner' indicates a deliberate intention on the part of the learner and/or teacher (thereby excluding immersion contexts). 'Language' here is human, written/spoken language (therefore not including Sign Languages or computer languages, for example). 'Welsh second language' is used with the specific meaning applied to learners who are following a curriculum solely or predominantly delivered through the medium of English. This differentiates them from second language speakers of Welsh who acquire the language through a predominantly Welsh-medium curriculum and/or within an immersion setting. Welsh L2 does not, therefore, necessarily reflect learners' linguistic background and exposure to Welsh prior to education, nor does it include those learners for whom Welsh may be a second (or other) language who are in Welsh-medium or immersion provision.
- **Intervention / Exposure:** classroom-based or -related teaching approaches and methods. Theories of language learning were beyond the remit of the study, as were interventions beyond the reach of the classroom. 'Approach' and 'method' are terms whose definition has changed over time and is contested (Hall, 2011). In this project, 'approach' is taken as a set of values, principles, and beliefs about factors that drive learning, and 'method' is taken as the systematic engagement of learners with language (definitions informed by e.g. Hall, 2011; Chapter 22 of Richards & Rodgers, 2014). Techniques and activities were included, only where they exemplified methods and approaches.
- **Comparisons:** Where the literature permits, comparisons between methods/approaches were made on the basis of 'effectiveness'. The interpretation of 'effectiveness' which informs this investigation is the deployment of a teaching method or approach in a classroom that enhances learner performance, in a measurable and sustained way, in one or more of the competences identified in the methodology.
- **Outcome:** language competence, particularly (given the emphasis in *Successful Futures*) with reference to 'transactional competence', and interpretations of the components of this.

Project Team

1.7 The REA took place between November 2017 and March 2018. The core project team of two senior researchers and four research assistants was based at Swansea University. An advisory team fed into the methodological process at critical stages, and most influentially in developing the investigative framework and contributing to the supplementary searches (see **Section 2. Methodology**). The team was constituted in order to maximise the reach of expertise in areas relevant to the project objectives and methodology. Half the team members have previous experience as classroom language teachers (of Welsh, Irish, English, French, Catalan and Spanish), and team members' research and scholarly activity covers areas including: Second language acquisition; Language education policy; Language pedagogy; Content and Language Integrated Learning (CLIL); Computer Assisted Language Learning (CALL); Motivation studies; Language assessment; Teacher education; Inter-relationship between child second language/foreign language learning, vocabulary and literacy development; The effects of age in second language learning; Formulaic language acquisition; Language learning in minority language contexts; Language acquisition and bilingualism amongst children and adults. A full list of the project team members is in **Annex A: Project team members, roles and affiliations**.

Approach and structure of the report

1.8 The potential scope of the report is vast, given the quantity of research literature published in the field of language learning and teaching. As an illustration of this, a search on the Scopus Database titles, keywords and abstracts for "[approaches OR methods] + [language] + [learning OR teaching]" returns over 43,000 items. In order, then, to adhere to the principled, selective approach demanded by an REA, to target the literature most acutely relevant to the REA, and to meet the 'rapidity' criteria of the REA, a sophisticated search mechanism was required (Objective A in **Project objectives** above). **Section 2. Methodology** of the report sets out the method we devised for this, and the quality assurance measures and contingencies we put in place. Findings then had to be synthesised and presented in a way that was fit for purpose (Objectives B and C above), and here the challenge was threefold:

i) A framework driven by conventional taxonomies of method and approach, established a priori, would risk excluding research on a) methods that do not sit obviously under the traditional (often unhelpfully broad) banners of 'task-based learning', 'grammar translation', and so on, and b) practice emerging from the "principled eclecticism" and "principled pragmatism" (see e.g. Kumaravadivelu 1994) of the post-method era.

ii) Effective language teaching and learning is by no means solely generated by planned, teacher-controlled, classroom-based endeavours; a vast array of additional influences and variables play critical roles in language learning, and these must be acknowledged.

iii) "Transactional competence" (Donaldson, 2015) is presented as a measure of effectiveness in the policy context, but the term itself is under-defined (generating only one 'hit' in Scopus, for example), and its components must be identified.

1.9 **Section 2. Methodology** of the report sets out the methodology used to address these challenges, and presents the rationale for organising the project findings (**Section 3. Findings**) according to a thematic approach based on language competences. Finally, in connection with point (ii) in the paragraph above and Objective D in **Project objectives**, **Section 4. Notes on excluded literature and future research imperatives** of the report focuses on areas of research which, though excluded from the main business of this project, are critical considerations in the formation of policy for Welsh language education in Wales.

2. Methodology

Defining the research question

2.1 The aim of the project was to provide a balanced assessment of research literature evidence in order to help inform the Welsh Government's planning and delivery of Welsh language provision for learners aged 3-16. In order to do this in such a way as to maximise both fitness-for-purpose and feasibility, precise framing of a research question and establishment of research parameters was essential. The social and psychological complexity of language learning means the breadth and quantity of research literature with relevance to the topic is vast. This was evident in the fact that initial scoping of policy frameworks (see **Policy background and development**) and consultation within the project team, generated a preliminary set of twelve broad areas of relevance to the WG's policy ambition. These were:

i) Effectiveness of approaches/methods when applied to the young language learner context;
ii) Immersion and CLIL (Content and language Integrated Learning);
iii) Assessment of learning;
iv) Quality and intensity of learners' exposure to language;
v) Age and cognitive development;
vi) Practitioner skills and training;
vii) The processes by which 'transactional competence' develops;
viii) Development of bi- and multi-literacy;
ix) Cognitive and social advantages of language learning and bilingualism;
x) Motivation and attitude;
xi) Role of technology in language learning;
xii) Individual learner differences.

2.2 That these areas are of core relevance to the policy ambition context of the project, and relate directly to considerations of teaching methods and approaches is undeniable. However, trial database searches encompassing all these themes yielded hundreds of thousands of results, indicating that this approach would risk generating wide-ranging and unfocused findings, and was in any case not feasible within the parameters of a REA. It was therefore decided to constrain the scope of the research question to a tight interpretation of approach and method (see ***Phase 2a – Screening***), while acknowledging that consideration of iii-xii above would also

be essential to informing policy development (see **Section 4. Notes on excluded literature and future research imperatives**).

2.3 This decision narrowed our focus to i) effectiveness of approaches/methods when applied to the young language learner context; and ii) immersion and CLIL (Content and language Integrated Learning). On closer scrutiny of ii), it was decided that while CLIL's explicit attention to language development (as well as academic content) render it relevant to this REA, the theoretical foundation of 'Immersion' (that exposure without explicit language-focused instruction will lead to acquisition) is not compatible with the "deliberate, within class, teaching of second languages" parameter established in **Project objectives** above, and is therefore not within the scope of the REA.

2.4 The question addressed by this review is, then: *What teaching approaches and methods are effective in developing young learners' second language competence, according to high quality empirical evidence?*

Methodological approach

2.5 To answer this research question, a rigorous approach to literature search, screening, scrutiny and synthesis was required. The expertise and experience of the project team supplemented this, enabling the authors to tap into state-of-the-art knowledge across areas relevant to our project focus, and to make contingency for crucial literature not captured by the database search (for example because article titles do not transparently reflect content, or items fall outside search engine capture zones, or are incompatible with our search terms).

2.6 The following sub-sections outline the objective investigative framework devised for the project. This was based on the parameters outlined in **Project objectives** of the introduction, and was informed by the REA considerations in Table 2.1, the GSR (Government Social Research) *Rapid Evidence Assessment toolkit* and EPPI (Evidence for Policy and Practice Information) Centre guidance.[5] The investigation proceeded in four phases:

[5] Government Social Research Service and Evidence for Policy and Practice Information Centre (n.d.). *Rapid Evidence Assessment toolkit*. Retrieved from http://webarchive.nationalarchives.gov.uk/20140402164155/http://www.civilservice.gov.uk/networks/gsr/resources-and-guidance/rapid-evidence-assessment

- Phase 1: Literature search
- Phase 2: Screening, selection and categorisation
- Phase 3: Detailed data extraction and weighing of evidence
- Phase 4: Clustering items and synthesising key findings

2.7 Methods used in each phase are detailed in sub-sections **Phase 1 – Search** to **Phase 4 – Clustering items and synthesising key findings**. A member of the project team was assigned the role of Quality Assurer, and scrutinised the methodological approach of each phase; quality assurance measures were also embedded in the methodology processes. Considerations driving the REA methods, and the means by which they were addressed, are presented in Table 2.1.

Table 2.1: Methods for addressing key REA considerations

Methodological consideration	Addressed by…
maximise **completeness** of REA	• conduct initial search across range of databases • define keywords for searches and trial them robustly • use project team's combined expertise to identify items supplementing initial search results
maximise **rapidity and achievability** of REA	• adopt exclusion criteria to limit search to recent research (published post 2000) in English or Welsh • secure support from research librarians and subject specialists • limit grey literature review to items signalled by expert team members and Welsh Government steering group, that meet the inclusion criteria set out in the section on supplementary manual searches • fast-track review papers (e.g. systematic reviews/meta-analyses/state-of-art summaries)
maximise **relevance of focus of** included studies	• attend closely to relevance of focus in key word selection • screen abstracts by relevance of study focus • record study focus and assessment in data extraction form
maximise **relevance of study design** of included studies	• screen abstracts by relevance of topic focus and research design • record study design assessment in data extraction form
assess **quality** of included studies	• use a matrix approach in data extraction process, considering study design, validity, and reliability (informed by Petticrew & Roberts, 2006).
avoid **bias** in REA	• define a priori search and scrutiny principles • search across range of databases • adhere to data extraction protocols • use a provisional list of thematic sections as indicative, and adapt/supplement these according to themes emerging from the literature
maximise **comprehensibility** of REA report	• organise syntheses by thematic sections, based on a consideration of competences and skills • include summary of 'key considerations'

Phase 1 – Search

Database searches

2.8 **Databases used:** The databases used in the search phase were:

- Scopus: a global abstract and citation database of peer-reviewed research titles in the fields of science, technology, medicine, social sciences and arts and humanities
- LLBA (Linguistics and Language Behaviour Abstracts – abstracts and indexes the international literature in linguistics and related disciplines in the language sciences)
- Web of Science Service for UK Education
- ERIC (Educational Resources Information Centre – major contributors include journal publishers, grey literature and book sources and individuals through an on-line submission system)
- National Library of Wales Catalogue (NLW is a legal deposit library with a particular interest in Welsh and Welsh language material)
- Swansea University Library Catalogue (Access to all the deposits in the University libraries including Welsh language material)

2.9 These databases/platforms were selected from an initial larger set, in collaboration with a subject librarian. Considerations in selection included coverage (size and scope of database), reach of search (the most efficient was for search terms to be sought in title, abstract and keywords), accessibility of abstracts for rapid scrutiny. The last two catalogues listed above were included to facilitate Welsh language searches.

2.10 The search and inclusion criteria below are presented according to search protocols on Scopus. The search terms were adapted to fit other database search engines as appropriate.

2.11 **Search terms (English language):** With close attention to the project parameters (**Project objectives**), the investigative framework (**Investigative framework – definition and interpretation of terms**) and a tight interpretation of the research question (**Defining the research question**), search terms were identified and trialled. Trialling attended to feasibility (number of results) and capture (a test set of expected items, drawn from the recommended literature from expert members of the project team, was used to ensure that relevant literature was captured). The set

of search terms emerging from this process consisted of five sets of “relevance” keywords linked by “AND” operators, yielding results relevant on the five counts: population, context, subject domain, activity and ‘effectiveness’ variable. The terms were:

For relevance of target population - (child* OR pupil* OR student* OR new speaker*)

AND

For relevance of context - (elementary OR secondary OR high school* OR young OR primary OR early years)

AND

For relevance of subject domain - (second language* OR foreign language* OR modern language* OR heritage language* OR minority language* OR regional language* OR L2)

AND

For relevance of activity - (teach* OR learn* OR instruct* OR pedagog* OR acqui*)

AND

For relevance of target variable - (succe* OR achiev* OR improv* OR attain* OR

effect* OR gain* OR increas* OR grow*)

2.12 **Search terms (Welsh language):** In order to capture Welsh language outputs, a set of Welsh search terms was created for use with the National Library of Wales and Swansea University Catalogues. Following trialling (as above), the terms were confirmed as:

For relevance of target population - (plant OR plentyn OR blant OR blentyn OR phlant OR phlentyn OR mhlant OR mhlentyn)

AND

For relevance of subject domain - (iaith OR ieith*)

AND

For relevance of activity - (dysg* OR ddysg* OR nysg* OR addysg*)

2.13 **Inclusion criteria:** To maximise the relevance and quality of search results, the following inclusion criteria were applied:

- Date of publication = post-2000. A non-date-limited key word search using Scopus revealed a sharp increase between 2000 and 2001 in the number of outputs relating to teaching young language learners. 2001 was the European Year of Languages, and a significant year in language teaching pedagogy, with initiatives that changed and reinvigorated language teaching research: for example, Cameron's Teaching Languages to Young Learners (2001) was first published; the Common European Framework of Reference for Languages was launched (Council of Europe, 2001); the National Foundation for Educational Research published Teaching Modern Languages: Policy and Practice in England, Wales and Northern Ireland (Boyd, 2001).
- Subject domain: "SOCIAL SCIENCES", "ARTS & HUMANITIES"
- Publication type: "ARTICLE", "BOOK CHAPTER", "ARTICLE IN PRESS", "REVIEW", "BOOK"
- Language: "ENGLISH", "WELSH"
- "Peer reviewed": some databases (Scopus) only include peer reviewed outputs; others (ERIC, LLBA) have this as an optional filter.

Supplementary searches

2.14 To mitigate the risk that relevant items would be missed through the database searches (for example, because they are in non-electronic or -indexed journals, or do not feature our search terms), and because books and book chapters do not have searchable abstracts, a supplementary search was conducted. This would also enable the capture of grey literature, and any highly relevant research items lying outside the inclusion zones set for the electronic searches. The supplementary searches were as follows:

i. Expert team members were asked to provide items that perform strongly against the criteria of relevance (as set out in Table 2.2 below), but which do not necessarily meet the inclusion criteria in ***Database searches*** above (i.e. the list may include items published before 2001; in a language other than English or Welsh, that do not have a specific focus on (but can inform) teaching of young learners, etc).

ii. Expert team members were asked to identify grey literature such as reports commissioned by public bodies including key reviews focusing on similar contexts and research imperatives; research project reports; unpublished dissertations; reading lists from internationally recognised, practitioner-focused training programmes. The inclusion criteria for grey literature were that it must have relevance to the REA, and carry some assurance of quality, such as having been produced by/for public bodies or research councils whose funding protocols demand quality assurance scrutiny, or by accredited professional bodies, or (for academic research) have met the criteria for PhD award.

iii. Special editions/sections of journals on relevant topics were scrutinised.

iv. A library catalogue search was conducted, using the search term "Young Language Learners" and restricted to books.

2.15 In connection with (i) and (ii) above, instructions to expert team members were to submit a list of publications which the team member considered to be of key relevance/importance to this project. With regard to grey literature, team members were asked to focus on items they considered most useful according to their own knowledge and experience. They were asked to record their list of publications on a spreadsheet, together with information to support the sorting, selection and categorisation of items for further scrutiny and synthesis. This information comprised: author(s); year of publication; full APA reference; publication type; weblink (if available); "which language competences and/or skills does this relate to?"; "what is the main focus in terms of teaching approach/method?"; "what is/are the key finding(s)?".

2.16 The decision was made not to give a pre-determined list for either the competence/skill or the approach/method question, in order not to risk excluding relevant items, and to enable categories identified by expert panel members to be triangulated with, and to inform, the indicative categories used in **Phase 2 – Screening, selection and categorisation** below.

Quality assurance at Phase 1

2.17 Quality assurance measures taken during Phase 1 included:

- Search terms were tested and scrutinised with research librarian (using a set of test articles) and with Quality Assurer, in order to ensure fitness-for-purpose and appropriate application.
- Database searches were conducted by two members of the team, and results compared.

2.18 In Phase 2 of the project, outlined below, search results were manually screened for relevance.

Phase 2 – Screening, selection and categorisation

2.19 Items emerging from the electronic searches in Phase 1 were imported into Endnote reference storage platform, and subjected to screening and categorisation. A programme was created to facilitate the transfer between Endnote and Excel of screening and scrutiny information at Phases 2 and 3.

Phase 2a – Screening

2.20 For an item to be included for scrutiny and assessment at Phase 3 of the study, it needed to satisfy the criteria for relevance set out in Table 2.2. A tick box approach was chosen to aid rapidity, whilst assuring rigour and reliability (see below).

Table 2.2: Relevance criteria used at Phase 2a - Screening

Phase 2a – Does the publication…..	
….relate to the effectiveness of something that is identifiable as an approach or method? (for definition of terms see Section 1.6)	yes/no
….have relevance to teaching second language Welsh in Wales – i.e. does it • engage with pedagogy in a classroom context? AND • have relevance to learners between the ages of 3 and 16? AND • have relevance to teaching of non-dominant target languages?	yes/no

2.21 If the answer to either of the above questions was no, then the article was excluded from any further treatment in the project. For quality assurance and in order to assess consistency of judgement across the team, a 20% sample was screened

using titles and abstracts by four team members, and records were compared for inter-rater reliability. While the team were in agreement about which articles to exclude from further data extraction, at this stage it was necessary to reassess the screening process in order to aid efficiency. A 1% sample of the overall articles from the electronic search were assessed according to the inclusion criteria in Table 2.2 by four team members, using titles only. The team were accurately able to exclude 75% of the articles based on title alone, and decisions were consistent across team members. Title-only screening allowed for a much more rapid screening process. Those articles where title alone did not give sufficient information for exclusion were then assessed using abstracts.

Phase 2b – Identification of thematic domains, and preliminary categorisation

2.22 In order to inform the structure of the in-depth review and synthesis at Phases 3 and 4, in the second stage of Phase 2, preliminary categorisation data was extracted from items where the answer was 'yes' to the questions in Table 2.2 (i.e., the inclusion criteria were met). This enabled the identification of emerging themes and clusters of focus in the literature, and would facilitate a coherent approach to the synthesis of findings. While the aim was to capture emerging themes (a pre-determined set of categories would impose constraints on emphasis or inclusion), it was necessary to decide on a broad thematic approach which would maximise the usefulness of the findings. Research in the area of language learning and teaching can be categorised according to a wide range of dimensions: target language, age of learner, resource requirement, and so on. Categorisation by approach or method was possible, but was considered problematic because of the risk that relevant evidence may be excluded, or novel approaches would be forced into ill-fitting categories and because of the implicit assumption that a single approach can be effective across the full range of language skills and competences.

2.23 The framework adopted therefore was one based on language competence. This reflected the key policy ambition of "transactional competence"; as noted earlier, definitions of "transactional competence" are elusive, but the term can be sensibly interpreted as a composite of key identifiable language competences. The decision was taken, therefore, to categorise research evidence according to a list of competences, and to structure the resulting report in the same way. The following provisional taxonomy of eight thematic competences was compiled, informed by

competence descriptors in the robustly researched CEFR framework, IELTS Band Descriptors and Cambridge DELTA Young Learner Profiles:[6]

- Spoken interaction and production
- Written production
- Reading comprehension
- Listening comprehension
- E-language, technology and on-line interaction
- Metalinguistic skills
- Autonomous learning skills
- Willingness to communicate

2.24 It was anticipated that changes would be made to this list as themes emerged from the literature that had met the inclusion criteria, and indeed a preliminary scrutiny of the items emerging from the Phase 2a screening revealed two further categories for inclusion:

- Grammatical awareness
- Vocabulary knowledge

2.25 These categories were used to cluster research outputs in the second stage of Phase 2. Specifically, for those sources which passed the screening process in 2a, information was extracted to answer the following questions:

- Which (if any) of the competence-based thematic domains (listed above) is this item relevant to?
- Are any other thematic domains addressed in this item? (if so what are they?)
- What is the focus of the item in terms of method/approach?
- What is the focus of the item in terms of topic/research question?
- What is the focus of the item in terms of findings?

2.26 This stage provided the information necessary to evaluate the fitness for purpose of the categories identified a priori; to facilitate the identification of additional domains, or necessary amendments to those listed; and to identify key topics within each category. This in turn enabled us to prepare an outline of the final report, with findings organised by theme, and to operate data extraction, at Phase 3,

[6] CEFR – Common European Framework of Reference for Languages https://www.coe.int/en/web/common-european-framework-reference-languages/
IELTS – International English Language Testing System https://www.examenglish.com/IELTS/index.html
DELTA – Diploma for Teaching English to Speakers of Other Languages
http://www.cambridgeenglish.org/teaching-english/teaching-qualifications/delta/

thematically, so that outputs addressing common themes/competences could be grouped. This made it possible to note replications, contradictions, conflicting and supporting evidence within each theme, and thus enhanced the quality of evidence assessment.

Quality assurance at Phase 2

2.27 In addition to the consistency checks detailed in ***Phase 2a – Screening*** above, a practice was implemented whereby different stages of screening were conducted by different team members so that each article included for detailed data extraction in the study was scrutinised by at least two team members. The protocols for Phase 2 were developed in consultation with the Quality Assurer.

Phase 3 – Detailed data extraction and weighing of evidence

2.28 Full texts of all of the studies emerging from Phase 2 were obtained and grouped according to their thematic domain. The ten thematic competences listed in ***Phase 2b – Identification of thematic domains, and preliminary categorisation*** above were refined in response to patterns of representation within the papers. Many papers addressing Willingness to Communicate also addressed Spoken Interaction and Production, so these themes were merged, along with Listening. Technology was initially listed as a competence in its own right ("E-language, technology and on-line interaction"), but closer scrutiny of outputs revealed that in almost all cases, the focus was either on technology-supported learning, where the technology was the medium for a particular method/approach, or on affective dependent variables such as motivation or attitude, which, though of crucial importance, are outside the scope of this project.

2.29 The refined list of competences, though derived in a bottom up, principled way, emerges as reflecting conventional taxonomies of language skill:

- Vocabulary competence
- Grammatical competence
- Reading competence
- Writing competence
- Speaking and Listening competence
- General language competence

2.30 The items emerging from Phase 2 were then scrutinised for quality and relevance and coded to enable synthesis of findings. The EPPI Data Extraction form[7] (see **Annex B: Data Extraction Form (adapted from EPPI-Centre 2007)** was adapted to fit the purpose of the REA, and the question-driven scrutiny it generated was trialled by the research team for feasibility before being used. The form uses 57 questions to extract data from a research item, with the questions clustered under the following headings:

- data management details
- study aims and context
- study focus and sample
- if the study is an intervention…
- methods – data analysis
- methods – data collection and analysis
- results and conclusions
- quality

2.31 A final heading, 'weight of evidence', requires assessments to be made regarding the 'trustworthiness', 'appropriateness' and 'relevance' of the target item; the assessment therefore, though not calculated quantitatively, is driven by attention to the quality and relevance considerations targeted by the previous 57 questions. Broadly, the quality assessment was based on research rigour and integrity, and validity of evidence. Assessment of appropriateness of design and analysis was based on the degree of empirical focus, element of comparison, etc. Assessment of relevance was based on the degree of comparability with the context of teaching Welsh in Wales, and the specific parameters set out in the investigative framework. The 'weight of evidence' section of the form can be seen in Table 2.3, and the full form is **Annex B: Data Extraction Form (adapted from EPPI-Centre 2007)**.

[7] EPPI-Centre (2007) *Review Guidelines for Extracting Data and Quality Assessing Primary Studies in Educational Research. Version 2.0 London:* EPPI-Centre, Social Science Research Unit

Table 2.3: Final section of Data Extraction Form (adapted from EPPI-Centre 2007)

Weight of evidence	I
I.1 Taking account of all quality assessment issues, can the study findings be trusted in answering the study question(s)?	I.1.1 high trustworthiness I.1.2 medium trustworthiness I.1.3 low trustworthiness
I.2 Appropriateness of research design and analysis for addressing the REA.	I.2.1 high I.2.2 medium I.2.3 low
I.3 Relevance of particular focus of the study (including conceptual focus, context, sample and measures) for addressing the question, or sub-questions, of this specific systematic review.	I.3.1 high I.3.2 medium I.3.3 low
I.4 Overall weight of evidence as relevant to the REA.	I.4.1 high I.4.2 medium I.4.3 low

2.32 The composite weighting, recorded in the last row of Table 2.3 was calculated according to the system outlined in Table 2.4 below.

Table 2.4: Calculation of overall weight of evidence (I.4 in Table 2.3 above)

I.4 = HIGH	IF	(I.1 = HIGH) AND other scores are (HIGH + HIGH) or (MEDIUM + HIGH)
I.4 = MEDIUM	IF	(I.1 = HIGH) AND other scores (in any order) are (MEDIUM + MEDIUM) or (MEDIUM + LOW) or (HIGH + LOW)
I.4 = MEDIUM	IF	(I.1 = MEDIUM) AND other scores (in any order) are (HIGH + HIGH) or (HIGH + MEDIUM) or (HIGH + LOW) or (MEDIUM + LOW) or (MEDIUM + MEDIUM)
I.4 = LOW	IF	(I.1 = MEDIUM) AND other scores are (LOW + LOW)
I.4 = LOW	IF	(I.1 = LOW)

Quality assurance at Phase 3

2.33 The detailed data extraction form was trialled by three members of the team in order to compare answers to one paper and interpretations of data categories. A 10% sample of detailed data extracted was checked by another member of the team. A 10% sample of the final items included was checked for consistency, and had been given the same weighting assessment by two team members, working independently.

Phase 4 – Clustering items and synthesising key findings

2.34 To ensure that the REA findings focus on the highest quality and the most relevant outputs, it was agreed to include in the syntheses of findings only those items that scored a composite HIGH AND scored HIGH for quality AND HIGH for relevance, i.e.

I.4 HIGH = (I.1 = HIGH) AND (I.2 = HIGH or MEDIUM) AND (I.3 = HIGH)

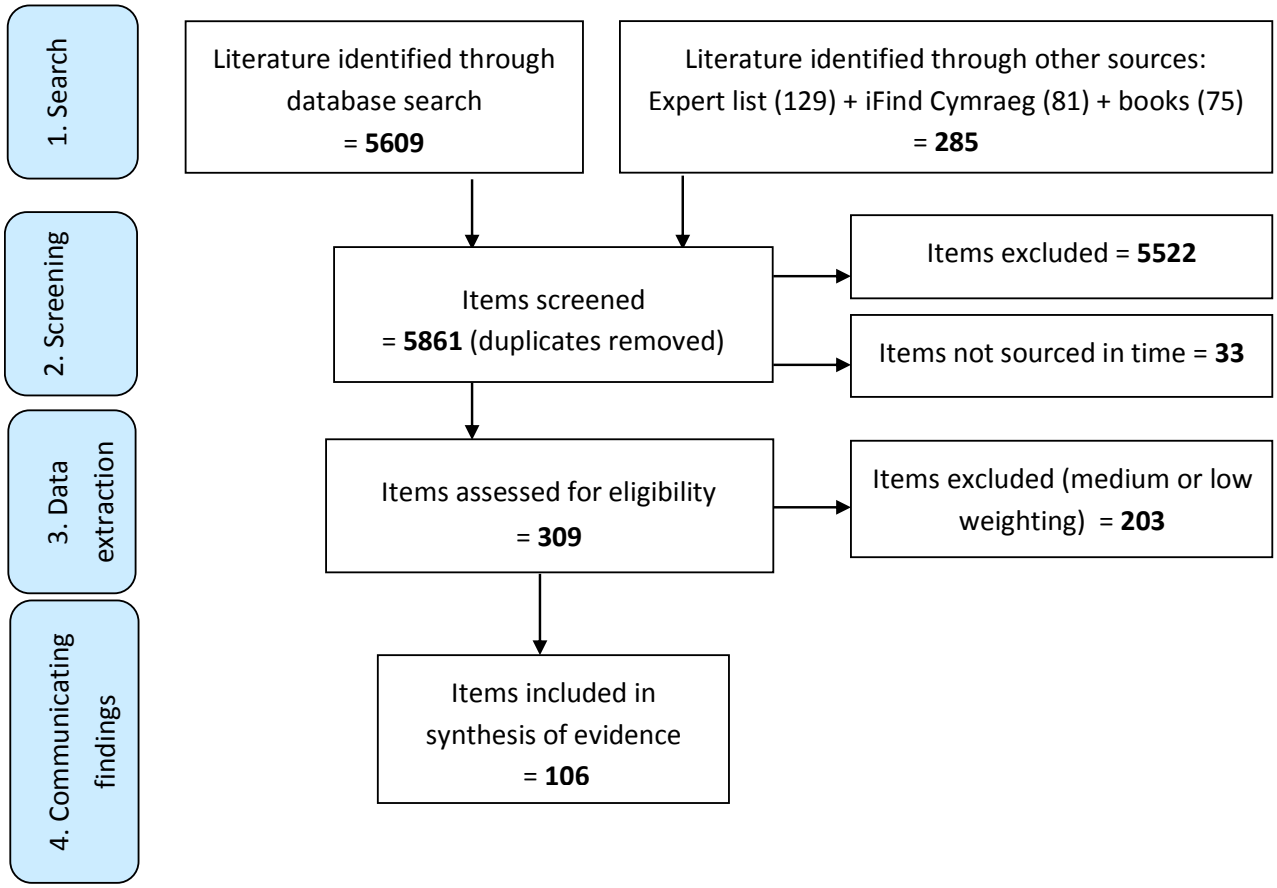
2.35 All such items were included in Phase 4, and these were grouped according to the competence-oriented categorisations developed in Phases 2 and 3 (see Sections ***Phase 2b – Identification of thematic domains, and preliminary categorisation*** and **Phase 3 – Detailed data extraction and weighing of evidence**). For each item, the key findings (regarding effectiveness of approach or method) were synthesised, and findings sharing a focus were grouped, for ease of comparison and reference between confirmatory, contradictory or otherwise connected findings. The findings are presented, organised first by competence area and then grouped according to their focus, in Section 3 of the REA.[8]

[8] All items that met the Phase 2 inclusion criteria are listed in Annex C, with an indication of those included in the research syntheses in Section 3 below.

3. Findings

3.1 This section of the REA opens with a Prisma diagram[9] (Figure 3.1) reporting the outcomes of each of the methodological stages described in **Section 2. Methodology**, in terms of the number of items emerging from each phase.[10] The following six sub-sections present research syntheses of evidence from the 106 studies that emerged from the search, screening, data extraction and weight of evidence assessments detailed in **Section 2. Methodology**. The evidence is grouped into the six areas of language competence identified in **Phase 3 – Detailed data extraction and weighing of evidence** above: vocabulary, grammatical, reading, writing, speaking and listening, and general language competence (Sub-sections **Vocabulary competence** to **General language competence**).

Figure 3.1: Flowchart of items included at each methodological stage[11]



[9] Flowchart adapted from*:* Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group (2009).
[10] At Phases 1 and 2 books are counted as a single item; after screening, individual book chapters are counted.
[11] The numbers given in Figure 3.1 are for unique items; in the following 'competence' sections and in Annex C, items that have relevance to more than one competence area may be listed more than once.

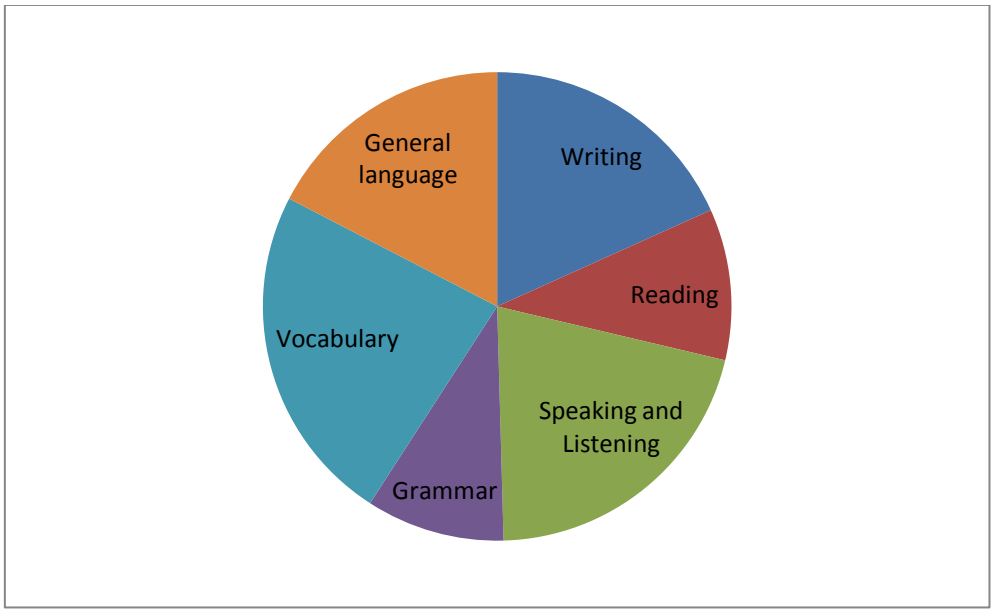
3.2 In the Section **Summary of findings applied to the Welsh context**, the application of evidence findings to the Welsh context is considered.

3.3 The competence focus of most eligible items was easily identified; in most cases it was the dependent variable measured in an empirical study, and in a few it was the area of competence that emerged as being influenced by an intervention or condition. Where items had relevance to more than one competence, they were considered under both the relevant headings (the 'number of items' summary tables at the start of Sub-sections **Vocabulary competence** to **General language competence** include these, so the numbers reported there will exceed the totals given in Figure 3.1). Items with relevance to more than two competences were allocated to the 'General language competence' category.

3.4 Any conclusions drawn from the findings in this section must take into account the derivation of evidence, i.e., that the research syntheses below are restricted to items which emerged from Phase 3 scrutiny (see **Phase 4 – Clustering items and synthesising key findings** above) as highly trustworthy, highly relevant to the context of the REA, and appropriate to the REA in terms of research design. The findings here, then, are from an apposite, but relatively small, evidence base. It should also be noted that there is an imbalance in the quantity of research across contexts. For example, there is more research on some age cohorts within the 3-16 range than others, much more research on some target languages (notably English) than others (notably non-dominant languages), and a more frequent focus on some competences (e.g. vocabulary knowledge) than others (e.g. listening competence). The findings reported below do not attempt to address that imbalance.

3.5 The distribution of the 106 items included in the syntheses below, between the six competence areas, is shown in Figure 3.2.

Figure 3.2: Distribution by competence of items included in the synthesis of evidence



Vocabulary competence

Table 3.1: Summary of items (vocabulary competence)

Number of items identified at Phase 2	72
Number of items not sourced	5
Number of items excluded from evidence synthesis at Phase 3	40
Number of items included in evidence synthesis	27

3.6 There are five key areas of focus in the literature relating to effective methods for teaching vocabulary to students aged 3-16. These are i) CLIL learning contexts (6 studies); ii) Intentional focus on form and meaning of individual words (7 studies); iii) Input-only, and input plus interaction instruction (8 studies); iv) Involvement load (2 studies); v) Integrating imagery, gesture and movement (4 studies).

The effect of CLIL instruction on vocabulary development

3.7 Six studies included for synthesis here relate to vocabulary development in CLIL contexts. The national and first language contexts vary, as does the age of the learners, but the studies have in common that they compare vocabulary size or gain in English language CLIL and non-CLIL experiences. Their commonalities and differences are summarised in Table 3.2.

Table 3.2: Studies of CLIL and vocabulary learning

Study	context and L1 (if given)	compared conditions	vocabulary tested	n per condition	age (yrs)
Merikivi and Pietilä (2014)	Finland, L1 Finnish	CLIL vs "mainstream"	receptive, productive	75, 88 CLIL 74, 93 non-CLIL	13, 16
Agustín-Llach and Canga Alonso (2016)	Spain	CLIL vs EFL	receptive	58 CLIL 49 non-CLIL	9-12
Gierlinger and Wagner (2016)	Austria, L1 83% German, 17% other	extra 60-80 hours CLIL	receptive	39 CLIL 48 non-CLIL	12-14
Jiménez Catalán and Ruiz de Zarobe (2009)	Spain L1 Spanish	CLIL vs EFL	receptive	65 CLIL 65 non-CLIL	11-12
Sylven (2010)	Sweden	CLIL vs non-CLIL	receptive	137 CLIL 84 non-CLIL	15-16
Tragant, Marsol, Serrano and Llanes (2016)	Spain	CLIL vs EFL	productive	22 (within subject)	8

3.8 Three studies: Jiménez Catalán and Ruiz de Zarobe (2009), Sylven (2010), and Agustín-Llach and Canga Alonso (2016), found CLIL learners' scores to be significantly higher than those of non-CLIL learners on a general test of receptive vocabulary. Merikivi and Pietilä (2014)'s CLIL learners outperformed the non-CLIL group in both receptive and productive vocabulary scores at 6th and 9th grades, with large differences particularly in the younger group. However, the other three papers note that the difference between group scores was small, and in fact Agustín & Canga found no difference in their youngest (9-10) group. Gierlinger and Wagner (2016) cast further doubt on CLIL's contribution to vocabulary growth; they measured vocabulary gains of a group receiving an extra 60-80 hours of CLIL tuition, against a control group, and found that both groups' receptive vocabulary size grew, with no significant effect of the CLIL treatment. Finally, Tragant, Marsol,

Serrano, and Llanes' within-subject comparison (2016) tested learners on words encountered in the class EFL (term one) and science (term two) textbooks, and found vocabulary gains to be significantly higher in the EFL context than the CLIL context. These CLIL studies urge caution in interpreting apparent gains from CLIL treatment, and suggest that contributory factors include hours of target language exposure, earlier exposure to target language, the quality of vocabulary input in CLIL (and the frequency-based vocabulary measures used in the studies), and the fact that CLIL is sometimes an opt-in reflecting students' (and parents') motivation to learn (in these cases) English.

Focusing attention and intentional learning activity on form and meaning of individual vocabulary items

3.9 Seven of the studies included in this REA compare Focus on Form (FonF), where attention is given to an item during a communicative activity, with the more traditional Focus on Forms (FonFs), where the learning focus is the items themselves. Laufer (2006) claims the role of the latter to be of major importance in learning conditions that do not emulate L1 input. She found that 16-year-old learners performed better (72% success) on target item recall when they had studied and practised items from a word list than when they had read a text containing the words, supported by dictionary look-up (47%). A follow-up study (Laufer & Girsai, 2008), differentiated not only between FonFs and FonF, but also between contrastive (L1-L2) and non-contrastive FonFs instruction. They found that 15-16 year old L1 Hebrew learners of English (n=75) receiving contrastive instruction in the form of L2>L1 and L1>L2 translation tasks, out-performed learners in the FonF and the non-contrastive FonFs conditions. (See 3.13 – 3.15 for further consideration of the role of L1).

3.10 Findings from Shintani (2013), who set out to compare the effectiveness of FonF and FonFs on 6-year-old learners' productive vocabulary knowledge, challenge those of Laufer, but replicate those of de la Fuente (2006) (whose participants were university students). Shintani's task-based teaching activity represented FonF, and a PPP approach (Presentation = repetition of target; Practice = drill; Production = in a game) represented FonFs. Target items were 24 nouns, for which there was no significant difference in uptake between the two groups, and 12 adjectives, for which the FonF group out-performed the FonFs group.

3.11 A possible explanation for these contradictory findings comes from studies that suggest the effectiveness of FonF depends on the timing and nature of teacher-led interventions. Alcón (2007) analysed audio recordings from a year of English language classes, along with learner diaries from 14-15 year old Spanish/Catalan L1 participants (n=12), and identified instances of 'pre-emptive' and 'reactive' focus on form by the teacher. She found that pre-emptive FonF led to 'noticing' items (as reported in learner diaries), and there was a positive correlation between noticing and post-test production of items. On the other hand, Hennebry, Rogers, Macaro and Murphy's study of 262 14-year old learners of French (2017) found that vocabulary instruction (including in L1) after a listening activity led to more effective recall of vocabulary than a listening only condition.

3.12 Several studies emphasise the role of the teacher in determining the timing and nature of the focus on vocabulary; indeed, Graham, Courtney, Marinis and Tonkyn (2017) found in comparing an oracy focussed approach with an oracy plus literacy approach to teaching French to 9 and 10 year olds (n=252), that teacher expertise was a more influential variable than teaching approach.

Input-Only, and Input plus Interaction Instruction

3.13 Eight of the studies meeting our weighting criteria investigated the effectiveness of L2 input conditions. Williams and Thomas (2017) assessed uptake of Welsh vocabulary by 4-5 year old English speakers in four 20-minute interventions applied three times weekly for 6 weeks. The interventions were: i) watching 15 Welsh TV programmes; ii) as i, but with teacher interaction; iii) "storytime", where the same stories as in the TV programmes were read aloud in Welsh; iv) as iii), but in English (control group). Post-tests on vocabulary from the programmes/stories found that the control group was outperformed by all other conditions. The highest vocabulary scores were seen for watching the TV programmes with teacher interaction, and there was no difference in performance between those who watched the TV programmes and those who listened to the same stories read by a teacher. The authors conclude that TV programmes, especially when viewed in interaction with a caregiver, can facilitate language uptake. Lin (2014) also reports vocabulary uptake from input-only exposure: 45 Chinese L1 learners of English, aged 10-11, listened to a story read aloud by the teacher four times, and by the third time had learned four of 8 target (non-) words.

3.14 Comparing 6-8 year olds' input-based instruction (listen-and-do card selection tasks, n=13) with production-based instruction (matched tasks but with cued production, n=11), Shintani (2011) found no significant difference in performance on three of four post-task vocabulary tests. Shintani (2012) uses the same context and finds that teacher input is modified, and learners' voluntary production of target items increases as the input-based task is repeated nine times over five weeks. Her finding that learners' negotiation of input pushes vocabulary gains from an input-only task is supported by Luan and Sappathy's study of 10-11 year old L1 Malay learners of English (2011).

3.15 Evidence suggests that input can be made more effective by including strategic use of L1. Lee and Macaro (2013) investigated use of L1 in vocabulary instruction. They tested vocabulary recall and recognition of 12-year old Korean learners of English. The teacher used code-switching into Korean to give information about new lexical items for 223 learners, while 220 received English only instruction. For acquisition and retention, the code-switching group significantly outperformed the English only group in both recall and recognition. Lesniewska and Pichette (2016) similarly note the apparent effectiveness of using L1 in activities to enhance L2 uptake, as do Camo and Ballester (2015), who tested the acquisition of 20 target items by 10-11 year old learners of English. Their experiment and control groups listened to a story, and as a target word appeared, the experiment group were shown and heard the word in both L2 and L1 (Catalan), whereas the control group were only exposed to the L2 item. In post-test picture matching tasks, the experiment group scored higher for accuracy and speed of response.

Involvement Load

3.16 Huang, Willson and Eslami (2012) aimed to investigate the effectiveness of output tasks in development of learner vocabulary. Their meta-analysis of 12 studies included six with high school learners. They found that vocabulary gains were greatest where the involvement load of the task was high, where a combination of output tasks were undertaken, and where more time was spent on task. Involvement load is a motivational-cognitive construct; a task requiring learners to need, search for and evaluate the meaning of a word is interpreted as having high involvement load. Task-based learning (TBL) and digital game based learning (DGBL) are candidates for high involvement load. From a meta-analysis of ten studies, 7 of which focus on young learners, Chen, Tseng and Hsaio (2018) find

that the greater the fun and adventure-challenge component in a game, the more likely it is to be effective in terms of vocabulary uptake.

Integrating imagery, gesture and movement

3.17 In the Keyword method of learning new vocabulary, learners attend to phonetic or orthographic features of the target item, link these to a familiar 'keyword', usually in the L1, and create a mental image representing the link. Dolean (2014) investigates the effectiveness of this method in teaching 101 Romanian L1 learners of Italian, aged 9-10, and finds that presenting learners with the image of the target word, alongside a keyword image, leads to significantly better L2>L1 translation performance in an immediate post-test, and a follow up study with 24 7-8 year olds and 21 13-14 year olds found a significant positive effect of keyword presentation, including in a delayed post-test. Dolean and Dolghi (2016), teaching 30 imageable items to 6-7 year old Romanian learners of English (n=34) found that a Keyword-instructed group significantly outperformed a Total Physical Response-instructed group, with a large effect size.

3.18 Investigating the integration of gesture and movement in learning, Porter (2016) found an initial significant advantage to gesture elaboration while teaching formulaic utterances to 4-7 year olds (n=40), but that recall dropped considerably (though remained higher than control group scores) at a 2-week delayed post test. Mavilidi, Okely, Chandler, Cliff and Paas (2015) compare four learning conditions for teaching 14 Italian words to 111 children (mean age 4.9). The conditions were simultaneous to visual and oral word presentation, and were: integrated physical exercise (children enact actions); non-integrated physical exercise (unrelated to item); gesture (gestures to act words while seated); conventional (repeat words while seated). Free- and cued-recall test scores were low for all conditions. The integrated group performed significantly better than other groups for free recall, but still their average score was below three out of 14 words recalled. In cued recall, no significant difference was found between the two physical exercise groups, but they both performed significantly better than the other conditions, and the gesture condition produced higher scores than the conventional condition.

Summary of Vocabulary Competence findings

- Learners in CLIL contexts tend to perform better on vocabulary tests than non-CLIL learners, but evidence for this is not consistent, and can be attributed to factors other than teaching method/approach.
- Focusing attention and intentional learning activity on form and meaning of individual vocabulary items enhances vocabulary uptake, but this must be strategically applied.
- Input-only instruction (including L2 television) is effective in terms of vocabulary uptake, and learning gains are enhanced further when input is supplemented with interaction.
- Vocabulary learning is facilitated by tasks with high involvement load.
- Integrating creative imagery can boost vocabulary uptake; integrating gesture and movement yields more modest gains.
- Timing and variety of mini-interventions in learning has a significant impact on effective learning, regardless of approach/method.

Grammatical competence

Table 3.3: Summary of items (grammatical competence)

Number of items identified at Phase 2	36
Number of items not sourced	7
Number of items excluded from evidence synthesis at Phase 3	18
Number of items included in evidence synthesis	11

3.19 There are three key areas of focus in the literature relating to effective methods for teaching grammar to students ages 3-16. These are i) Explicit instruction (7 studies); ii) CLIL (2 studies); iii) Oracy and literacy (2 studies).

Explicit Instruction

3.20 The term “explicit instruction” is taken here to mean any kind of instruction in which learners come to an explicit awareness of target language features. In deductive instruction, these features are brought directly to learners’ attention by the

instructor; in inductive instruction, they are led to discover the target features for themselves, and these are later confirmed by the instructor.[12]

3.21 Hanan (2015) found that explicit instruction with either a focus on form-meaning correspondence, or on form only, was effective for learners of German aged 9-11, who made substantial gains on written, oral, and metalinguistic tasks. The preference for explicit instruction is supported by Tode (2007), who found that explicit instruction for Japanese learners of English aged 12 and 13 led to immediate performance gains as compared to implicitly-instructed subjects, but that these gains did not persist to delayed post-test.

3.22 Lichtman (2013) provides evidence from children aged 8- 17 showing that explicit instruction is particularly effective for tasks that test explicit knowledge of grammar, whereas implicit instruction leads to stronger performance in tasks believed to tap implicit knowledge, such as a story rewriting task. Findings by Toth and Guijarro-Fuentes (2013) expand on this, with evidence that explicit instruction leads to improvement in tests that target implicit knowledge for learners aged 15-17. Tammenga-Helmantel, Arends, and Canrinus' (2014) study of 981 Dutch children aged 12-15 learning three different L2s found that any kind of exposure to a target form, whether explicit, implicit, or incidental, leads to gains on grammaticality judgement tests (GJTs) and the correct use of the target item in writing tests. They found a weak preference for explicit instruction in two contexts only: GJT scores for learners of English, and writing test scores for learners of German. Ho and Binh (2014) support these findings, in that both traditional grammar translation method (GTM) teaching and communicative-style explicit instruction were found to increase grammatical knowledge in participants aged 12, although only the communicative-style instruction led to gains on an oral production task.

3.23 Ho and Binh (2014) also found that an inductive teaching method led to gains in grammatical competence on both a grammar test and an oral production test. Hanan (2015) provides evidence that deductive methods increase gains across a range of test types. Tammenga-Helmantel, Bazhutkina, Steringa, Hummel and Suhre (2016) treated deductive and inductive methods as variables with learners aged 15-18, and found that inductive instruction is slightly more effective for performance on a GJT, but not on a writing test. Tammenga-Helmantel et al. (2014) found weak evidence for the superiority of deductive instruction over implicit

[12] Although the terms 'inductive' and 'deductive' instruction are not used in all studies, treatments are described in such a way that this categorisation can be made.

instruction on an English GJT, and for the superiority of inductive instruction over implicit instruction on a German performance test. The effect sizes here are small, however, and other tests in English, German, and Spanish showed no preference for either type of instruction.

Content and Language Integrated Learning (CLIL)

3.24 Ibarrola (2012) found CLIL students learning English to be ahead of their counterparts (who were following a curriculum of English as a school subject (ESS) only) on a range of grammatical measures when the former were aged 13 and the latter aged 15. They conclude that CLIL instruction, begun in this case at age 12, provides a significant advantage in the development of grammatical competence. The findings of Martínez Adrián and Gutiérrez Mangado (2009) are more equivocal, showing equivalent performance for CLIL learners and ESS learners aged 14 in all but one outcome measure.

3.25 They argue that the relatively small effect of CLIL versus ESS instruction in their study may be the result of the CLIL learners' having experienced only 363 hours of extra exposure at the time of the study. They suggest that the effects of CLIL may be tied to the amount of exposure to the target language. This conclusion is also drawn by Ibarrola (2012) , who argues that the positive effect of CLIL on grammatical competence and rate of acquisition of the target language is largely a result of increased exposure. She further argues that her results support the commencement of CLIL instruction around age 12, hypothesizing that this is an age when the level of cognitive maturity means that greater exposure can be transformed into faster rate of learning.

Oracy and Literacy

3.26 Campfield and Murphy (2017) found that providing eight-year-old learners of English with input rich in prosodically-marked features led to better results on a GJT testing understanding of English word order. This is interpreted as support for the 'prosodic bootstrapping hypothesis', which argues that prosodic features in continuous speech underline the grammatical functions of different word types.

3.27 Graham, Courtney, Marinis and Tonkyn (2017) considered the relative merits of an oracy-based and a literacy-based approach to teaching primary school children aged 9-10, and found that there was little difference in outcome between the two. Learners with lower L1 (English) literacy scores, however, were slightly benefitted by a literacy approach to the teaching of French. The study found that the teacher's

level of training, and the number of hours of instruction, were far more important variables than type of instruction. Pupils with a teacher with degree-level French made significant progress in grammatical competence at all test points, whereas those whose teacher had GCSE-level French or below failed to progress between school years 5 and 6. Those who received 60 minutes or more of instruction per week achieved test scores in year 5 that were barely achieved by students receiving less instruction in year 7. The authors conclude that type of instruction is not a decisive factor in children's grammatical development in the L2 between primary and secondary school.

Summary of Grammatical Competence findings

- Explicit instruction is effective, but rarely more effective than other types of instruction, in developing grammatical competence.
- Both inductive and deductive types of explicit instruction are effective, but inductive may be slightly more effective under certain conditions.
- CLIL is more effective than teaching a language as a school subject for the development of some areas of grammatical competence.
- Although CLIL appears to be a particularly effective method, this may well be the result of the greater exposure to the target language enjoyed by these students, rather than of the methodology per se.
- There may be an optimum age (12?) at which to commence CLIL instruction.
- Attention to prosodic features (e.g. rhyme, rhythm) in oral input can aid the development of grammatical competence.
- Both oracy and literacy approaches in primary school can be effective in developing grammatical competence.
- Teacher language competence and number of hours' instruction are more influential factors than instruction type.

Reading competence

Table 3.4: Summary of items (reading competence)

Number of items identified at Phase 2	34
Number of items not sourced	2
Number of items excluded from evidence synthesis at Phase 3	20
Number of items included in evidence synthesis	12

3.28 There are two key areas of focus in the literature relating to effective methods for teaching reading skills to students ages 3-16. These are i) Development of lower-level processing skills such as orthographic decoding (5 studies); ii) Working with higher-level processing skills including strategy development and attention to cognitive load (7 studies).

Developing lower-level processing skills

3.29 Takeda (2002) investigated the effects of six months of phonics instruction on Japanese L1 learners of English (aged 12-13). The treatment group out-performed a control group in improvements in (i) reading words, listening to three possible pronunciations of each, and identifying the correct pronunciations; and (ii) reading aloud words printed on the test paper. Phonics instruction was shown to be a significant factor in the treatment group's improvement on both tests. A qualitative questionnaire completed by the treatment group did not clearly demonstrate improvements in comprehension as opposed to simply being able to read words aloud.

3.30 Similar results were found by Fonseca-Mora Jara-Jiménez and Gómez-Domínguez (2015), who studied an 11-week intervention with 7-8 year old beginner EFL students in Spain. They compared a control group (in which the teacher used the syllabic and global word approach) with two experimental groups receiving phonological training, one with music support. The phonological training programme included phonics instruction, and also phonological awareness development, particularly of sounds which are not distinguished in the learners' L1. Students in both experimental groups performed significantly better than students in the control group on tests (i) *naming upper and lowercase letters presented randomly to the student* and (ii) *identifying the initial sound of ten words read aloud*. In addition, a non significant trend was reported for the largest improvement on test (iii*) reading a dialogue with accuracy, speed and fluency*, to be found in the non-music phonological treatment group.

3.31 Porter's (2014) study of 9-11 year olds learning French in England somewhat supports this. The pedagogical approach implemented during this 23-week study involved four main elements: Simultaneous oracy and literacy development; Focus on L2 sounds; Systematic and explicit phonics instruction; and Experiencing L2 sound and print. As with Fonseca-Mora et al.'s (2015) intervention, Porter's (2014) phonological training involved attention to differences between the L1 and the target

language. Statistically significant increases were found in both reading aloud and reading comprehension scores in week 21 of the intervention. Unlike the studies reported by Takeda (2002) and Fonseca-Mora et al. (2015) this was a one-group study – i.e. there was no comparison with a control group, and it is possible that benefits cannot fully be attributed to the intervention.

3.32 Two studies investigated technology-driven methods for developing oral reading fluency. In a study of 9-10 year old EFL students in Taiwan, Lan, Sung and Chang (2009) compared an experimental group learning reading skills using a computer-assisted system with a control group using paper-based versions of the same materials and activities. In both modes, instruction included individual learning (e.g. learning phonics rules and vocabulary, reading a paragraph of a text), cooperative learning (e.g. teaching each other the learned rules and vocabulary, putting paragraphs into the right order), and peer/teacher assessment. In the computer-assisted group, each student had a tablet, stylus and headset, and completed individual and cooperative activities and peer assessment by logging into the bespoke software. Oral reading fluency was tested before and after the 10-week treatment, and although significant improvements were made by both groups, there was no significant difference between groups. However, raw scores suggest that in the control group the instruction mainly benefited high-level ability students, whereas the computer-assisted instruction benefited most students. In addition, video recordings of the classes allowed observation of learner and teacher behaviour. Analysis of the time spent on different behaviours demonstrated that learning-unrelated behaviours (e.g. chatting, looking around, walking around, playing alone or with others or with equipment) were significantly more common in the control group than in the computer-assisted group. In addition, the experimental group appeared better able to focus on the activities allocated to them by the computer-assisted system (i.e. independent and cooperative learning activities, as and when appropriate), whereas the control group showed far more teacher-dependency.

3.33 Chen, Tan and Lo (2016) also studied Taiwanese EFL students, this time aged 13-14. The experimental group in this 8-week study practised repeated reading with the support of a digital pen, whilst the control group practised with the support of a peer and/or the teacher. For both groups, instruction consisted of a reading demonstration (via the digital pen or read aloud by the teacher), student imitation of the model and repeated practice, and peer assessment of oral readings (via a

Moodle course management system or face-to-face). The principal difference between the groups was that the digital pen allowed students in the experimental group to record their own readings and listen back to them, thus encouraging self-assessment in comparison to the model. Both groups of students significantly improved on oral reading fluency tests between pre- and post-tests. However, the experimental group made more significant progress than the control group. Qualitative data collected via interviews with a sample of students from the experimental group point to the benefits of self-learning via the digital pen system in terms of learners' active engagement with and control of their learning.

Developing higher-level processing skills

3.34 The six studies of interventions involving reading strategy instruction are unanimous in finding it plays a role in developing comprehension skills (Harris, 2007; Macaro & Erler, 2008; Macaro & Mutton, 2009; Manoli, Papadopoulou, & Metallidou, 2016; Martínez & de Zarobe, 2017; Mistar, Zuhairi, & Yanti, 2016). Strategy instruction was variously implemented with groups ranging in age from 10-11 to 15-16. In all cases, treatment groups outperformed control groups on measures of reading comprehension ability, and Martínez & de Zarobe (2017) found this to be the case whether students were being taught via CLIL or via EFL teaching approaches.

3.35 The strategies chosen for instruction varied somewhat between studies, but many were used in multiple studies. The most commonly implemented were:

- Predicting text content (four of the studies).
- Inferring meaning of unfamiliar words from their context (four studies).
- Using prior/background knowledge/common sense (three studies).
- Skimming (getting the gist) or looking for the main ideas (three studies).
- Scanning (identifying specific information) or looking for details (two studies).
- Not giving up (two studies).
- Sounding out an unfamiliar word/phrase to aid decoding (two studies).
- Checking deductions/guesses make sense (two studies).

3.36 All but one (Macaro & Mutton, 2009) of the six studies of strategy instruction involved the teaching of more than one strategy simultaneously.

3.37 The five studies that explicitly taught multiple reading strategies taught three (Mistar et al., 2016), four (Manoli et al., 2016), five (Martínez & de Zarobe, 2017), six (Macaro & Erler, 2008) and 18 (Harris, 2007) strategies at once. Whilst all

interventions were successful in improving students' comprehension skills, Harris (2007) found that the very large number of strategies in her study caused some problems for the learners, including difficulties in remembering all the different strategies and difficulties in selecting the most appropriate strategy from the large repertoire. No such problems were reported in the other four studies, suggesting that explicitly teaching three to six strategies at a time might be more appropriate.

3.38 Two studies investigated the effect on reading comprehension of manipulating cognitive load: Türk & Erçetin (2014) investigated the use of multimedia glosses during reading comprehension tasks, comparing two conditions of learner control over the presentation of glosses: learner choice of text, graphics or both; and simultaneous presentation of text and graphics (no learner choice). The study showed that students in the simultaneous group outperformed those who had a choice of glosses in reading comprehension tests. The authors conclude that this is explained by a reduction in cognitive load. The implication is that materials which are adapted in order to reduce the distribution of cognitive load could lead to better learning.

3.39 The learning materials in Macaro & Mutton's (2009) interventions were: (i) graded readers; and (ii) a code-switched text – i.e. an age-appropriate novel in the learners' L1, with target language words embedded into the text in gradually increasing proportions. It should be noted that the code-switched text group also received strategy training, whereas the graded readers group did not. Students in both interventions made significant advances in reading comprehension in comparison with a control group who received normal teaching provision rather than dedicated reading time with adapted materials. Both types of materials used in the study have been adapted – for example, graded readers include vocabulary and grammar structures appropriate to ability level. However, the code-switched text helps learners notice and acquire 'little words' (function words) which can be important for comprehension but that might be overlooked in L2-only texts. It achieves this by "reduc[ing] the cognitive load on working memory by expecting the reader to hold the context not in L2 but in the much more familiar L1, thereby freeing up capacity to process the phrase, not only for its meaning but also for its form" (2009, p. 176).

Summary of Reading Competence findings

- Instruction for the development of reading skills needs to include explicit attention to both lower-level and higher-level processes; it cannot be assumed that either will be transferred from a student's L1 without instruction.
- Phonological training can help beginner learners to process word forms, but not necessarily word meanings.
- Technology can contribute to the development of oral reading fluency by facilitating student-centred learning.
- In contrast to paper-based learning, computer-assisted learning of reading enhances focus on task, reduces teacher dependency, and benefits lower ability learners.
- Strategy instruction is an effective method of developing reading comprehension skills.
- Multiple strategies can be taught together, but not in excessive quantities.
- Reading materials which are adapted to reduce cognitive load can scaffold comprehension.

Writing competence

Table 3.5: Summary of items (writing competence)

Number of items identified at Phase 2	51
Number of items not sourced	2
Number of items excluded from evidence synthesis at Phase 3	28
Number of items included in evidence synthesis	21

3.40 There are three key areas of focus in the literature relating to effective methods and approaches for teaching L2 writing skills to learners aged 3-16. These are i) The role of technology (6 studies); ii) Non-traditional pedagogical approaches (10 studies); iii) Taught strategies for writing (5 studies).

The role of technology in writing development

3.41 Although there are a large number of research works describing the positive impact of ICT (Information and Computer Technologies) on the general learning process, there are fewer indications that it is necessarily advantageous to students' writing in particular (Yunus, Nordin, Salehi, Embi, & Salehi, 2013). The importance of

understanding how best to use technology, rather than assuming it to be advantageous merely by its presence, cannot be overlooked. It is clear that students must be sufficiently supported, and that writing improvement will only substantially occur under suitable conditions.

3.42 Fidaoui, Bahous and Bacha's (2010) exploration of the use of CALL (Computer-Assisted Language Learning) indicated that the majority of participants felt it was an enjoyable (95.8 %) and motivational (64.6 %) influence on EFL writing acquisition. Students had weekly sessions for 50 minutes in a computer lab, in addition to access to one networked machine in their main classroom. Using CALL, they were better able to express their perspectives, gather and synthesise online information and develop "creative, neat, organised, error-free written products" (Fidaoui et al., 2010, p. 164). However, as a caveat, initial learner weaknesses in elementary computer and research skills prevented them from producing high-quality written work. The authors recommend careful planning of ICT-based writing work, and that students are properly monitored throughout. Users must be taught how to access reliable websites, to adequately paraphrase and avoid plagiarism from Internet sources, to focus on writing neatness and to use a range of relevant computer features as required.

3.43 These findings were supported by Yunus et al.'s (2013) study to an extent, which also noted the advantage of using ICT to stimulate student interest, develop their lexical knowledge and promote practical learning. However, further challenges were also identified, namely increased difficulty of controlling the class, ease of distraction and the tendency of pupils to write short-form responses to tasks. It was also revealed that teachers in that project were somewhat weak at handling these problems, and that planning to use computers in writing sessions was less than adequate.

3.44 It may be the case that the impact of using more specific elements of CALL / ICT is easier to ascertain. For example, Taylor, Lazarus and Cole (2005) investigated the use of drop-down menus to provide writing frames, which assisted students by providing options for part of a sentence they were attempting to formulate. Students were found to write at greater length, increase their accuracy and engage with tasks more enthusiastically. Progress was supported by PowerPoint presentations, in which grammar was addressed – an important additional part of the writing development taking place, which was not to be overlooked.

3.45 Other research has explored the use of Weblogs, concluding that a positive impact also depends on a considered approach (Sercu, 2013). Encouraging learners to contribute to an online blog led to a sense of achievement and successful collaboration, and increased motivation was reported once again. Seeking peer approval and the benefits of co-constructing a text were found to be especially appealing to participants. Nonetheless, students who were weaker than their classmates at using technology felt marginalised and somewhat less enthusiastic about participation – they struggled to reach the level required. This echoes the findings of Fidaoui et al. (2010) and underlines the importance of considering how to help those pupils with limited technological skills.

3.46 Using mobile devices was also found to potentially increase motivation to learn writing skills in class, and subsequently encourage achievement, as such activities were perceived as enjoyable and engaging (Hwang, Chen, Shadiev, Huang, & Chen, 2014) . Using mobiles for 'situated writing' about three familiar contexts (classroom, meal and playground) was challenging for the elementary students involved, but those who were asked to use mobiles demonstrated a higher performance when asked to describe the environment and express their ideas than those who were not. Finally, Facebook has also been suggested as a tool for developing writing skills (Buga, Capeneata, Chirasnel, & Popa, 2014). Findings showed that it has the potential to help students experiment with learning methods, and that participants completed written homework tasks, having not previously done so.

Non-traditional approaches to writing instruction

3.47 Contemporary instructional techniques have also been identified as potentially advantageous to writing. 'Flipped learning' involves giving students online materials prior to class, and using the lesson time to deepen understanding of these, for example through collaborative problem-solving activities. Abdelrahman, Dewitt, Alias and Rahman (2017) found that writing proficiency and pupil engagement improved using this approach, particularly due to the interactive nature of the tasks set. The teacher was able to allocate more class time to help the learners, which was cited as an especially positive outcome. The authors note the possibility that such an approach might not suit every student.

3.48 Ngo and Trinh (2011) evaluated the impact of a process-based approach to writing in a context where examined, product-based approaches are the norm. Process writing – involving idea generation, multiple drafts, emphasising the reader, collaboration and creativity – increased student writing performance and enthusiasm for written work. Focusing on content over correcting grammatical inaccuracies was particularly instrumental to motivation, in addition to prioritising a strong communicative message over error avoidance.

3.49 Another non-traditional approach explored was dialogue journal writing, which was shown to empower students, foster critical awareness and develop the notion of 'voice' (Ghahremani-Ghajar & Mirhosseini, 2005). Teaching literacy and oracy together has also been suggested as an innovative means of instruction, and one which is feasible without one element necessarily having a negative impact on the other (Porter, 2014). First language (L1) reading age and verbal working memory proficiency are identified as highly important to second language (L2) oracy and literacy education. However, all ranges of ability are capable of participating in L2 combined instruction, and progress can be made by lower level pupils too.

3.50 Bartan (2017) explored improvement of Turkish learners' English language writing through short story reading. Specifically, using reading to provide a model for writing was found to help students develop their language, content, organisational structure and communicative achievement. Collaborative translation was also investigated as a potential means of instruction (Bruton, 2007). Undertaking these tasks increased vocabulary development as students wrote, and they also became able to identify different text formats they were exposed to. This was described as a potentially useful skill to acquire for future use.

3.51 Finally, the relationship between CLIL and writing instruction must also be considered. Saladrigues and Llanes (2014) revealed a modest advantage for CLIL participants over their non-CLIL counterparts in written fluency, though not in syntactic complexity. Amount of L2 exposure played an important role in subsequent writing performance - whether or not the language content was chosen for a specific purpose or context. Other research supported this, indicating that CLIL students' writing contained more features of the target language discourse than that of non-CLIL learners when compared (Maxwell-Reid, 2010). Gené-Gil (2015) echoes this point, concluding that higher writing performance scores for CLIL participants than non-CLIL were evident. Despite this, the need to avoid simplistic conclusions about its effectiveness was underlined. Although overall the CLIL group

benefited, there were particular areas in which performance was lower – such as lexical complexity, for example. Ikeda (2013) reports that a 'weak' or 'soft' version of CLIL leads to improvement in writing fluency and complexity, but not in accuracy. Notes of caution accompany this finding: language acquisition was prioritised over content learning (possibly at the expense of the latter), gains were not solely attributable to CLIL lessons, and there are significant teacher training implications.

Taught strategies to promote writing skills

3.52 In addition to direct pedagogical approaches, strategies may be taught to assist learner writing. Layered instructional strategy - matching students to their learning needs rather than using one approach for all class members – was shown to significantly benefit the writing skills of those who received it (Shafqat, Idrees, & Gujjar, 2009). Similarly, accommodating multiple intelligences in the classroom was shown to improve writing acquisition more effectively than traditional methods (Gündüz & Ünal, 2016). In addition, participants reacted more positively to this when compared to regular practices.

3.53 Griva & Anastasiou (2009) investigated morphological awareness training as a route to spelling improvement in written tasks. Spelling did largely improve, but other aspects of learning to write which were not included in the training did fall somewhat behind. The authors conclude that although this does support earlier research findings, more work in the area is required. Other strategies taught to assist learner writing included mind-maps, brainstorming and pre-writing planning, facilitated using computers (Lan, Sung, Cheng, & Chang, 2015). Results showed that mind-mapping and drawing before writing increased grammatical knowledge significantly, and were particularly appealing to younger learners. Learners reacted positively to their use overall, and displayed awareness of why they had been engaging in such activities.

3.54 Finally, reducing learner anxiety about the writing process through creating a suitable environment has been shown to improve performance; and render writing instruction more effective (Tsiriotakis, Vassilaki, Spantidakis, & Stavrou, 2017). Specifically, using cognitive strategy instruction (Self-Regulated Strategy Development, SRSD) is a recommended approach to writing instruction. Developing student critical awareness and ability to learn in a social environment, for example, will better equip them to write independently later on.

Summary of Writing Competence findings

- Technology has a potentially important role to play in writing development if carefully implemented.
- Technology cannot be assumed to be advantageous merely by its presence; its use must be informed and planned.
- Learners require training in using ICT-based information for writing.
- Use of ICT requires specific classroom management techniques.
- Pupils with limited technological skills must be appropriately supported.
- Interactive writing tasks promote proficiency and engagement.
- Combining literacy and oracy teaching can be beneficial, especially in mixed ability classes.
- Strategy training (e.g. planning, morphological awareness, cognitive strategies) can improve writing performance.

Speaking and listening competence

Table 3.6: Summary of items (speaking and listening competence)

Speaking	
Number of items identified at Phase 2	62
Number of items not sourced	5
Number of items excluded from evidence synthesis at Phase 3	42
Number of items included in evidence synthesis	15
Listening	
Number of items identified at Phase 2	19
Number of items not sourced	4
Number of items excluded from evidence synthesis at Phase 3	6
Number of items included in evidence synthesis	9

3.55 Because of numerous overlaps in the content of these items, these topics have been integrated into a section on speaking and listening, organised into four key focus areas:, i) Interaction and negotiating meaning (12 studies); ii) Pronunciation (4 studies); iii) Listening (5 studies); iv) Speaking and listening strategies (3 studies). These counts, and those in the table above, include overlapping items.

Interaction and negotiating meaning

3.56 A number of interventions under the umbrella of 'creating meaningful interaction' demonstrate how meaningful interaction enhances speaking skills. In a secondary school context, Arslanyilmaz (2013) demonstrated that task-based CALL (with meaning-focused objectives) was more effective than form focused CALL (with objectives focused on word form) for improving production in terms of fluency, accuracy and complexity. Ho and Binh (2014), from the perspective of CLT, highlight how communicative grammar teaching can improve secondary students' oral production over a Grammar Translation Method (GTM). In this study, after being introduced to the grammar through exploration, guidance and comprehension check, the students in the experimental group took part in two further stages, a meaningful oral fluency task, and a production stage where students talked about their own lives. The students in the CLT group performed significantly better in an oral post-test than the students taught using a GTM. Working with younger learners (7-11) from the perspective of TBL, and specifically on how type of task repetition affects speaking, García Mayo and Imaz Agirre (2016), found that there was no difference in the effect of the type of task in terms of whether it was task repetition (with the same task and content) or procedural repetition (with the same task and different content) on negotiation of meaning strategies (such as use of clarification and confirmation checks). However, they did find that procedural repetition had the effect of creating more collaborative interactional patterns amongst learners than task repetition.

3.57 Well-targeted game-based activities can also provide opportunities for meaningful interaction that improve speaking skills. Focusing on primary school learners (mean age 7.41 years), Griva and Semoglou (2012) found that participating in physical activities and role play games provided a real reason for children to use the target language. The group taught in this way were found to be more effective in communicative activities including both listening and speaking than the experimental group where a PPP (presentation, practice and production) teaching approach was used. Also working with primary school learners, Young and Wang

(2014) investigated the different effects of drill-based or game-based CALL pronunciation practice, and found that learners' pronunciation showed a significant improvement when using the game-based platform. One key advantage of the game-based activities over the drill was that they afforded opportunities for lively interaction with peers. Though opportunities for meaningful interaction are highlighted in both these studies, it should be noted that the teaching activities and the games used were carefully targeted to suit the relevant learning objectives.

3.58 Three of the studies that explore the positive effects of meaningful interaction use CLIL as the independent variable. That is, the students use the language as a vehicle for learning content, so are engaged in this way in meaningful interaction. Fluency and accuracy are particular aspects of speech that seem to be higher for CLIL than non-CLIL students. Gallardo del Puerto and Gómez Lacabex (2013) report that in a story-telling task, CLIL students' speech was found to be more fluent and characterised by better grammar and vocabulary, supporting previous findings that CLIL instruction leads to more fluent and accurate oral production than non-CLIL instruction. Pérez Cañado and Lancaster (2017) report a CLIL group performing better on a spoken interaction task than non-CLIL students, with a greater propensity to employ more sophisticated structures in their language.

3.59 The above two studies report evidence that one area of oral production does not appear to be benefited by CLIL: both report pronunciation as no better for students who had been educated through CLIL than those who had not. Gallardo del Puerto, Gómez Lacabex and García Lecumberri (2009) similarly found no statistical difference between CLIL and non-CLIL learners in terms of degree of foreign accent and intelligibility, but alongside this report that CLIL students' foreign accent was rated as significantly 'less irritating' by judges. The authors note that there could be non-phonological factors accounting for the 'irritability' judgement. One reason that pronunciation differs from other aspects of oral production amongst CLIL learners can be related to the input the learners receive and how 'native-like' this is. CLIL learners in Spain, where the studies reported here took place, are often taught by non-native speakers and so there may be a 'lack of authentic input' (Gallardo del Puerto et al., 2009).

3.60 Several studies investigate the effectiveness of computer-mediated interaction on learning. Satar and Ozdener (2008), working in the context of a secondary school, showed that synchronous computer-mediated communication, whether voice chat or text chat, had positive effects on both spoken production and anxiety levels. It

was noted that text chat may be more appropriate to lower level learners due to more time being afforded to construct utterances while higher proficiency learners can gain more from voice chat to improve fluency. Another important finding from this research is that the learners preferred to work with non-native speaker (NNS) interactants. For younger learners, robots with speech recognition software can be an effective means for students to improve conversation skills in a foreign language. Wu, Wang and Chen (2015) investigate the effect of meaningful communication with a single teaching assistant robot in a class, which uses forms of total physical response, communicative language teaching and storytelling techniques. He reports improvement in learning outcomes (as well as motivation/confidence) of children (aged 8-9) who worked with the robot compared to a control group . Working with primary aged children (7-9) in a study that allows children to have more one-to one-interaction time with a robot, Wang, Young and Jang (2013) report the benefits of students working with tangible learning companions that have the appearance of soft toys. Though it is reported that the students' performance improved, the robots used in this study used the audio-lingual method, and were suitable only for more basic interactions with beginner level students. Crucially, all of these studies emphasise the importance of creating an environment where students are free to interact and make mistakes.

Pronunciation

3.61 Training for pronunciation of individual words was found to be as effective using ASR (automatic speech recognition) as using a teacher for 11-year olds, even where time for training with the computer amounted to half that with the teacher (Neri, Mich, Gerosa, & Giuliani, 2008). The system used in this study trained children on individual words and provided a simple reject/ accept response. Authors suggest that the positive effect of the intervention on the experimental group is that the children training with the computer assisted programme had the 'undivided' attention of the programme. While not comparing CALL with a teacher, Young and Wang (2014) similarly note the value of ASR software. Their dual game- and drill-based system enabled 7-9 year olds to repeat a level if word-level pronunciation was incorrect. The game-based activities were shown to improve the children's pronunciation over the drill activities, but the dual system was important in allowing students of different levels choice in their study and the ability to return to material they were unsure of.

3.62 One aspect of pronunciation activities that proves particularly beneficial to learners is the opportunity to repeat tasks over time. This is evident in the two CALL studies noted above, and improved precision in pronunciation by primary aged children using CALL for a a read-aloud task was shown by Nutta, Feyten, Norwood, Meros, Yoshii and Ducher (2002). In this study, the children worked with an interactive story and used the computer to record their own versions of the narration. They were able to use the recording function on the computer to pause and rerecord their output, leading to more repeats and greater precision. Repetition of task is equally possible in a non-CALL environment. In Jung, Kim and Murphy's (2017) study on collaborative priming tasks with adolescents, task repetition (with the same procedure and content) was as effective in an immediate post-test on the production of lexical stress as procedural repetition (repetition of the same task but with different content). However, in the delayed post-test, the task repetition was twice as effective as the procedural repetition. This suggests that a) hearing and producing the same words containing the target-stress patterns repeatedly, coupled with collaborative meaning-focused tasks, aided the learners in improving lexical stress, and b) knowledge of word stress might be more fully proceduralized and automatized through repeated practice with the same words. The authors note how, as was reported by Trofimovich, McDonough and Foote (2014) these findings provide evidence of the learning value of interaction-based alignment in pronunciation (where learners can be guided by the pronunciation of peers).

Listening

3.63 Studies focusing on types of listening input evidence how supported authentic material, adapted materials and moderated teacher input can all aid comprehension. Vulchanova, Aurstad, Kvitnes and Eshuis's (2015) study of secondary school students watching L2 animations demonstrated that subtitles in the target language aid comprehension over L1 subtitles.[13] This, the authors note, contradicts earlier research where the use of L1 subtitles was shown to be more beneficial (for example, Bianchi & Ciabattoni, 2008; Guillory, 1998). A possible explanation given for this apparent contradiction is that Norwegian learners in Vulchanova et al.'s study may have been more advanced than learners previously studied. Nevertheless, findings regarding the L2 as the best language for subtitles should be treated with caution. Also evidencing the importance of multiple input

[13] Different results were found for learners of over 16 included in Vulchanoa et al.'s study but only the results for those learners falling into the relevant age group are reported here.

channels, Cabrera and Martínez (2001) investigated strategies used by teachers to modify spoken input in the primary classroom. They found that while input from the linguistic point of view is important, there is no improvement in comprehension without simultaneous interactive adjustment (such as repetitions, comprehension checks and gestures). This theme of interactivity as well as modification in terms of linguistic input for young learners is further supported by Verdugo and Belmonte (2007) who found that though the use of multimedia stories did improve learners' (aged 6) comprehension, the stories need to be of a slow enough pace for learners.

3.64 There is evidence in the studies that pre-listening activities are important in aiding listening comprehension and that some types of pre-listening activity are more effective than others. Li, Wu and Lin (2017), in order to prepare 14-16 year old students for a listening task, used an interactive brainstorming technique (think-pair-share), developed by Lyman (1981). They found that students who used a collaborative technique for prediction outperformed the control group who did not. In terms of the type of brainstorming activity, those students who used picture brainstorming in the first activity did better than those who used brainstorming with words. With a similar age group, Rouhi, Nabavi and Mohebbi (2014) likewise found that topic preparation was effective in aiding comprehension, as was repeated listening. However, the study found that previewing questions did not seem to have any effect on listening comprehension scores.

Speaking and listening strategies

3.65 Tisma (2016) reports beneficial effects of strategy instruction (SI) on listening comprehension and on the metacognitive awareness of 14-16 year olds after five weeks of training. Also working with secondary age children, Harris (2007) in preliminary findings of a study on strategy interaction found SI had a positive influence on performance and motivation. However, the study suggests that SI needs to be tailored effectively to ability level as students found the listening strategies they were taught to be complex. Nutta et al. (2002), working with primary age children, found that while the CALL approach they were investigating showed no significant learning gains for students in terms of language production, what did improve, as a result of the opportunities the CALL approach afforded the children, were language learning strategies. Learners spent longer perfecting their answers when reading aloud, highlighting that as well as teaching language strategies, providing opportunities for learners to develop language learning strategies can have beneficial effects on output.

Summary of Speaking and Listening Competence findings

- Approaches that create meaningful interaction more positively affect speaking skills than activities where the focus is on form.
- Game-based activities (in contrast to e.g. drill-based) have a positive effect on pronunciation and communicative achievement.
- Spoken production (though not necessarily pronunciation) tends to be more advanced among CLIL than non-CLIL students. However, this cannot be attributed to CLIL alone.
- Online interaction and interaction with robots can be equally as effective as face-to-face interaction in creating an effective learning environment for spoken interaction.
- Opportunities to interact and to make language errors are important to the development of speaking skills.
- Use of automatic speech recognition (ASR) technology to provide feedback aids pronunciation and is more efficient but not necessarily more effective than a teacher.
- Task repetition aids pronunciation.
- Supporting authentic listening material (even with L2 input in a different medium) and moderating input both aid comprehension.
- Pre- listening support aids comprehension.
- Strategy Instruction or providing opportunities for students to develop strategies for speaking and listening improves both competences.

General language competence

Table 3.7: Summary of items (general language competence)

Number of items identified at Phase 2	89
Number of items not sourced	11
Number of items excluded from evidence synthesis at Phase 3	58
Number of items included	20

3.66 The previous sections have addressed the effectiveness of approaches and methods for developing specific aspects of learner competence. Twenty of the items meeting the inclusion criteria for this REA relate to multiple aspects of competence, or to general competence in the L2. This section includes a relatively high proportion of grey literature, in the form of commissioned pedagogic reviews and reports. It is important to note that the synthesis below only refers to findings from those reports that are within the parameters of this REA; other important themes emerging from those reports include the impact of: teachers' competence and confidence in target language, and consequent opportunities for learners to use/hear target language spontaneously (Cable et al., 2010; Driscoll et al., 2004; Ofsted, 2011); a whole school policy promoting the teaching of the target language (Cable et al., 2010; Driscoll et al., 2004); effective teacher training and ongoing skills development (Cable et al., 2010; Lo Bianco, 2009); motivating learners and making the language relevant to them in terms of their interests, community and identity (Lo Bianco, 2009; Edelenbos, Johnstone, & Kubanek, 2007; Ofsted, 2011); second language literacy development to sustain L2 proficiency (Harris & Ó Duibhir, 2011).These themes are discussed further in Section 4 "Excluded literature and future research imperatives".

3.67 Five themes with direct relevance to effective methods and approaches for general L2 proficiency development emerge from the literature: i) CLIL and partial CLIL (6 studies); ii) Amount and distribution of instruction time (4 studies); iii) Isolated and integrated Form Focused Instruction (3 studies); iv) Technology and language learning (2 studies); v) Use of songs in the language classroom (3 studies).

CLIL and partial CLIL

3.68 The widespread adoption of CLIL methods in some national contexts offers opportunities for the analysis of very large data sets. De Diezmas (2016) analyses English language test results in writing, oral production and interaction, reading, and listening, that were generated by a census of all schools in Castilla-La Mancha, including CLIL (>1,900 learners) and non-CLIL (>17,100 learners). All learners studied English in infant (270 hours) and primary (450 hours) education, and CLIL learners received an additional 250 hours (average) of target language. A significant difference was found between CLIL and non-CLIL learners in oral production and interaction, but not in the other three tests. Dalton-Puffer (2011) , synthesising research on learning outcomes of CLIL, also finds that CLIL learners outperform others in spontaneous oral production. Additionally, she finds that they

score higher on vocabulary tests, but adds the caveat that they often arrive in the programme with larger L2 vocabularies, and score high on frequency based vocabulary tests due to the subject-specific vocabulary they are exposed to. She finds that evidence of the effectiveness of CLIL for improving morpho-syntactic knowledge is mixed, and that evidence relating to the effect of CLIL on writing ability is unclear.

3.69 The confounding variables noted by Dalton-Puffer and others are addressed by Artieda, Roquet and Nicolás-Conesa (2017), who enhanced the robustness of their comparison of formal instruction with and without CLIL, by controlling for the effects of confounding variables identified in other CLIL studies: ability, motivation to study the L2, socio-economic status, age, and previous exposure to L2. CLIL in this study was 'partial', accounting for 2 hours per week of science class. The study found that in some aspects of performance (listening comprehension, writing accuracy), the older learners (13-14 years, n=50) outperformed younger learners (12-13 years, n=50) even when the latter, but not the former, received +CLIL instruction. When the learners in the +CLIL and -CLIL groups were matched for age, the +CLIL group scored significantly higher in assessments of reading comprehension, lexical richness, and linguistic and communicative competence.

3.70 A more effective example of partial CLIL is reported in Coral, Lleixà and Ventura (2016)'s analysis of state language competence assessments in multilingual schools in Catalonia. A key finding relates to the teaching of a Physical Education in CLIL (PE-in-CLIL) programme, which yields significant language gains. The authors identify critical contributors to this success: methods of teaching PE were changed to incorporate language learning techniques; co-operative learning was combined with meaningful hands-on tasks, to maximise opportunities for learner talking time.

3.71 Harris and Ó Duíbhir (2011) argue that there is evidence that CLIL has a positive impact on language proficiency and that it 'enables learners to encounter language in context and use it for authentic communication'. The Irish (L2) situation at primary school level favours the use of CLIL, which may be enacted through (i) informal use by teachers either in the Irish language classroom or in other subjects where there is use made of the language (ii) a wider programme of Irish medium provision in other subject areas (iii) partial immersion (up to fifty per cent of contact hours). All teachers at this level in Ireland have to have competence in Irish to fully qualify as a teacher in the primary sector. Edelenbos, Johnstone and Kubanek (2007) also call for partial/total immersion with more contact hours and intensity, finding that this

factor yields the greatest level of proficiency in the target language (with the caveat, however, that it is 'unlikely to be generally applicable').

Amount and distribution of instruction time

3.72 The intensity and timing of second language instruction, as well as the total amount of class time, is considered by a number of authors. Harris and Ó Duíbhir (2011) note as a main finding that short-term intensive language programmes are more effective than 'drip feed programmes' taking place over a longer time period. This is supported by the findings of Collins, Halter, Lightbown and Spada (1999), who compare language gains in two time distribution conditions: 2 hours of language class per day for 10 months, and the same number of contact hours condensed into 5 months (n=700). Language gains were evident in all learners, but those on the intensive programme scored higher across all language skills tested. They note as a possible confounding factor that the groups were not matched on academic achievement level, and findings from a replication by Collins and White (2011) add weight to this possible confound. Collins and White examined the acquisition of English by 11-year-old French L1 learners (n=230) in two 400-hour programmes, one delivered across 10 months and one concentrated into 5 months. Language development was compared across the two contexts four times via a battery of comprehension and production measures, and found that both groups improved, with no significant difference in learning outcomes between the groups.

3.73 Mitchell and Myles' (2018 (forthcoming)) report on 38 observed hours of language classes across one school year, for 8 year old English L1 learners of French. The classes took a predominantly oral approach, using age-appropriate games such as role-play, stories, songs and crafts. Post tests included a receptive vocabulary test (based on classroom input) and an Elicited Imitation (EI) test to measure general proficiency. Significant gains were made in the Elicited Imitation Test, but not in vocabulary. The authors consider these outcomes to be limited, and conclude that a time allowance of 38 hours per school year will make appropriate language gains challenging, regardless of the teaching approach and quality.

Isolated and integrated Form Focused Instruction

3.74 Communicative Language Teaching (CLT) approaches, including Content Based Instruction (CBI) and forms of CLIL, create opportunities for language development during meaning-focused activities. Form focused instruction (FFI) can be included before, after, or during CLT and CBI classes, and this necessitates consideration of

which is more effective: isolated or integrated FFI. Spada and Lightbown (2008) review evidence of the comparative effectiveness of these approaches, and conclude that each fulfils a different role. Integrated FFI, which might range from responsive feedback to planned, repeated elicitation of target language structures, can boost reliable usage of recently acquired language. Isolated FFI lessons might usefully focus on language elements identified as challenging (possibly due to L1 influence), but will be most beneficial when included in a programme of study alongside CLT and CBI classes. Somewhat in line with this, Harris and Ó Duíbhir (2011) call for balance between communicative and analytic approaches, noting that there are contradictory results regarding courses with a communicative orientation.

3.75 Spada and Lightbown (2008) report research evidence to suggest age differentials in learners' responses to FFI: older children will derive more benefit from it, whereas young learners sometimes need little or no FFI instruction. They do not specify the age threshold for this, but Elgün-Gündüz, Akcan and Bayyurt (2012) test this finding with 11-12 year olds, in an investigation of isolated and integrated FFI in primary school English classrooms in Turkey. They compare the vocabulary, grammar, and writing development of learners in a school implementing isolated FFI (n=50) and a school implementing integrated FFI (n=70). Learners receiving integrated FFI performed better in all measures than students receiving isolated FFI.

Technology and language learning

3.76 Data Driven Learning (DDL) does not, in theory, require technology support, but the strength of the approach is in its capacity to present learners with large data sets of multiple examples of target items, which they can explore to find patterns of use. Boulton and Cobb (2017) provide a systematic meta-analysis of 64 studies on the effectiveness of a corpus linguistics approach to second language learning, and find that the greatest gains happen when learners operate a concordance themselves, directly or through a CALL (Computer Assisted Language Learning) programme. DDL is found to be beneficial for learner language development, with large effect sizes reported for both between- and within-group comparisons. The authors reflect that while the approach, which often requires learners to investigate patterns of use for single vocabulary items, seems time consuming, it develops language sensitivity, noticing, and inductive skills, and encourages autonomous learning and engagement with authentic language.

3.77 Grgurović, Chapelle and Shelley (2013) focus on more traditional CALL methods in their meta-analysis of effectiveness studies on computer technology-supported language learning. They include 37 studies comparing approaches supported by and not supported by computer technology, and find a small but significant difference in favour of CALL instruction.

Use of songs in the language classroom

3.78 Driscoll et al (2004) 's systematic review of the characteristics of effective foreign language teaching found support for a ludic approach, including songs, in the primary school language classroom. However, Davis' (2017) review of evidence focusing specifically on the effectiveness of songs in teaching young language learners (3-12), finds that theory-driven links between songs and language development do not materialise strongly in practice. He reports some evidence that in addition to motivating learners, use of songs can promote vocabulary acquisition, and there is tentative evidence to suggest that integration of music into language learning can lead to communicative competence gains. Campfield and Murphy (2017), also cited in Davis, found that rhythm-salient input helped to develop structural knowledge of English in L1 Polish learners.

A note on approaches and methods

3.79 Many of the outputs included in this review, and in particular the commissioned reports featured in this section, explicitly state that the effectiveness of any method or approach is less influential than the skill and competence of the teacher delivering it. The Bauckham (2016) modern foreign languages pedagogy review concentrating on teaching practice in key stages 3 and 4 comes to the conclusion that 'no single approach to teaching languages represents 'the best way' in all circumstances'. Similarly, Edelenbos, Johnstone & Kubanek (2007) conclude that there is 'no clear evidence which relates different models of provision for ELL [Early Language Learning] to the outcomes of these models.' As noted at the beginning of this section, the reports included in our review draw attention to many factors affecting language learning in schools, but methods and approaches do not feature strongly in this list (Bauckham, 2016; Edelenbos et al., 2007; Harris & Ó Duibhir, 2011; Lo Bianco, 2009; Ofsted, 2011).

Summary of General Language Competence findings

- It is difficult to draw any strong conclusions in terms of the direct effect of CLIL as in most studies the amount of exposure to language was greater than for the control non-CLIL groups.
- Some evidence suggests that short-term intensive language programmes may be more effective than 'drip feed programmes'.
- Form focused instruction should be strategically deployed, and may be more beneficial to older children.
- Data driven learning can develop language sensitivity, noticing, and inductive skills, and encourages autonomous learning and engagement with authentic language.
- There is little evidence to suggest that use of songs in class promotes language uptake.
- Approaches and methods can be less influential on learning than factors such as teachers' language confidence, opportunities for language exposure, societal and educational context for learning, effective teacher training and ongoing skills development, motivating learners and making the language relevant to them in terms of their interests, community and identity. These themes are explored in more depth in Section 4.

Summary of findings applied to the Welsh context

3.80 In applying the findings from sections **Vocabulary competence** to **General language competence** to the context of Welsh language (L2) pedagogy (project objective C), the following key observations emerge:

- Maximising exposure to the language is common to many of the findings. This could be achieved by use of Welsh elsewhere in the curriculum (CLIL), within the school as part of a school language policy, or through the use of technology. Calls for CLIL such as Harris and Ó Duíbhir (2011)'s could be considered in the Welsh context although findings also suggest that CLIL may be more effective precisely because it tends to equate with an increase in contact hours. CLIL is shown to be most effective when combined with greater exposure to the language and when this happens consistently for a longer duration. If there is a subsequent reduction in contact hours or exposure to the language, then the benefits of CLIL can be eroded. On that basis, there would need to be an agreed

optimum number of teaching time/hours to maximise contact with Welsh consistently across all Key Stages.

- The educational context in which learning happens and incidental use of the language throughout the school were found to be influential factors. Therefore, any movement towards developing CLIL in the Welsh context would require negotiations with the wider school community to ensure support for developments of this nature, including a realisation that there would be implications for teachers' training / development.
- Oracy and literacy approaches to teaching and learning a language at primary level can both be effective in developing grammatical competence and writing skills. These can be supplemented by the incidental use of prosodic features at this level too.
- Meaningful interaction, particularly with an emphasis on use of authentic materials, has a positive effect on speaking skills, and may be just as effective when conducted on-line as face-to-face; this is important for teaching and learning in communities where there is less opportunity for face-to-face interaction. There is evidence that focus on form can also be beneficial, especially at Key Stages 3 and 4.
- Planned use of technology and data driven learning is shown to have potential in developing skills and promoting more general language awareness. In the Welsh context, this could also involve greater use of (and cooperation with) other authentic Welsh language sources such as S4C and Radio Cymru (broadcasts and web-based materials).
- There is a need to ensure that any advances in teaching Welsh at primary level are sustained and continue to be developed effectively at secondary level.
- Affective factors: although these are outside the scope of this REA, many findings refer to the importance of motivation and enjoyment in maximising achievement. It is clear, therefore, that the wider societal and educational context of learning Welsh as a second language and its association with learners' motivations(s) need further exploration. This is discussed further in the next section of the report.

4. Notes on excluded literature and future research imperatives

4.1 This section of the report deals briefly with a number of key psycholinguistic, sociolinguistic and educational issues which were excluded from the scope of the main report (see **Defining the research question**) but which are relevant to the future of effective Welsh language education and will merit continuing research attention. The issues discussed are:

- Bilingualism and the "cognitive advantage";
- Identity and language learning motivation;
- Cross linguistic relations, translanguaging,[14] and biliteracy;
- Effectiveness of immersion education.

Bilingualism and the "cognitive advantage"

4.2 There is a long historical debate over the relevance of bilingualism to intelligence, cognitive development and cognitive resilience. Recent research has argued for a bilingual advantage in certain types of executive functioning in the brain (e.g. Bialystok & Viswanathan, 2009; Grundy & Timmer, 2017), and for a protective effect of bilingualism during aging, including delayed onset of dementia symptoms (e.g. Bialystok et al., 2014). Considerable numbers of studies have supported these claims, including studies of later, non-balanced bilingualism (e.g. Vega-Mendoza et al., 2015), and these have naturally aroused considerable popular interest. However others have failed to find any clear cognitive advantage for bilinguals (including the Welsh-English bilingualism study of Gathercole et al., 2014). Critics of this research tradition have argued that the apparent bilingual advantage may be correlational rather than causal, and/or that it may be a result of social confounds, and/or of publication bias (de Bruin et al., 2015; Valian, 2014). However the proponents of "bilingual advantage" have pursued the debate, putting forward proposals on how the issues can be pursued and disentangled in a rigorous and better contextualised manner (e.g. Bak, 2016), for example through longitudinal research such as the Lothian Birth Cohort study (Bak et al., 2014). This is a key background issue for proponents of bilingualism in language education, and further research in locally relevant conditions is needed to inform longer term policy development.

[14] See Lewis, Jones, and Baker (2012) for definition and discussion of Translanguaging.

Identity and language learning motivation

4.3 In an age of global English, proponents of all types of education involving other languages must confront issues of attitudes, motivation and identity both within school systems and in the wider social context (Ushioda, 2017). European surveys of school students have consistently documented more positive motivation for English and less positive motivations for other foreign languages (Csizer & Lukacs, 2010; Busse, 2017); longitudinal research in Ireland has shown an association between decline in wider social support for Irish language instruction, and in primary school learners' achievement in the language (Harris et al., 2006). Regarding regional minority languages, motivation has been theorised differently by different researchers. A study of learners of Gaelic in Nova Scotia claims that learners are aiming to develop a "rooted" L2 self (MacIntyre et al., 2017) i.e. they are primarily motivated to become bilingual in order to transmit local cultural heritage. However, Lasagabaster (2017) argues that school aged learners of minority languages in Spain (Basque, Catalan), including new speakers, are moving beyond traditional exclusive language identities and developing a more "cosmopolitan" and plurilingual identity in which the regional language sits more comfortably alongside English and Spanish (and that this identity is reflected in increasing translanguaging behaviour). It is also known that attitudes (as well as achievement) of students in partial or total immersion programmes are much more favourable toward minority languages and related identities. This may be because populations are not identical (e.g. reflecting parental choice) but it nonetheless reinforces the vital importance of immersion programmes in the production of minority language speakers. For all types of learner, classroom experiences and relationships with languages teachers can influence motivation (Graham et al., 2016), and Dörnyei and associates have proposed a range of classroom motivational strategies grounded in research, which would merit further research investigation in a minority languages context (Dörnyei, 2001; Dörnyei & Kubanyiova, 2014). Lasagabaster (2017) also stresses the dynamism of language attitudes and motivation, and the need to track these systematically, for minority language education to meet the changing sociocultural expectations of young people (and to face challenges such as engagement with new migrant populations).

Cross linguistic relations, translanguaging and biliteracy

4.4 Contemporary theorists of language acquisition increasingly take the view that multilingualism is a "normal" condition and monolingualism the exception (Douglas Fir Group, 2016). Correspondingly they are interested in relations between language systems and mutual influences of different languages within the multilingual mind (Sharwood Smith & Truscott, 2014), with some theorists arguing strongly for conceptualising multilingual proficiency in terms of a single underlying dynamic system (Cook & Li Wei, 2016). These views have implications for language education, supporting the development of metalinguistic awareness, classroom translanguaging and the coordinated promotion of biliteracy as routes to ultimate effective performance in different languages (Leonet, 2017; Lyster, 2015; Moore & Sabatier, 2014). These educational ideas relate positively to the development of plurilingual identities incorporating positive views of minority languages, but their full application and evaluation would evidently require an innovative programme of interventionist classroom research. In particular educational evidence regarding the impact of L1 oracy and literacy development on L2 and vice versa, is complex and context dependent (Murphy et al., 2015; Shiel et al., 2010; Sparks, 2012).

Effectiveness of immersion education

4.5 The effectiveness of immersion education is a longstanding focus of bilingual education research including in Wales. Content learning in quality immersion programmes typically keeps pace (at least) with that in other programme types (Cheng et al., 2010; Shiel et al., 2010; Steele et al., 2017). Regarding language development, a common finding is that immersion education produces high level receptive skills in L2 and may also promote L2 oral fluency, but that grammatical accuracy in L2 production can lag behind by comparison. Various interventionist research studies have pursued this issue and made recommendations on styles of L2 pedagogy offering appropriate focus on form for immersion students (Lapkin & Swain, 2004; Ó Duibhir et al., 2016). Some Canadian researchers view biliteracy and elements of translanguaging in immersion contexts, when accompanied by metalinguistic reflection, as contributing to developing high level L2 proficiency longer term (Cummins, 2014; Lyster, 2015), while others argue for continuing language separation (Ballinger et al., 2017). Other researchers have explored the challenge facing immersion teachers of balancing content instruction and L1/L2 development (Cammerata & Tedick, 2012; Li et al., 2016; Turnbull et al., 2011), of responding to individual differences in immersion settings, and of resourcing high

quality early years immersion in minority languages (Hickey & De Mejia, 2014). The nature and adequacy of teacher qualifications in immersion education remains a general concern (Murphy & Evangelou, 2016). A number of issues require ongoing research in the area of immersion education in order to inform development of curriculum and pedagogy. These include: the quality of interaction in the immersion classroom, the dialogic construction of content knowledge and the extent and nature of translanguaging; children's engagement with classroom resources in all relevant languages; and the learning outcomes being achieved.

5. Concluding remarks

5.1 The research evidence synthesised in this report has emerged from a principled and systematic search and scrutiny of literature, and has met a set of strict inclusion criteria, and has been assessed as trustworthy, appropriate and relevant to the context of this REA. This process has controlled for the quality and relevance of the items included, but it is important to note that the distribution of those items across competences, the degree to which particular approaches and methods are represented, and the number of items addressing each approach, are dependent entirely on the outcome of the methods described in **Section 2. Methodology**. Distribution of items across competences is reasonably even. The method/approach themes, though, are not distributed across conventional categories. This is partly a function of the date inclusion criteria; by 2001 the "post methods" era was underway, and research on the global effectiveness of isolated methods was rare. It is possible that an interaction exists between research quality and type of method/approach investigated; research into CLIL, for example, has become increasingly sophisticated, and the co-existence of CLIL and non-CLIL language education systems (e.g. in Spain) makes for comparative studies with large sample sizes, and therefore statistically powerful research. It is also the case that CLIL is straightforwardly identified as a method (though not without definitional challenges), and one relevant to a bilingual context, and this too may explain why it emerged strongly from our searches.

5.2 The inclusion criteria we used favoured evidence emerging from high quality individual papers; because each output was screened and scrutinised in isolation, any evidence in collections of research papers which were individually of moderate quality, but which complemented and supported each other's findings, was not captured. Indeed, evidence reported in an REA will almost inevitably be defined by, and therefore limited to, the scope and distribution of published research within the specified timeframe, and once search and selection protocols are defined, there is little capacity to ensure balanced representation across themes and focuses as research meeting the inclusion criteria emerge. In this report we note that there is a dearth of studies targeting the teaching of lesser-used languages, and in contrast, studies with English as a target language are legion.

5.3 A major constraint of an REA is its rapidity; in this case, within five months a robust set of protocols and methods had to be established, and 5861 research outputs were processed. The time invested in honing procedural protocols was well spent, and supported the fast, consistent and high-quality decision-making required of the team. The time resource constraint necessitated a relatively narrow and exclusive focus on approaches and methods and on competence in this report. This must be acknowledged in interpretations of the findings we report; as we see in the papers scrutinised here and in **Section 4. Notes on excluded literature and future research imperatives**, methods and approaches are inextricably linked with affective, cognitive and societal factors, and these must be considered together in order to use the findings in a way that is ecologically valid.

5.4 That said, a number of evidence clusters emerge strongly from the REA. Effective language teaching requires a supportive environment, that facilitates exposure to the target language, including in non-traditional and creative ways. Attention must be paid to learners' linguistic landscape, and the societal and educational context for learning. Technological developments and tools offer valuable opportunities for language teaching and learning, but only if they are used in an informed way. Teacher education too should be strategic, well-informed from a language learning perspective, and carefully targeted, to produce teachers who are responsive to learning needs and processes. The effectiveness of a method or approach is critically dependent on the context in which it is used, and effective language education demands mastery of a comprehensive range of approaches and methods including those reported in this REA, and the ability to deploy them strategically.

References

Abdelrahman, L. A. M., Dewitt, D., Alias, N., & Rahman, M. N. A. (2017). Flipped learning for ESL writing in a Sudanese school. *Turkish Online Journal of Educational Technology, 16*(3), 60-70.

Agustín-Llach, M. P., & Canga Alonso, A. (2016). Vocabulary growth in young CLIL and traditional EFL learners: Evidence from research and implications for education. *International Journal of Applied Linguistics, 26*(2), 211-227.

Alcón, E. (2007). Incidental focus on form, noticing and vocabulary learning in the EFL classroom. *International Journal of English Studies, 7*(2), 41-60.

Arslanyilmaz, A. (2013). Computer-assisted foreign language instruction: task based vs. form focused. *Journal of Computer Assisted Learning, 29*(4), 303-318. doi:10.1111/jcal.12003

Artieda, G., Roquet, H., & Nicolás-Conesa, F. (2017). The impact of age and exposure on EFL achievement in two learning contexts: formal instruction and formal instruction + content and language integrated learning (CLIL). *International Journal of Bilingual Education and Bilingualism*, 1-19. doi:10.1080/13670050.2017.1373059

Bak, T. H. (2016). The impact of bilingualism on cognitive ageing and dementia. *Linguistic Approaches to Bilingualism, 6*(1-2), 205-226.

Bak, T. H., Nissan, J. J., Allerhand, M. M., & Deary, I. J. (2014). Does bilingualism influence cognitive aging? *Annals of Neurology, 75*(6), 959-963.

Ballinger, S., Lyster, R., Sterzuk, A., & Genesee, F. (2017). Context-appropriate crosslinguistic pedagogy: Considering the role of language status in immersion education. *Journal of Immersion and Content-based Language Education*, *5*(1), 35-57.

Bartan, Ö. S. (2017). The Effects of Reading Short Stories in Improving Foreign Language Writing Skills. *Reading Matrix: An International Online Journal, 17*(1), 59-74.

Bauckham, I. (2016). *Modern Foreign Languages Pedagogy Review: A review of modern foreign languages teaching practice in key stage 3 and key stage 4.* Teaching Schools Council. Retrieved from https://www.tscouncil.org.uk/modern-foreign-languages-report/

Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition*, *112*(3), 494–500.

Bialystok, E., Craik, F. I., Binns, M. A., Ossher, L., & Freedman, M. (2014). Effects of bilingualism on the age of onset and progression of MCI and AD: Evidence from executive function tests. *Neuropsychology*, *28*(2), 290.

Bianchi, F., & Ciabattoni, T. (2008). Captions and Subtitles in EFL Learning: an investigative study in a comprehensive computer environment. In A. Baldry, M. Pavesi, C. Taylor Torsello, & C. Taylor (Eds.), *From Didactas to Ecolingua* (pp. 69-90). EUT - Edizioni Università di Trieste.

Boulton, A., & Cobb, T. (2017). Corpus Use in Language Learning: A Meta-Analysis. *Language Learning, 67*(2), 348-393.

Boyd, S. (2001). *Teaching Modern Languages: Policy and Practice in England, Wales and Northern Ireland*. National Foundation for Educational Research. Retrieved from https://www.nfer.ac.uk/publications/91173/91173.pdf

Bruton, A. (2007). Vocabulary Learning from Dictionary Referencing and Language Feedback in EFL Translational Writing. *Language Teaching Research, 11*(4), 413-431.

Buga, R., Cāpeneatā, I., Chirasnel, C., & Popa, A. (2014). Facebook in foreign language teaching - a tool to improve communication competences. *Procedia - Social and Behavioral Sciences 128,* 93-98. doi: 10.1016/j.sbspro.2014.03.124

Busse, V. (2017). Plurilingualism in Europe: Exploring attitudes toward English and other European languages among adolescents in Bulgaria, Germany, the Netherlands, and Spain. *The Modern Language Journal, 101*(3), 566-582.

Cable, C., Driscoll, P., Mitchell, R., Sing, S., Cremin, T., Earl, J., . . . Heins, B. (2010). *Languages learning at key stage 2: A longitudinal study.* Department for Children, Schools and Families. Retrieved from http://oro.open.ac.uk/20630/

Cabrera, M. P., & Martínez, P. B. (2001). The effects of repetition, comprehension checks, and gestures on primary school children in an EFL situation. *ELT Journal, 55*(3), 281-288.

Cameron, L. (2001) *Teaching Languages to Young Learners.* Cambridge: Cambridge University Press.

Cammarata, L., & Tedick, D. J. (2012). Balancing Content and Language in Instruction: The Experience of Immersion Teachers. *The Modern Language Journal*, *96*(2), 251-269.

Camo, A. C., & Ballester, E. P. (2015). The effects of using L1 translation on young learners' foreign language vocabulary learning. *Elia-Estudios De Linguistica Inglesa Aplicada, 15*(15), 109-134. doi:10.12795/elia.2015.i15.06

Campfield, D. E., & Murphy, V. A. (2014). Elicited imitation in search of the influence of linguistic rhythm on child L2 acquisition. *System, 42*, 207-219.

Campfield, D. E., & Murphy, V. A. (2017). The influence of prosodic input in the second language classroom: does it stimulate child acquisition of word order and function words? *Language Learning Journal, 45*(1), 81-99. doi:10.1080/09571736.2013.807864

Chen, C. M., Tan, C. C., & Lo, B. J. (2016). Facilitating English-language learners' oral reading fluency with digital pen technology. *Interactive Learning Environments, 24*(1), 96-118. doi:10.1080/10494820.2013.817442

Chen, M. H., Tseng, W. T., & Hsiao, T. Y. (2018). The effectiveness of digital game-based vocabulary learning: A framework-based view of meta-analysis. *British Journal of Educational Technology, 49*(1), 69-77. doi:10.1111/bjet.12526

Cheng, L., Li, M., Kirby, J. R., Qiang, H., & Wade-Woolley, L. (2010). English language immersion and students' academic achievement in English, Chinese and mathematics. *Evaluation & Research in Education, 23*(3), 151-169.

Collins, L., & White, J. (2011). An intensive look at intensity and language learning. *Tesol Quarterly, 45*(1), 106-133. doi:10.5054/tq.2011.240858

Collins, L., Halter, R. H., Lightbown, P. M., & Spada, N. (1999). Time and the distribution of time in L2 instruction. *TESOL Quarterly, 33*(4), 655-680. doi:10.2307/3587881

Cook, V., & Li Wei (Eds.). (2016). *The Cambridge handbook of linguistic multi-competence*. Cambridge: Cambridge University Press.

Coral, J., Lleixà, T., & Ventura, C. (2016). Foreign language competence and content and language integrated learning in multilingual schools in Catalonia: an *ex post facto* study analysing the results of state key competences testing. *International Journal of Bilingual Education and Bilingualism*, *21*(2), 139-150. doi:10.1080/13670050.2016.1143445

Council of Europe / Conseil de l'Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment.* (Cambridge: Cambridge University Press).

Csizér, K., & Lukács, G. (2010). The comparative analysis of motivation, attitudes and selves: The case of English and German in Hungary. *System*, *38*, 1–13.

Cummins, J. (2014). Rethinking pedagogical assumptions in Canadian French immersion programs. *Journal of Immersion and Content-Based Language Education*, *2*(1), 3-22.

Dalton-Puffer, C. (2011). Content-and-Language integrated learning: From practice to principles? *Annual Review of Applied Linguistics, 31*, 182-204. doi:10.1017/S0267190511000092

Davis, G. M. (2017). Songs in the young learner classroom: A critical review of evidence. *ELT Journal, 71*(4), 445-455.

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science, 26*(1), 99-107.

de Diezmas, E. N. M. (2016). The impact of CLIL on the acquisition of L2 competences and skills in primary education. *International Journal of English Studies, 16*(2), 81-101.

de la Fuente, M. J. (2006). Classroom L2 vocabulary acquisition: Investigating the role of pedagogical tasks and form-focused instruction. *Language Teaching Research, 10*(3), 263-295.

Denison, S., & Xu, F. (2014). The origins of probabilistic inference in human infants. *Cognition, 130*(3), 335-347. doi:10.1016/j.cognition.2013.12.001

Dolean, D. D. (2014). Using the keyword method in the classroom: Is the interacting imagery necessary? *System, 45*, 17-26. doi:10.1016/j.system.2014.04.003

Dolean, D. D., & Dolghi, A. (2016). Teaching young FL learners new vocabulary: A comparison between the efficiency of Keyword Method and Total Physical Response. *International Journal of English Linguistics, 6*(6), 1-7. doi:10.5539/ijel.v6n6p1

Donaldson, G. (2015). *Successful Futures: Independent Review of Curriculum and Assessment Arrangements in Wales*. Retrieved from https://gov.wales/docs/dcells/publications/150225-successful-futures-en.pdf

Dornyei, Z. (2001). *Motivational strategies in the language classroom.* Cambridge: Cambridge University Press.

Dornyei, Z., & Kubanyiova, M. (2014). *Motivating learners, motivating teachers: Building vision in the language classroom*. Cambridge: Cambridge University Press.

Douglas Fir Group (2016). A transdisciplinary framework for SLA in a multilingual world. *The Modern Language Journal, 100*(S1), 19-47.

Driscoll, P., Jones, J., Martin, C., Graham-Matheson, L., Dismore, H., & Sykes, R. (2004). A systematic review of the characteristics of effective foreign language teaching to pupils between the ages 7 and 11. In *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

Edelenbos, P., Johnstone, R., & Kubanek, A. (2007). *The main pedagogical principles underlying the teaching of languages to very young learners. Languages for the children of Europe: Published research, good practice and main principles.* European Commission. Retrieved from http://ec.europa.eu/education/languages/pdf/doc425_en.pdf

Elgun-Gunduz, Z., Akcan, S., & Bayyurt, Y. (2012). Isolated form-focused instruction and integrated form-focused instruction in primary school English classrooms in Turkey. *Language Culture and Curriculum, 25*(2), 157-171. doi:10.1080/07908318.2012.683008

EPPI-Centre (2007) Review Guidelines for Extracting Data and Quality Assessing Primary Studies in Educational Research. Version 2.0 London: EPPI-Centre, Social Science Research Unit.

Fidaoui, D., Bahous, R., & Bacha, N. N. (2010). CALL in Lebanese elementary ESL writing classrooms. *Computer Assisted Language Learning, 23*(2), 151-168. doi:10.1080/09588221003666248

Fonseca-Mora, M. C., Jara-Jiménez, P., & Gómez-Domínguez, M. (2015). Musical plus phonological input for young foreign language readers. *Frontiers in Psychology, 6:*286. doi:10.3389/fpsyg.2015.00286

Gallardo del Puerto, F., & Gómez Lacabex, E. (2013). The impact of additional CLIL exposure on oral English production. *Journal of English Studies, 11*, 113-131.

Gallardo del Puerto, F., Gómez Lacabex, E., & García Lecumberri, M. L. (2009). Testing the effectiveness of content and language integrated learning in foreign language contexts: The assessment of English pronunciation. In Y. Ruiz de Zarobe & R. M. Jiménez Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe.* (pp. 63-80). Bristol: Multilingual Matters.

García Mayo, M. D. P., & Imaz Agirre, A. (2016). Task repetition and its impact on EFL children's negotiation of meaning strategies and pair dynamics: An exploratory study. *Language Learning Journal, 44*(4), 451-466. doi:10.1080/09571736.2016.1185799

Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Viñas Guasch, N., . . . Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology, 5*, 11.

Gené-Gil, M., Juan-Garau, M., & Salazar-Noguera, J. (2015). Development of EFL writing over three years in secondary education: CLIL and non-CLIL settings. *Language Learning Journal, 43*(3), 286-303.

Ghahremani-Ghajar, S. S., & Mirhosseini, S. A. (2005). English class or speaking about everything class? Dialogue journal writing as a critical EFL literacy practice in an Iranian high school. *Language, Culture and Curriculum, 18*(3), 286-299. doi:10.1080/07908310508668748

Gierlinger, E. M., & Wagner, T. A. (2016). The more the merrier - Revisiting CLIL-based vocabulary growth in secondary education. *Latin American Journal of Content & Language Integrated Learning, 9*(1), 37-63. doi:10.5294/laclil.2016.9.1.3

Government Social Research Service and Evidence for Policy and Practice information Centre (n.d.). *Rapid Evidence Assessment toolkit.* Retrieved from http://webarchive.nationalarchives.gov.uk/20140402164155/http://www.civilservice.gov.uk/networks/gsr/resources-and-guidance/rapid-evidence-assessment

Graham, S., Courtney, L., Marinis, T., & Tonkyn, A. (2017). Early language learning: The impact of teaching and teacher factors. *Language Learning, 67*(4), 922-958.

Graham, S., Courtney, L., Tonkyn, A., & Marinis, T. (2016). Motivational trajectories for early language learning across the primary–secondary school transition. *British Educational Research Journal, 42*(4), 682-702.

Grgurović, M., Chapelle, C. A., & Shelley, M. C. (2013). A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL, 25*(2), 165-198.

Griva, E., & Anastasiou, D. (2009). Morphological strategies training: The effectiveness and feasibility of morphological strategies training for students of English as a foreign language with and without spelling difficulties. *Journal of Writing Research, 1*(3), 199-223. doi:10.17239/jowr-2009.01.03.2

Griva, E., & Semoglou, K. (2012). Estimating the effectiveness and feasibility of a game-based project for early foreign language learning. *English Language Teaching, 5*(9), 33-44. doi:10.5539/elt.v5n9p33

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, *33*, 325-340.

Guillory, H. G. (1998). The effects of keyword captions to authentic French video on learner comprehension. *Calico Journal, 15*, 89-108.

Gündüz, Z. E., & Ünal, I. D. (2016). Effects of multiple intelligences activities on writing skill development in an EFL context. *Universal Journal of Educational Research, 4*(7), 1687-1697.

Hall, G. (2011). *Exploring English language teaching*. London: Routledge.

Hanan, R. E. (2015). *The effectiveness of explicit grammar instruction for the young foreign language learner: A classroom-based experimental study* (Unpublished doctoral dissertation). University of York, UK.

Harris, J., & Ó Duibhir, P. (2011). *Effective Language Teaching: A synthesis of research.* Dublin: National Council for Curriculum and Assessment.

Harris, V. (2007). Exploring Progression: Reading and Listening Strategy Instruction with Near-Beginner Learners of French. *Language Learning Journal, 35*(2), 189-204.

Hennebry, M., Rogers, V., Macaro, E., & Murphy, V. (2017). Direct teaching of vocabulary after listening: is it worth the effort and what method is best? *The Language Learning Journal, 45*(3), 282-300.

Hickey, T. M., & de Mejía, A.-M. (2014). Immersion education in the early years: a special issue. *International Journal of Bilingual Education and Bilingualism*, *17*(2), 131-143.

Ho, P. V. P., & Binh, N. (2014). The effects of communicative grammar teaching on students' achievement of grammatical knowledge and oral production. *English Language Teaching, 7*(6), 74-86.

Huang, S., Willson, V., & Eslami, Z. (2012). The effects of task involvement load on L2 incidental vocabulary learning: A meta-analytic study. *The Modern Language Journal, 96*(4), 544-557.

Hwang, W. Y., Chen, H. S. L., Shadiev, R., Huang, R. Y. M., & Chen, C. Y. (2014). Improving English as a foreign language writing in elementary schools using mobile

devices in familiar situational contexts. *Computer Assisted Language Learning, 27*(5), 359-378. doi:10.1080/09588221.2012.733711

Ibarrola, A. L. (2012). Faster and further morphosyntactic development of CLIL vs. EFL Basque-Spanish bilinguals learning English in high-school. *International Journal of English Studies, 12*(1), 79-96.

Ikeda, M. (2013). Does CLIL work for Japanese secondary school students? Potential for the 'weak' version of CLIL. *International CLIL Research Journal, 2*(1), 31-43.

Jiménez Catalán, R. M., & Ruiz de Zarobe, Y. (2009). The receptive vocabulary of EFL learners in two instructional contexts: CLIL versus non-CLIL instruction. In Y. Ruiz de Zarobe & R. M. Jiménez Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe* (pp.81-92). Bristol: Multilingual Matters.

Jung, Y., Kim, Y., & Murphy, J. (2017). The role of task repetition in learning word-stress patterns through auditory priming tasks. *Studies in Second Language Acquisition, 39*(2), 319-346.

Kumaravadivelu, B. (1994). The postmethod condition: (E)merging strategies for second/foreign language teaching. *TESOL quarterly*, *28*(1), 27-48. doi:10.2307/3587197

Lan, Y. J., Sung, Y. T., & Chang, K. E. (2009). Let us read together: Development and evaluation of a computer-assisted reciprocal early English reading system. *Computers & Education, 53*(4), 1188-1198.

Lan, Y. J., Sung, Y. T., Cheng, C. C., & Chang, K. E. (2015). Computer-supported cooperative prewriting for enhancing young EFL learners' writing performance. *Language Learning & Technology, 19*(2), 134-155.

Lapkin, S., & Swain, M. (2004). What underlies immersion students' production: The case of *avoir besoin de*. *Foreign Language Annals 37*(3): 349-355.

Lasagabaster, D. (2017). Language learning motivation and language attitudes in multilingual Spain from an international perspective. *The Modern Language Journal, 101*(3), 583-596.

Laufer, B. (2006). Comparing Focus on Form and Focus on FormS in second-language vocabulary learning. *Canadian Modern Language Review, 63*(1), 149-166.

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics, 29*(4), 694-716. doi:10.1093/applin/amn018

Lee, J. H., & Macaro, E. (2013). Investigating age in the use of L1 or English-only instruction: Vocabulary acquisition by Korean EFL learners. *Modern Language Journal, 97*(4), 887-901. doi:10.1111/j.1540-4781.2013.12044.x

Leonet, O., Cenoz, J., & Gorter, D. (2017). Challenging minority language isolation: Translanguaging in a trilingual school in the Basque country. *Journal of Language, Identity & Education*, *16*(4), 216-227.

Lesniewska, J., & Pichette, F. (2016). Songs vs. stories: Impact of input sources on ESL vocabulary acquisition by preliterate children. *International Journal of Bilingual Education and Bilingualism, 19*(1), 18-34.

Lewis, G., Jones, B., & Baker, C. (2012). Translanguaging: Origins and development from school to street and beyond. *Educational Research and Evaluation*, *18*(7), 641-654.

Li, C., Wu, M., & Lin, W. (2017). The use of a “Think-Pair-Share” brainstorming advance organizer to prepare learners to listen in the L2 classroom. *International Journal of Listening*, 1-14. doi:10.1080/10904018.2017.1394193

Li, J., Steele, J., Slater, R., Bacon, M., & Miller, T. (2016). Teaching Practices and Language Use in Two-Way Dual Language Immersion Programs in a Large Public School District. *International Multilingual Research Journal, 10*(1), 31-43.

Lichtman, K. (2013). Developmental comparisons of implicit and explicit language learning. *Language Acquisition: A Journal of Developmental Linguistics, 20*(2), 93-108.

Lin, L. C. (2014). Learning word meanings from teachers' repeated story read-aloud in EFL primary classrooms. *English Language Teaching, 7*(7), 68-81. doi:10.5539/elt.v7n7p68

Lo Bianco, J. (2009). *Second languages and Australian schooling*. Melbourne, Australia: Australian Council for Educational Research.

Luan, N. L., & Sappathy, S. M. (2011). L2 vocabulary acquisition: The impact of negotiated interaction. *GEMA Online Journal of Language Studies, 11*(2), 5-20.

Lyman, F. T. (1981). The responsive classroom discussion: The inclusion of all students. In A. Anderson (Ed.), *Mainstreaming Digest.* (pp. 109–113). College Park, MD: University of Maryland Press.

Lyster, R. (2015). Using form-focused tasks to integrate language across the immersion curriculum. *System, 54*, 4-13.

Macaro, E., & Erler, L. (2008). Raising the achievement of young-beginner readers of French through strategy instruction. *Applied Linguistics, 29*(1), 90-119. doi:10.1093/applin/amm023

Macaro, E., & Mutton, T. (2009). Developing reading achievement in primary learners of French: Inferencing strategies versus exposure to 'graded readers'. *Language Learning Journal, 37*(2), 165-182. doi:10.1080/09571730902928045

MacIntyre, P. D., Baker, S. C., & Sparling, H. (2017). Heritage passions, heritage convictions and the rooted L2 self: Music and Gaelic language learning in Cape Breton, Nova Scotia. *Modern Language Journal*, *101*, 501–516.

Manoli, P., Papadopoulou, M., & Metallidou, P. (2016). Investigating the immediate and delayed effects of multiple-reading strategy instruction in primary EFL classrooms. *System, 56*, 54-65. doi:10.1016/j.system.2015.11.003

Martínez Adrián, M., & Gutiérrez Mangado, M. J. (2009). The Acquisition of English syntax by CLIL learners in the Basque country. In Y. Ruiz de Zarobe & R. M. Jiménez Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe* (pp. 176-196). Bristol: Multilingual Matters.

Martínez, A. G., & de Zarobe, Y. R. (2017). Comparing the benefits of a metacognitive reading strategy instruction programme between CLIL and EFL primary school students. *Elia, 17*, 71-92. doi:10.12795/elia.2017.i17.04

Mavilidi, M. F., Okely, A. D., Chandler, P., Cliff, D. P., & Paas, F. (2015). Effects of integrated physical exercises and gestures on preschool children's foreign language vocabulary learning. *Educational Psychology Review, 27*(3), 413-426. doi:10.1007/s10648-015-9337-z

Maxwell-Reid, C. (2010). Content and Language Integrated Learning (CLIL): the influence of studying through English on Spanish students' first-language written discourse. *Text & Talk, 30*(6), 679-699. doi:10.1515/text.2010.033

Merikivi, R., & Pietilä, P. (2014). Vocabulary in CLIL and in mainstream education. *Journal of Language Teaching & Research, 5*(3).

Mistar, J., Zuhairi, A., & Yanti, N. (2016). Strategies training in the teaching of reading comprehension for EFL learners in Indonesia. *English Language Teaching, 9*(2), 49-56.

Mitchell, R., & Myles, F. (forthcoming, 2018). Foreign language learning in an Anglophone setting: classroom engagement and attainable learning outcomes. *Apples: Applied Language Studies*.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement*. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

Moore, D. & Sabatier, C. (2014). Les approches plurielles et les livres plurilingues. De nouvelles ouvertures pour l'entrée dans l'écrit en milieu multilingue et multiculture. *Nouveaux cahiers de la recherche en education*, *17*(2), 32-65.

Murphy, V. A., & Evangelou, M. (Eds.) (2016). *Early childhood education in English for speakers of other languages*. London: British Council.

Murphy, V.A., Macaro, E., Alba, S. & Cipolla, C. (2015). The influence of L2 learning on first language literacy skills. *Applied Psycholinguistics, 36,* 1133-1153.

Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008). The effectiveness of computer assisted pronunciation training for foreign language learning by children. *Computer Assisted Language Learning, 21*(5), 393-408. doi:10.1080/09588220802447651

Ngo, C. M., & Trinh, L. Q. (2011). Lagging behind writing pedagogical developments: the impact of implementing process-based approach on learners' writing in a Vietnamese secondary education context. *Journal on English Language Teaching, 1*(3), 60-71.

Nutta, J. W., Feyten, C. M., Norwood, A. L., Meros, J. N., Yoshii, M., & Ducher, J. (2002). Exploring new frontiers: What do computers contribute to teaching foreign languages in elementary school? *Foreign Language Annals, 35*(3), 293-306. doi:10.1111/j.1944-9720.2002.tb01855.x

Ó Duibhir, P., Ní Dhiorbháin, A., & Cosgrove, J. (2016). An inductive approach to grammar teaching in grade 5 & 6 Irish immersion classes. *Journal of Immersion and Content-Based Language Education*, *4*(1), 33-58.

Ofsted. (2011). *Modern languages: Achievement and challenge 2007–2010*. Retrieved from https://www.gov.uk/government/publications/modern-languages-achievement-and-challenge-2007-to-2010.

Pérez Cañado, M. L., & Lancaster, N. K. (2017). The effects of CLIL on oral comprehension and production: A longitudinal case study. *Language, Culture and Curriculum, 30*(3), 300-316.

Petticrew, M., & Roberts, H. (2006). *How to appraise the studies: an introduction to assessing study quality*. Oxford: Blackwell.

Porter, A. (2014). *An early start to French literacy: Learning the spoken and written word simultaneously in English primary schools.* (Unpublished doctoral dissertation). University of Southampton, UK.

Porter, A. (2016). A helping hand with language learning: teaching French vocabulary with gesture. *Language Learning Journal, 44*(2), 236-256. doi:10.1080/09571736.2012.750681

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge: Cambridge University Press.

Richardson, W. S., Wilson, M. C., Nishikawa, J., & Hayward, R. S. (1995). The well-built clinical question: a key to evidence-based decisions. *ACP Journal Club*, *123*(3), A12-A12.

Rouhi, A., Nabavi, S.-M., & Mohebbi, H. (2014). The effects of previewing questions, repetition of input, and topic preparation on listening comprehension of Iranian EFL learners. *Iranian Journal of Language Teaching Research, 2*(2), 73-85.

Saladrigues, G., & Llanes, A. (2014). Examining the impact of amount of exposure on L2 development with CLIL and non-CLIL teenage students. *Sintagma, 26*, 133-147.

Satar, H. M., & Ozdener, N. (2008). The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat. *Modern Language Journal, 92*(4), 595-613.

Sercu, L. (2013). Weblogs in foreign language education: real and promised benefits. In L. G. Chova, A. L. Martínez, & I. C. Torres (Eds.), *7th International Technology, Education and Development Conference* (pp. 4355-4366).

Shafqat, A., Idrees, M., & Gujjar, A. A. (2009). Effectiveness of layered instructional strategy for teaching English at secondary level. *i-Manager's Journal of Educational Technology, 6*(1), 82.

Sharwood Smith, M., & Truscott, J. (2014). *The multilingual mind: A modular processing perspective*. Cambridge: Cambridge University Press.

Shiel, G., Gilleece, L., Clerkin, A., & Millar, D. (2011). *The 2010 National Assessments of English reading and mathematics in Irish-medium schools: Summary Report.* Dublin: Education Research Centre.

Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research, 15*(2), 137-158. doi:10.1177/1362168810388692

Shintani, N. (2012). Repeating input-based tasks with young beginner learners. *RELC Journal, 43*(1), 39-51. doi:10.1177/0033688212439322

Shintani, N. (2013). The effect of *focus on form* and *focus on forms* instruction on the acquisition of productive knowledge of L2 vocabulary by young beginning-level learners. *Tesol Quarterly, 47*(1), 36-62. doi:10.1002/tesq.54

Spada, N., & Lightbown, P. M. (2008). Form-focused instruction: Isolated or integrated? *TESOL quarterly, 42*(2), 181-207.

Sparks, R. L. (2012). Individual Differences in L2 Learning and Long-Term L1–L2 Relationships. *Language Learning*, *62*, 5-27.

Sylven, L. K. (2010). Teaching in English or English teaching? On the effects of content and language integrated learning on Swedish learners' incidental vocabulary acquistion. *Gothenburg Studies in English, 97*, 1-294.

Takeda, C. (2002). The application of phonics to the teaching of reading in junior high school English classes in Japan. *TESL Reporter, 35*(2), 16-36.

Tammenga-Helmantel, M., Arends, E., & Canrinus, E. T. (2014). The effectiveness of deductive, inductive, implicit and incidental grammatical instruction in second language classrooms. *System, 45*, 198-210. doi:10.1016/j.system.2014.06.003

Tammenga-Helmantel, M., Bazhutkina, I., Steringa, S., Hummel, I., & Suhre, C. (2016). Comparing inductive and deductive grammatical instruction in teaching German as a foreign language in Dutch classrooms. *System, 63*, 101-114. doi:10.1016/j.system.2016.09.002

Taylor, A., Lazarus, E., & Cole, R. (2005). Putting languages on the (drop down) menu: innovative writing frames in modern foreign language teaching. *Educational Review, 57*(4), 435-455. doi:10.1080/00131910500278322

Tišma, D. J. (2016). Improving listening comprehension skills relying on metacognitive strategies – Focus on vocabulary and specific L2 instruction. *Zbornik Instituta za Pedagoska Istrazivanja, 48*(2), 301-320. doi:10.2298/zipi1602301j

Tode, T. (2007). Durability problems with explicit instruction in an EFL context: the learning of the English copula "be" before and after the introduction of the auxiliary "be". *Language Teaching Research, 11*(1), 11-30.

Toth, P. D., & Guijarro-Fuentes, P. (2013). The impact of instruction on second-language implicit knowledge: Evidence against encapsulation. *Applied Psycholinguistics, 34*(6), 1163-1193.

Tragant, E., Marsol, A., Serrano, R., & Llanes, A. (2016). Vocabulary learning at primary school: a comparison of EFL and CLIL. *International Journal of Bilingual Education and Bilingualism, 19*(5), 579-591. doi:10.1080/13670050.2015.1035227

Trofimovich, P., McDonough, K., & Foote, J. A. (2014). Interactive alignment of multisyllabic stress patterns in a second language classroom. *TESOL Quarterly, 48*(4), 815-832.

Tsiriotakis, I. K., Vassilaki, E., Spantidakis, I., & Stavrou, N. A. M. (2017). The examination of the effects of writing strategy-based procedural facilitative environments on students' English foreign language writing anxiety levels. *Frontiers in Psychology, 7*. doi:10.3389/fpsyg.2016.02074

Türk, E., & Erçetin, G. (2014). Effects of interactive versus simultaneous display of multimedia glosses on L2 reading comprehension and incidental vocabulary learning. *Computer Assisted Language Learning, 27*(1), 1-25.

Turnbull, M., Cormier, M., & Bourque, J. (2011). The first language in science class: A quasi-experimental study in late French immersion. *The Modern Language Journal, 95*, 182-198.

Ushioda, E. M. A. (2017). The impact of global English on motivation to learn other languages: Toward an ideal multilingual self. *The Modern Language Journal, 101*(3), 469-482.

Valian, V. (2014). Bilingualism and cognition. *Bilingualism: Language and Cognition, 18*(1), 3-24.

Vega-Mendoza, M., West, H., Sorace, A., & Bak, T. H. (2015). The impact of late, non-balanced bilingualism on cognitive performance. *Cognition, 137*, 40-46.

Verdugo, D. R., & Belmonte, I. A. (2007). Using digital stories to improve listening comprehension with Spanish young learners of English. *Language Learning & Technology, 11*(1), 87-101.

Vulchanova, M., Aurstad, L. M. G., Kvitnes, I. E. N., & Eshuis, H. (2015). As naturalistic as it gets: subtitles in the English classroom in Norway. *Frontiers in Psychology, 5*. doi:10.3389/fpsyg.2014.01510

Wang, Y. H., Young, S. S. C., & Jang, J.-S. R. (2013). Using tangible companions for enhancing learning English conversation. *Educational Technology & Society, 16*(2), 296-309.

Welsh Government (2015). *A curriculum for Wales – a curriculum for life*. Retrieved from http://gov.wales/docs/dcells/publications/151021-a-curriculum-for-wales-a-curriculum-for-life-en.pdf

Welsh Government (2017). *Cymraeg 2050: Welsh language strategy.* Retrieved from http://gov.wales/docs/dcells/publications/170711-welsh-language-strategy-eng.pdf

Welsh Government (2017). *Education in Wales: Our national mission.* Retrieved from http://gov.wales/docs/dcells/publications/170926-education-in-wales-en.pdf

Williams, N., & Thomas, E. M. (2017). Exploring minority language input sources as means of supporting the early development of second language vocabulary and grammar. *Applied Psycholinguistics, 38*(4), 855-880. doi:10.1017/s0142716416000473

Wu, W. C. V., Wang, R. J., & Chen, N. S. (2015). Instructional design using an in-house built teaching assistant robot to enhance elementary school English-as-a-foreign-language learning. *Interactive Learning Environments, 23*(6), 696-714. doi:10.1080/10494820.2013.792844

Young, S. S.-C., & Wang, Y.-H. (2014). The game embedded CALL system to facilitate English vocabulary acquisition and pronunciation. *Educational Technology & Society, 17*(3), 239-251.

Yunus, M. M., Nordin, N., Salehi, H., Embi, M. A., & Salehi, Z. (2013). The use of information and communication technology (ICT) in teaching ESL writing skills. *English Language Teaching, 6*(7), 1-8. doi:10.5539/elt.v6n7p1

Annex A: Project team members, roles and affiliations

Table A.1: Project team members, roles and affiliations

Name	Role	Institution
Tess Fitzpatrick	Project manager; author	Swansea University
Steve Morris	Deputy project manager; author	Swansea University
Lizzy Tanguay	Research Assistant; author	Swansea University
Jenny Needs	Research Assistant; author	Swansea University
Bethan Tovey	Research Assistant; author	Swansea University
Tony Clark	Research Assistant; author	Swansea University
Emma Marsden	Consultant; advisory board member	University of York
Imma Miralpeix	Consultant; advisory board member	University of Barcelona
Ros Mitchell	Consultant; advisory board member; author	University of Southampton
Vicky Murphy	Consultant; advisory board member	Oxford University
Enlli Thomas	Consultant; Quality Assurer	Bangor University
Kevin McManus	advisory board member	Pennsylvania State University
Pádraig Ó'Duibhir	advisory board member	Dublin City University
Alison Wray	advisory board member	Cardiff University
Fiona Hardy	advisory board member (teacher education)	Swansea University
Alex Lovell	advisory board member (CLIL)	Swansea University
Cornelia Tschichold	advisory board member (CALL)	Swansea University
Vivienne Rogers	advisory board member	Swansea University

In addition to the team members listed above, we would like to acknowledge the support of Federica Barbieri, Ian Glen, Susan Glen, Sam Oakley and all colleagues in the Department of Applied Linguistics and the Department of Welsh at Swansea University who facilitated work on this research project.

Annex B: Data Extraction Form (adapted from EPPI-Centre 2007)

Table B.1: Data Extraction Form (adapted from EPPI-Centre 2007)

Data management details	**A**
A.1 Identifier *Your initials + number of item LT1, LT2, LT3 etc*	A.1.1 details
A.2 Date of review	A.2.1 details yyyy/mm/dd
A.3 Paper citation details	A.3.1 APA citation A.3.2 doi if available
A.4 Status	A.4.1 Published - Please use this keyword if the report has an ISBN or ISSN number. A.4.2 Published as a report or conference paper - Please use this code for reports which do not have an ISBN or ISSN number (e.g. 'internal' reports; conference papers) A.4.3 Unpublished - e.g. thesis or author manuscript
A.5 Language of publication (please specify)	A.5.1 English A.5.2 Welsh A.5.3 Other – details
Study aims and context	**B**
B.1 Study aims (explicit/implicit/unclear) (please specify – write in author's own description if there is one). Elaborate on which aspects are reviewer's interpretation.	B.1.1 explicit B.1.2 implicit B.1.3 unclear
B.2 Which language learning contexts is this relevant to?	B.2.1.heritage language (learning context is a region where a different language is dominant) B.2.2 second language (learning context is region where the target language is dominant) B.2.3 foreign language (learning context is a region where target

	language is not widely used) B.2.4 Other – please specify
B.3 Which age group of learners is this relevant to?	B.3.1 3-5 B.3.2 5-7 B.3.3 7-11 B.3.4 11-14 B.3.5 14-16 B.3.6 16+
B.4 Is there a clearly-stated theoretical underpinning to the work?	B.4.1 not applicable B.4.2 yes (please specify) B.4.3 no
B.5 Funded (by whom)?	B.5.1 yes (please specify) B.5.2 no
B.6 What are the research questions/hypotheses? Use words of author(s) here	B.6.1 details
Study focus and sample	**C**
C.1 What is the focus of study according to the primary thematic categories (competences)?	C.1.1 i - spoken C.1.2 ii - written C.1.3 iii - reading C.1.4 iv - listening C.1.5 v - e-language/technology C.1.6 vi – metalinguistic skills C.1.7 vii – autonomous skills C.1.8 viii – willingness to communicate
C.2 What is the focus of study according to the thematic domains initially identified?	C.2.1 i C.2.2 ii C.2.3 iii C.2.4 iv C.2.5 v C.2.6 vi C.2.7 vii C.2.8 viii C.2.9 ix C.2.10 x

	C.2.11 Other – please specify
C.3 Was a specific method or approach investigated?	C.3.1 yes (please specify) C.3.2 no
C.4 What is the target language?	C.4.1 specify language
C.5 What is the learners' L1(s)?	C.5.1 specify language
C.6 How old were the participants at the time of the study?	C.6.1 3-5 C.6.2 5-7 C.6.3 7-11 C.6.4 11-14 C.6.5 14-16 C.6.6 16- 18 C.6.7 18+
C.7 What was the number of participants?	C.7.1 not applicable (e.g. non-empirical paper, study of policies, documents etc.) C.7.2 explicitly stated (please specify) C.7.3 implicit (please specify) C.7.4 not stated/ unclear (please specify)
C.8 What was the institutional setting for the study?	C.8.1 Not applicable (e.g. study of policies, documents, etc.) C.8.2 Nursery school / early years setting C.8.3 Primary school C.8.4 Secondary school C.8.5 Other (please specify)
C.9 What is known about ILDs within the sample (SES, specific learning needs, intellectual attainment, etc.)	C.9.1 Not applicable (e.g. study of policies, documents etc.) C.9.2 Explicitly stated (please specify) C.9.3 Implicit (please specify) C.9.4 Not stated/unclear (please specify)
If the study is of an intervention	**D**

D.1 Is an intervention being studied?	D.1.1 Not applicable (no programme or intervention, please go to section E) D.1.2 Yes (please specify) D.1.3 No (please specify) D.1.4 Not stated/ unclear (please specify)
D.2 Describe the intervention/variable being studied *Describe the intervention in detail, whenever possible copying the authors' description from the report word for word.*	D.2.1 Details
D. 3 What is the aim of the intervention?	D.3.1 Not stated D.3.2 Not explicitly stated (Write in, as worded by the reviewer) D.3.3 Stated (Write in, as stated by the authors)
D.4 What is the duration of the intervention?	D.4 details
D. 5 Who provides the intervention (teacher, parent, researcher, materials provider)?	D.5.1 teacher D.5.2 parent D.5.3 researcher, D.5.4 materials provider
D.6 What training/skills are needed to perform the intervention?	D.6 Details
Methods – Data analysis	**E**
E.1 Is the study cross-sectional or longitudinal?	E.1.1 not applicable E.1.2 cross-sectional E.1.3 longitudinal
E.2 When were outcome measures taken (pre-post-delayed)?	E.2.1 n/a E.2.2 pre E.2.3 post E.2.4 delayed
E.3 What is the methodology of study (group, case, cohort, systematic…)? *NB Studies may use more than one method – please code each method used and note the respective outcomes for each method.*	E.3.1 A Random experiment with random allocation to groups E.3.2 B Experiment with non-random allocation to groups E.3.3 C One group pre-post test

A= *i). compare two or more groups which receive different interventions or different intensities/levels of an intervention with each other; and/or with a group which does not receive any intervention at all* *AND* *ii) allocate participants (individuals, groups, classes, schools, LEAs etc.) or sequences to the different groups based on a fully random schedule (e.g. a random numbers table is used). If the report states that random allocation was used and no further information is given then please keyword as RCT. If the allocation is NOT fully randomised (e.g. allocation by alternate numbers by date of birth) then please keyword as a non-randomised controlled trial* ***B=****Please use this code if the evaluation compared two or more groups which receive different interventions, or different intensities/levels of an intervention to each other and/or with a group which does not receive any intervention at all BUT DOES NOT allocate participants (individuals, groups, classes, schools, LEAs etc.) or sequences in a fully random manner. This keyword should be used for studies which describe groups being allocated using a quasi-random method (e.g. allocation by alternate numbers or by date of birth) or other non-random method* ***C=****Please use this code where a group of subjects e.g. a class of school children is tested on outcome of interest before being given an intervention which is being evaluated. After receiving the intervention the same test is administered again to the same subjects. The outcome is the difference between the pre and post test scores of the subjects.* ***D=****Please use this code where one group of subjects is tested on outcome of interest after receiving the intervention which is being evaluated* ***E=****Please use this code where researchers prospectively study a sample (e.g. learners), collect data on the different aspects of policies or practices experienced by members of the sample (e.g. teaching methods, class sizes), look forward in time to measure their later outcomes (e.g. achievement) and relate the experiences to the outcomes achieved. The purpose is to assess the effect of the different experiences on outcomes.* ***F=****Please use this code where researchers compare two or more groups of individuals on the basis of their current situation (e.g. 16 year old pupils with high current educational performance compared to those with average educational performance), and look back in time to examine the statistical association with different policies or practices which they have experienced (e.g. class size; attendance at single sex or mixed sex schools; non-school activities etc.).* ***G=*** *please use this code where researchers have used a*	E.3.4 D one group post-test only E.3.5 E Cohort study E.3.6 F Case-control study E.3.7 G Statistical survey E.3.8 H Views study E.3.9 I Ethnography E.3.10 J Systematic review E.3.11 K Other review (non systematic) E.3.12 L Case study E.3.13 M Document study E.3.14 N Action research E.3.15 O Methodological study E.6.16 P Secondary data analysis

questionnaire to collect quantitative information about items in a sample or population e.g. parents views on education ***H=*** *Please use this code where the researchers try to understand a phenomenon from the point of the 'worldview' of a particular, group, culture or society. In these studies there is attention to subjective meaning, perspectives and experience'.* ***I=*** *please use this code when the researchers present a qualitative description of human social phenomena, based on fieldwork* ***J=*** *please use this code if the review is explicit in its reporting of a systematic strategy used for (i) searching for studies (i.e. it reports which databases have been searched and the keywords used to search the database, the list of journals hand searched, and describes attempts to find unpublished or 'grey' literature; (ii) the criteria for including and excluding studies in the review and, (iii) methods used for assessing the quality and collating the findings of included studies.* ***K=*** *Please use this code for cases where the review discusses a particular issue bringing together the opinions/findings/conclusions from a range of previous studies but where the review does not meet the criteria for a systematic review (as defined above)* ***L=*** *please use this code when researchers refer specifically to their design/ approach as a 'case study'. Where possible further information about the methods used in the case study should be coded* ***M=****please use this code where researchers have used documents as a source of data e.g. newspaper reports* ***N=****Please use this code where practitioners or institutions (with or without the help of researchers) have used research as part of a process of development and/or change. Where possible further information about the research methods used should be coded* ***O=****please use this keyword for studies which focus on the development or discussion of methods; for example discussions of a statistical technique, a recruitment or sampling procedure, a particular way of collecting or analysing data etc. It may also refer to a description of the processes or stages involved in developing an 'instrument' (e.g. an assessment procedure).* ***P=*** *Please use this code where researchers have used data from a pre-existing dataset to answer their 'new' research question*	
E.4 How were group comparisons achieved?	E.4.1 Not applicable (not more than one group) E.4.2 Prospective allocation into more than one group, e.g. allocation to different

	interventions, or allocation to intervention and control groups E.4.3 No prospective allocation but use of pre-existing differences to create comparison groups, e.g. receiving different interventions or characterised by different levels of a variable such as social class E.4.4 Other (please specify) E.4.5 Not stated/ unclear (please specify)
Methods – Data collection and analysis	**F**
F.2 How were the study samples identified/recruited?	F.2.1 details
F.3 Was consent sought?	F.3.1 Not applicable (please specify) F.3.2 Participant consent sought F.3.3 Parental consent sought F.3.4 Other consent sought F.3.5 Consent not sought F.3.6 implicit F.3.7 Not stated/unclear (please specify)
F.4 Which variable does the study aim to examine?	F.4.1 Explicitly stated (please specify) F.4.2 Implicit (please specify) F.4.3 Not stated/ unclear
F.5 What data was collected?	F.5.1 details
F.6 How was the data collected?	F.6.1 test F.6.2 Curriculum-based assessment F.6.3 Focus Group interview F.6.4 One-to-one interview (face to face or by phone) F.6.4 Observation F.6.5 Self-completion questionnaire F.6.6 self-completion report or

	diary F.6.7 Examinations F.6.9 Practical test F.6.10 Other documentation F.6.11 Not stated/ unclear (please specify) F.6.12 Please specify any other important features of data collection
F.7 Are assessments of reliability/replicability reported?	F.7.1 yes (please specify) F.7.2 no (please specify)
F.8 Are assessments of validity reported?	F.8.1 yes (please specify) F.8.2 no (please specify)
F.9 Where were the data collected?	F.9 Details
F.10 What rationale is given for data analysis methods?	F.10.1 details
F.11 What data analysis methods were used?	F.11.1 Explicitly stated (please specify) F.11.2 Implicit (please specify) F.11.3 Not stated/unclear (please specify) F.11.4 Please specify any important analytic or statistical issues
F.12 What statistical methods (if any) were used?	F.12.1 details
F.13 For qualitative methods, has complexity and diversity been addressed?	F.13.1 yes (please specify) F.4.2 no
Results and conclusions	**G**
G.1 What results do the authors report?	G.1.1 details
G.2 What population do the authors aim to extrapolate to?	G.2.1 not specified G.2.2 details
G.3 What do the authors conclude about the study findings? *Provide sufficient detail here to allow for synthesis stage to be completed*	G.3.1 details
Quality	**H**

H.1 Are the aims of the study clear? Please refer to your answers for section B here	H.1.1 yes (please specify) H.1.2 no (please specify)
H.2 Is the methodology adequately reported? *Consider your answers to section E and F here*	H.2.1 yes (please specify) H.2.2 no (please specify)
H.3 Is the study replicable?	H.3.1 yes (please specify) H.3.2 no (please specify)
H.4 Are the data traceable?	H.4.1 yes (please specify) H.4.2 no (please specify)
H.5 Is there selective reporting? *Do the authors report all variables as noted in their aims?*	H.5.1 yes (please specify) H.5.2 no (please specify)
H.6 Are there ethical concerns?	H.6.1 yes (please specify) H.6.2 no (please specify)
H.7 Is the methodology justified by the authors?	H.7.1 yes (please specify) H.7.2 no (please specify)
H.8 Is the validity of the research tools established?	H.8.1 yes (please specify) H.8.2 no (please specify)
H.9 To what extent can research error/bias be ruled out?	H.9.1 almost fully (please specify) H.9.2 a little (please specify) H.9.3 not at all
H.10 How justifiable are the conclusions? *Take the above into consideration and note to what extent the review agrees with the author about the conclusions.*	H.10.1 almost fully (please specify) H.10.2 a little (please specify) H.10.3: not at all
H.11 How generalizable are the findings?	H.11.1 details
Weight of evidence	**I**
I.1 Taking account of all quality assessment issues, can the study findings be trusted in answering the study question(s)?	I.1.1 high trustworthiness I.1.2 medium trustworthiness I.1.3 low trustworthiness
I.2 Appropriateness of research design and analysis for addressing the REA	I.2.1 high I.2.2 medium I.2.3 low

I.3 Relevance of particular focus of the study (including conceptual focus, context, sample and measures) for addressing the question, or sub-questions, of this specific systematic review	I.3.1 high I.3.2 medium I.3.3 low
I.4 Overall weight of evidence as relevant to the REA.	I.4.1 high I.4.2 medium I.4.3 low

Annex C: Items that met the Phase 2 inclusion criteria

Items that met the Phase 2 inclusion criteria (see **Phase 2 – Screening, selection and categorisation**), organised by competence.

- For criteria determining inclusion in or exclusion from syntheses of findings (Phase 3 weighting) see **Phase 3 – Detailed data extraction and weighing of evidence** and **Phase 4 – Clustering items and synthesising key findings**.
- To maximise usefulness of the table below, multiple entries are included for items that have relevance to more than one competence area.
- Note that the accuracy of bibliographical information for 'not sourced' items cannot be guaranteed.

Table C.1: Items that met the Phase 2 inclusion criteria

Included in synthesis of findings?	Item (for full bibliographical details of 'included' texts see References section)
Vocabulary	
Included	Agustín-Llach & Canga Alonso (2016) Alcón (2007) Camo & Ballester (2015) Chen, Tseng & Hsiao (2018) de la Fuente (2006) Dolean (2014) Dolean & Dolghi (2016) Gierlinger & Wagner (2016) Graham, Courtney, Marinis & Tonkyn (2017) Hennebry, Rogers, Macaro & Murphy (2017) Huang, Willson & Eslami (2012) Jiménez-Catalán & Ruiz de Zarobe (2009) Laufer (2006) Laufer & Girsai (2008) Lee & Macaro (2013) Lesniewska & Pichette (2016) Lin (2014) Luan & Sappathy (2011) Mavilidi, Okely, Chandler, Cliff & Paas (2015) Merikivi & Pietilä (2014) Porter (2016)

	Shintani (2011) Shintani (2012) Shintani (2013) Sylven (2010) Tragant, Marsol, Serrano & Llanes (2016) Williams & Thomas (2017)
Excluded	Aghlara, L., & Tamjid, N. H. (2011). The effect of digital games on Iranian children's vocabulary retention in foreign language acquisition. *Procedia-Social and Behavioral Sciences, 29,* 552-560. Amin Afshar, M. & Mojavezi, A. (2017). The effect of aural and visual storytelling on vocabulary retention of Iranian EFL learners. *English Language Teaching, 10*(4), 92-99. Behlol, M., & Kaini, M. M. (2011). Comparative effectiveness of contextual and structural method of teaching vocabulary. *English language teaching*, *4*(1), 90-97. Bilen, D., & Tavil, Z. M. (2015). The effects of cooperative learning strategies on vocabulary skills of 4th grade students. *Journal of Education and Training Studies*, *3*(6), 151-165. Choi, M. L. & Ma, Q. (2015). Realising personalised vocabulary learning in the Hong Kong context via a personalised curriculum featuring "student-selected vocabulary". *Language and Education, 29*(1), 62-78. Davoudi, M., & Yousefi, D. (2016). The effect of Keyword Method on vocabulary retention of senior high school EFL learners in Iran. *Journal of Education and Practice*, *7*(11), 106-113. Ellis, R., & Heimbach, R. (1997). Bugs and birds: Children's acquisition of second language vocabulary through interaction. *System*, *25*(2), 247-259. Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning, 44*(3), 449–491. Fahrurrozi (2017). Improving students' vocabulary mastery by using Total Physical Response. *English Language Teaching, 10*(3), 118-127. Fernández Fontecha, A. (2014). Receptive vocabulary knowledge and motivation in CLIL and EFL. *Revista de Lingüística y Lenguas Aplicadas, 9*, 23-32. Franciosi, S. J., Yagi, J., Tomoshige, Y., & Ye, S. (2016). The effect of a simple simulation game on long-term vocabulary retention. *Calico Journal*, *33*(3), 355-379.

Goundareva, I. (2011). Effect of translation practice on vocabulary acquisition in L2 Spanish. *Working Papers of the Linguistics Circle of the University of Victoria*, *21*(1), 145-154.

Harwood, N. (2002). Taking a lexical approach to teaching: Principles and problems. *International Journal of Applied Linguistics*, *12*(2), 139-155.

Heras, A., & Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research*, *19*(1), 70-88.

Huang, C. S., Yang, S. J., Chiang, T. H., & Su, A. Y. (2016). Effects of situated mobile learning approach on learning motivation and performance of EFL students. *Journal of Educational Technology & Society*, *19*(1), 263-276.

Jafari, S., & Chalak, A. (2016). The role of WhatsApp in teaching vocabulary to Iranian EFL learners at junior high school. *English Language Teaching*, *9*(8), 85-92.

Jiménez-Catalán, R., Ruiz de Zarobe, Y. & Cenoz, J. (2006). Vocabulary profiles of English foreign language learners in English as a subject and as a vehicular language. *Vienna English Working Papers, 15*(3) 23-27.

Jingjit, M. (2015). The effects of multimedia learning on Thai primary pupils' achievement in size and depth of vocabulary knowledge. *Journal of Education and Practice*, *6*(33), 72-81.

Lee, K. R., & Kwon, S. (2014). Effects of vocabulary memorizing tools on L2 learners' vocabulary size. *The Journal of AsiaTEFL*, *11*(2), 125-148.

Lin, C. C. Hsiao, H. S., Tseng, S. P., & Chan, H. J. (2014). Learning English vocabulary collaboratively in a technology-supported classroom. *TOJET: The Turkish Online Journal of Educational Technology*, *13*(1), 162-173.

Madill, M. T. R. (2014). The path to English literacy: Analyzing elementary sight word procurement using computer assisted language learning (CALL) in contrast to traditional methodologies. *Teaching English with Technology*, *14*(3), 30-57.

Mazzoni, E., & Benvenuti, M. (2015). A robot-partner for preschool children learning English using socio-cognitive conflict. *Journal of Educational Technology & Society*, *18*(4), 474-485.

Miyandehi, M. S., Babaie, H., & Saharkhiz, A. (2016). On the effects of a specialized vocabulary-training application on the development of English vocabulary among Iranian high school students. *Modern Journal of Language Teaching Methods*, *6*(9), 230-239.

Morimoto, S., & Loewen, S. (2007). A comparison of the effects of image-schema-based instruction and translation-based instruction on the acquisition of L2 polysemous words.

Language Teaching Research, *11*(3), 347-372.

Nakata, T. (2008). English vocabulary learning with word lists, word cards and computers: Implications from cognitive psychology research for optimal spaced learning. *ReCALL*, *20*(1), 3-20.

Nakata, T. (2017). Does repeated practice make perfect? The effects of within-session repeated retrieval on second language vocabulary learning. *Studies in Second Language Acquisition, 39*(4), 653-679.

Nam, T. T., & Trinh, L. Q. (2012). Powerpoint as a potential tool to learners' vocabulary retention: empirical evidences from a Vietnamese secondary education setting. *i-Manager's Journal on English Language Teaching*, *2*(4), 15-22.

Nova, J. C., Chavarro, C. I. O., & Córdoba A. T. Z. (2017). Educational videos: a didactic tool for strengthening English vocabulary through the development of affective learning in kids. *Gist: Education and Learning Research Journal*, (14), 68-87.

Öz, H., & Efecioglu, E. (2015). Graphic novels: An alternative approach to teach English as a foreign language. *Journal of Language and Linguistic Studies*, *11*(1), 75-90.

Petursdottir, A. I., & Hafliđadóttir, R. D. (2009). A comparison of four strategies for teaching a small foreign language vocabulary. *Journal of Applied Behavior Analysis*, *42*(3), 685-690.

Safataj, M., & Amiryousefi, M. (2016). Effect of homonyms instruction on vocabulary development and retention of Iranian young female elementary EFL learners through call-mediated tasks. *Modern Journal of Language Teaching Methods*, *6*(9), 192-207.

Sozler, S. (2012). The effect of memory strategy training on vocabulary development of Austrian secondary school students. *Procedia-Social and Behavioral Sciences*, *46*, 1348-1352.

Sun, Y., & Dong, Q. (2004). An experiment on supporting children's English vocabulary learning in multimedia context. *Computer assisted language learning*, *17*(2), 131-147.

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture*, *8*(2), 219-235.

Tsou, W., Wang, W., & Li, H. Y. (2002). How computers facilitate English foreign language learners acquire English abstract words. *Computers & Education*, *39*(4), 415-428.

Vungthong, S., Djonov, E., & Torr, J. (2017). Images as a resource for supporting vocabulary learning: A multimodal analysis of Thai EFL tablet apps for primary school children. *TESOL Quarterly*, *51*(1), 32-58.

	Xanthou, M. (2011). The impact of CLIL on L2 vocabulary development and content knowledge. *English Teaching*, *10*(4), 116. Yamagata, S. (2018). Comparing core-image-based basic verb learning in an EFL junior high school: Learner-centered and teacher-centered approaches. *Language Teaching Research*, *22*(1), 65-93. Yousefi, A., Yekta, R. R. & Farahmandian, H. (2014). The effect of modern lyrical music on second language vocabulary acquisition. *Mediterranean Journal of Social Sciences, 5*(23), 2583-2586.
Not sourced	Esquiliche Mesa, R., Bruton, A., & Ridgway, T. (2007). The effect of task-based reading on FL vocabulary learning. *ITL-International Journal of Applied Linguistics*, 153(1), 1-24. Naeini, N. N. & Shahrokhi, M. (2016). The effect of methodology on vocabulary learning by Iranian female and male elementary students: a comparison of TPR and Direct Method. *Modern Journal of Language Teaching Methods*, 6(1), 562-584. Otieva, R. (2011). Pragmatically-oriented approach in teaching foreign languages. *International Journal of the Humanities*, 9(7), 55-60. Taşbaşi, N., Eski, H. & Eski, G. (2015). A visual content based mobile software for vocabulary learning in secondary education. *Turkish Online Journal of Educational Technology*, 61-64. Xanthou, M. (2010). Current trends in L2 vocabulary learning and instruction: Is CLIL the right approach? In A. Psaltou-Joycey and M. Mattheoudakis (Eds.), *Advances in Research on Language Acquisition and Teaching: Selected Papers* (pp.459-471). Thessaloniki: GALA.
Grammar	
Included	Campfield & Murphy (2017) Graham, Courtney, Marinis & Tonkyn (2017) Hanan (2015) Ho & Binh (2014) Ibarrola (2012) Lichtman (2013) Martínez Adrián & Gutiérrez Mangado (2009) Tammenga-Helmantel, Arends & Canrinus (2014) Tammenga-Helmantel, Bazhutkina, Steringa, Hummel & Suhre (2016)

	Tode (2007) Toth & Guijarro-Fuentes (2013)
Excluded	Abadikhah, S., & Zarrabi, F. (2011). The effect of output tasks on the acquisition of English verbal morphemes. *Theory & Practice in Language Studies, 1*(11), 1549-1560. Alcón, E., & Mayo, M. D. P. G. (2008). Incidental focus on form and learning outcomes with young foreign language classroom learners. In J. Philp, R. Oliver, & A. Mackey (Eds.), *Second Language Acquisition and the Younger Learner: Child's Play?* (pp.173-192). Amsterdam: John Benjamins Publishing. Alghamdi, R., & Gillies, R. (2013). The impact of cooperative learning in comparison to traditional learning (small groups) on EFL learners' outcomes when learning English as a foreign language. *Asian Social Science, 9*(13), 19-27. Alinte, C. (2013) Teaching grammar through music. *The Journal of Linguistic and Intercultural Education, 6*, 7-27. Alzu'bi, M. A. (2015). Effectiveness of inductive and deductive methods in teaching grammar. *Advances in Language and Literary Studies, 6*(2), 187-193. Behlol, M., & Dad, H. (2010). Effectiveness of structural method of teaching vocabulary. *English Language Teaching, 3*(3), 125-135. Dang, T. T. D., & Nguyen, H. T. (2013). Direct versus indirect explicit methods of enhancing EFL students' English grammatical competence: a concept checking-based consciousness-raising tasks model. *English Language Teaching, 6*(1), 112-121. Fatemipour, H., & Hemmati, S. (2015). Impact of consciousness-raising activities on young English language learners' grammar performance. *English Language Teaching, 8*(9), 1-10. Grim, F. (2014). Using culture beyond its borders: the use of content-enriched instruction and the effects of input enhancement on learning in high school French classes. *Innovation in Language Learning and Teaching, 8*(3), 219-238. Idek, S., Fong, L. L., & Sidhu, G. K. (2013). The use of consciousness-raising tasks in learning and teaching of subject-verb agreement. *English Language Teaching, 6*(6), 113-122. Mehrseresht, K., Gowhary, H., & Azizifar, A. (2015). A comparative study of the use of Persian vs. English in teaching English grammar to Iranian students in junior high school. *Advances in Language and Literary Studies, 6*(4), 221-225.

	NamazianDost, I., Bohloulzadeh, G., & Pazhakh, A. (2017). The effect of task-based language teaching on motivation and grammatical achievement of EFL junior high school students. *Advances in Language and Literary Studies, 8*(2), 243-259. Rikhtegar, O., & Gholami, J. (2015). The effects of pre-versus post-presentation input flooding via reading on the young Iranian EFL learners' acquisition of simple past tense. *English Language Teaching, 8*(3), 80-88. Sadeghi, K., & Dousti, M. (2012). The effect of length of exposure to CALL technology on young Iranian EFL learners' grammar gain. *English Language Teaching, 6*(2), 14-26. Shammari, Z. A., Sharoufi, H. A., & Yawkey, T. D. (2008). The effectiveness of direct instruction in teaching English in elementary public education schools in Kuwait: A research case study. *Education, 129*(1), 80-90. Shintani, N. (2012). Input-based tasks and the acquisition of vocabulary and grammar: A process-product study. *Language Teaching Research, 16*(2), 253-279. Shintani, N. (2014). The effectiveness of processing instruction and production-based instruction on L2 grammar acquisition: A meta-analysis. *Applied Linguistics, 36*(3), 306-325. Shintani, N., & Ellis, R. (2010). The incidental acquisition of English plural–s by Japanese children in comprehension-based and production-based lessons: A process-product study. *Studies in second language acquisition, 32*(4), 607-637.
Not sourced	Cumming, J. & Lyster, R. (2016). Integrating content-based instruction into high school foreign language classrooms. In Cammarata, L. (Ed.), *Content-based foreign language teaching: Curriculum and pedagogy for developing advanced thinking and literacy skills* (pp. 77-97). New York: Routledge. Moranski, K., & Toth, P. (2016). Small-group meta-analytic talk and Spanish L2 development. In M. Sato & S. Ballinger (Eds.) *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 291-316). Amsterdam: John Benjamins. Niemeier, S., & Reif, M. (2008). Making progress simpler? Applying cognitive grammar to tense-aspect teaching in the German EFL classroom. In S. de Knop & T. de Rycker (Eds.) *Cognitive Approaches to Pedagogical Grammar* (pp. 325-356). Amsterdam: John Benjamins. Pawlak, M. (2012). The effects of focus on forms and focus on form in teaching complex grammatical structures. *Indian Journal of Applied Linguistics, 38*(1), 35-56. Prorkar, R., & Gorjian, B. (2016). Using grammar techniques in teaching wh-questions to high school students. *Modern*

	Journal of Language Teaching Methods, 6(9), 70-79. Saadati, Z. (2011). The impact of fill-in-the-map and construct-a-map techniques on EFL learners' use of tenses in oral accuracy. *European Journal of Social Sciences, 22*(2), 299-308. Zhaku-Kondri, B. 2014. Using song lyrics in the classroom: assessing the utility of song lyrics for the acquisition of a foreign language. *The 1st International Conference on Research and Educatıon – Challenges Toward the Future* (ICRAE2013), Shkodra, Albania.
Reading	
Included	Chen, Tan & Lo (2016) Fonseca-Mora, Jara-Jiménez & Gómez-Domínguez (2015) Harris (2007) Lan, Sung, & Chang (2009) Macaro & Erler (2008) Macaro & Mutton (2009) Manoli, Papadopoulou & Metallidou (2016) Martínez & de Zarobe (2017) Mistar, Zuhairi & Yanti (2016) Porter (2014) Takeda (2002) Türk & Erçetin (2014)
Excluded	Chen, J. M., Chen, M. C., & Sun, Y. S. (2010). A novel approach for enhancing student reading comprehension and assisting teacher assessment of literacy. *Computers & Education, 55*(3), 1367-1382. Chiang, H. K., & Kuo, F. L. (2005). Active learning through Web and PDA. In *Information Technology: Research and Education, 2005. ITRE 2005. 3rd International Conference on IEEE* (pp. 490-494). Ghabanchi, Z., & Mirza, F. H. (2010). The effect of summarization on intermediate EFL learners' reading comprehension and their performance on display, referential and inferential questions. *Journal of College Teaching and Learning, 7*(9), 53-60. Hamidavi, N., Shekaramiz, M., & Gorjian, B. (2016). The effect of CLIL Method on teaching reading comprehension to junior high school students. *Modern Journal of Language Teaching Methods, 6*(9), 60-69.

Hickey, T. (2001). Reading in a second language: Teaching Irish reading. *Nóta ón Eagarthóir, 7*(2), 66.

Hickey, T. M., & Stenson, N. (2016). One step forward and two steps back in teaching an endangered language? *Revisiting L2 reading in Irish. Language, Culture and Curriculum, 29*(3), 302-318.

Hong, Z. W., Huang, Y. M., Hsu, M., & Shen, W. W. (2016). Authoring robot-assisted instructional materials for improving learning performance and motivation in EFL classrooms. *Educational Technology & Society, 19*(1), 337-349.

Huang, Y. N., & Hong, Z. R. (2016). The effects of a flipped English classroom intervention on students' information and communication technology and English reading comprehension. *Educational Technology Research and Development, 64*(2), 175-193.

Jing, J. (2012). Teaching English reading through MI theory in primary schools. *English Language Teaching, 6*(1), 132-140.

Krashen, S. (2009) Anything but reading. *Knowledge Quest, 37*, 18-25.

Majuddin, E. (2014). The effects of differing densities of glossing on vocabulary uptake and reading comprehension. *Advances in Language and Literary Studies, 5*(6), 132-152.

Moghaddam, N. N., & Mahmoudi, A. (2016). The effects of pre-reading activities on reading comprehension of Iranian EFL learners. *Advances in Language and Literary Studies, 7*(3), 235-242.

Ren, G. (2011). Can cloze tests really improve second language learners' reading comprehension skills? *Journal of College Teaching and Learning, 8*(4), 15-22.

Sawangsamutchai, Y., & Rattanavich, S. (2016). A comparison of seventh grade Thai students' reading comprehension and motivation to read English through applied instruction based on the genre-based approach and the teacher's manual. *English Language Teaching, 9*(4), 54-63.

Sheu, S. P. (2003). Extensive reading with EFL learners at beginning level. *TESL Reporter, 36*(2), 8-26.

Tajeddin, Z., & Tabatabaei, S. (2016). Concept mapping as a reading strategy: does it scaffold comprehension and recall? *The Reading Matrix: An International Online Journal, 16*(1), 194-208.

Tajeddin, Z., Tabatabaei, S., & Alemi, M. (2010). Scaffolding L2 reading comprehension and recall through concept maps accompanying texts. In *EDULEARN10 Proceedings* (pp. 2574-2587). IATED.

Torabian, A., & Tajadini, M. (2017). Fostering EFL learners' reading comprehension: animation film technique. *Advances in Language and Literary Studies, 8*(2), 55-63.

	Underwood, P., Myskow, G., & Hattori, T. (2012). The effect of speed reading instruction on Japanese high school students' English reading comprehension and vocabulary development. *Journal of International Education Research, 8*(1), 27-40. Zainuddin (2015) The effect of cooperative integrated reading and composition technique on students' reading descriptive text achievement. *English Language Teaching, 8*(5), 11-21.
Not sourced	Esquiliche Mesa, R., Bruton, A., & Ridgway, T. (2007). The effect of task-based reading on FL vocabulary learning. *ITL-International Journal of Applied Linguistics, 153*(1), 1-24. Zhaku-Kondri, B. 2014. Using song lyrics in the classroom: assessing the utility of song lyrics for the acquisition of a foreign language. *The 1st International Conference on Research and Educatıon – Challenges Toward the Future* (ICRAE2013), Shkodra, Albania.
Writing	
Included	Abdelrahman, Dewitt, Alias & Rahman (2017) Bartan (2017) Bruton (2007) Buga, Capeneata, Chirasnel & Popa (2014) Fidaoui, Bahous & Bacha (2010) Gené-Gil, Juan-Garau & Salazar-Noguera (2015) Ghahremani-Ghajar & Mirhosseini (2005) Griva & Anastasiou (2009) Gündüz & Ünal (2016) Hwang, Chen, Shadiev, Huang & Chen (2014) Ikeda (2013) Lan, Sung, Cheng & Chang (2015) Maxwell-Reid (2010) Ngo & Trinh (2011) Porter (2014) Saladrigues & Llanes (2014) Sercu (2013) Shafqat, Idrees & Gujjar (2009) Taylor, Lazarus & Cole (2005) Tsiriotakis, Vassilaki, Spantidakis & Stavrou (2017) Yunus, Nordin, Salehi, Embi & Salehi (2013)

Excluded	Agustín Llach, M. P. (2017). The effects of the CLIL approach in young foreign language learners' lexical profiles. *International Journal of Bilingual Education and Bilingualism, 20*(5), 557-573. Akinwamide, T. K. (2012) The influence of Process Approach on English as second language students' performances in essay writing. *English Language Teaching, 5,* 16-29. Anastasiadou, A. (2017). The impact of the application of a process writing component on the students' writing output in English as a Foreign Language. *Selected papers on theoretical and applied linguistics, 20*(3), 11-32. Attan, A., & Khalidi, H. (2015). Exploring peer revision as a strategy in the ESL writing classroom. In *SHS Web of Conferences* (Vol. 18). EDP Sciences. Beard, A. C. (2016). *Addysg Gymraeg ail iaith mewn Ysgolion cyfrwng-Saesneg: Astudiaeth i archwilio i ba raddau y mae amodau dysgu'r rhaglen Gymraeg ail iaith yng nghyfnodau Allweddol 2 a 3 yn gymwys i gynhyrchu siaradwyr yr iaith.* (Unpublished doctoral dissertation). University of Wales Trinity Saint David, UK. Biber, D., Nekrasova, T., & Horn, B. (2011). The effectiveness of feedback for L1-English and L2 writing development: A meta-analysis. In *TOEFL iBT Research Report No. TOEFLiBT-14.* Princeton, NJ: Educational Testing Service. Blain, S. (2001) Study of verbal peer feedback on the improvement of the quality of writing and the transfer of knowledge in francophone students in grade 4 living in a minority situation in Canada. *Language, Culture and Curriculum, 14,* 156-70. Bournot-Trites, M., Belliveau, G., Spiliotopoulos, V., & Séror, J. (2007). The role of drama on cultural sensitivity, motivation and literacy in a second language context. *Journal for Learning through the Arts, 3*(1), 9. Britsch, S. (2010) Photo-booklets for English language learning: incorporating visual communication into early childhood teacher preparation. *Early Childhood Education Journal, 38*, 171-177. Buitrago Campo, A. C. (2016) Improving 10th graders' English communicative competence through the implementation of the task-based learning approach. *PROFILE: Issues in Teachers' Professional Development, 18*, 95-110. Chen, T. (2016) Technology-supported peer feedback in ESL/EFL writing classes: a research synthesis. *Computer Assisted Language Learning, 29*, 365-397. Damayanti, I. L. (2017) From storytelling to story writing: The implementation of reading to learn (R2L) pedagogy to teach

English as a foreign language in Indonesia. *Indonesian Journal of Applied Linguistics, 6,* 232-245.

East, M. (2006) The impact of bilingual dictionaries on lexical sophistication and lexical accuracy in tests of L2 writing proficiency: A quantitative analysis. *Assessing Writing, 11,* 179-197.

Gutiérrez, K. G. C., Puello, M. N., & Galvis, L. A. P. (2015). Using pictures series technique to enhance narrative writing among ninth grade students at Institución Educativa Simón Araujo. *English Language Teaching, 8*(5), 45.

Hwang, W. Y., Shadiev, R., & Huang, S. M. (2011). A study of a multimedia web annotation system and its effect on the EFL writing and speaking performance of junior high school students. *ReCALL, 23*(2), 160-180.

Juriah, J. (2015). Implementing Controlled Composition to Improve Vocabulary Mastery of EFL Students. *Dinamika Ilmu, 15*(1), 137-162.

Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal, 99*(1), 1-18.

Le Bouthillier, J. (2015). Writing progress and processes of grade 7 at-risk French immersion students. *International Journal of Technology and Inclusive Education, 2*(1), 690-697.

Lee, I., & Wong, K. (2014). Bringing innovation to EFL writing: The case of a primary school in Hong Kong. *The Asia-Pacific Education Researcher, 23*(1), 159-163.

Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *Journal of Second Language Writing, 30,* 66-81.

Mak, B., Coniam, D., & Kwan, M. C. S. (2007). A buddy reading programme in Hong Kong schools. *ELT journal, 62*(4), 385-394.

Mehrpour, S., & Ghayour, M. (2017). The Effect of Educational Computerized Games on Learning English Spelling among Iranian Children. *The Reading Matrix: An International Online Journal, 17*(2),165-178.

Met, M. (2013). Toward a Renewed Focus. Literacy in Early Language Programs. *Learning Languages, 18*(1), 40-41.

Mohamed, A. R., Tumin, M., & Omar, H. (2008). Success structure for accelerated acquisition of English by young ESL learners. *Educational Research and Reviews, 3*(4), 169-181.

Salimi, A., Amini, D., & Yousefzadeh, M. (2015). Combining Explicit-Implicit with Focus on Form-Focus on FormS Feedback in EFL Learners' Written Task Accuracy. *International Journal of Applied Linguistics and English Literature, 4*(6), 113-118.

Shadiev, R., Hwang, W. Y., Yueh-Min, H., & Liu, T. Y. (2015).

	The impact of supported and annotated mobile learning on achievement and cognitive load. *Journal of Educational Technology & Society, 18*(4), 53-69. Sylvén, L. K., & Ohlander, S. (2015). The CLISS project: Receptive vocabulary in CLIL versus non-CLIL groups. *Moderna språk, 108*(2), 80-114. Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing, 16*(4), 255-272.
Not sourced	Van Gelderen, A. & R. Oostdam (2002) Improving linguistic fluency for writing: effects of explicitness and focus of instruction. L1 - *Educational Studies in Language and Literature, 2,* 239-270. Zhu, C. Y., Cheng, I. L., Chew, S. W., Wu, G. Y., Hsu, C. N., & Chen, N. S. (2017). The effect of children learning English vocabulary through a gesture-based system. In E. Popescu, M. K. Khribi, R. Huang, M. Jemni, N. S. Chen & D. G. Sampson (Eds.) *Innovations in Smart Learning* (pp. 209-214). Springer, Singapore.
Speaking	
Included	Arslanyilmaz (2013) Gallardo del Puerto & Gómez Lacabex (2013) Gallardo del Puerto, Gómez Lacabex & García Lecumberri (2009) García Mayo & Imaz Agirre (2016) Griva & Semoglou (2012) Ho & Binh (2014) Jung, Kim & Murphy (2017) Neri, Mich, Gerosa & Giuliani (2008) Nutta, Feyten, Norwood, Meros, Yoshii & Ducher (2002) Pérez Cañado & Lancaster (2017) Satar & Ozdener (2008) Verdugo & Belmonte (2007) Wang, Young & Jang (2013) Wu, Wang & Chen (2015) Young & Wang (2014)
Excluded	Alvarez-Marinelli, H., Blanco, M., Lara-Alecio, R., Irby, B. J.,

Tong, F., Stanley, K., & Fan, Y. (2016). Computer assisted English language learning in Costa Rican elementary schools: an experimental study. *Computer Assisted Language Learning, 29*(1), 103-126.

Buitrago Campo, A. C. (2016) Improving 10th Graders' English Communicative Competence through the Implementation of the Task-Based Learning Approach. *PROFILE: Issues in Teachers' Professional Development, 18*, 95-110.

Castaneda, M. E. (2013) "I am proud that I did it and it's a piece of me": Digital storytelling in the foreign language classroom. *CALICO Journal*, 30, 44-62.

Dewi, R. S., Kultsum, U., & Armadi, A. (2017). Using communicative games in improving students' speaking skills. *English Language Teaching, 10*(1), 63-71.

Duarte Romero, M., Tinjacá Bernal, L. M., & Carrero Olivares, M. (2012). Using songs to encourage sixth graders to develop English speaking skills. *Profile Issues in Teachers Professional Development, 14*(1), 11-28.

Forbes, K., & Fisher, L. (2018). The impact of expanding advanced level secondary school students' awareness and use of metacognitive learning strategies on confidence and proficiency in foreign language speaking skills. *The Language Learning Journal, 46*(2), 173-185.

Fotovatnia, Z., & Omidi, M. (2013). The effect of exposure to the visual medium on learning pronunciation and word stress of L2 learners. *Theory and Practice in Language Studies, 3*(5), 769.

Galimova, E. G., & Shvetsova, M. G. (2016). Modern educational technologies in teaching senior secondary pupils' communication in the form of a monologue and in generating their cognitive interest. *International Journal of Environmental and Science Education, 11*(8), 1951-1962.

García Mayo, M. D. P. & Lázaro Ibarrola, A. (2015). Do children negotiate for meaning in task-based interaction? Evidence from CLIL and EFL settings. *System, 54*, 40-54.

Georgopoulou, A. A., & Griva, E. (2012). Story-based learning: implementation of a pilot project for early foreign language learning. *Procedia-Social and Behavioral Sciences, 31*, 110-115.

Gil, I. R., de Larios, J. R., & Balibrea, Y. C. (2017). The effect of repeated reading on the pronunciation of young EFL learners. *Porta Linguarum: revista internacional de didáctica de las lenguas extranjeras, 27,* 7-19.

González-Bueno, M. & Quintana-Lara, M. (2011). The Teaching of L2 Pronunciation through Processing Instruction. *Applied Language Learning, 21*, 57-82.

Griva, E., Semoglou, K., & Geladari, A. (2010). Early foreign language learning: Implementation of a project in a game–

based context. *Procedia-Social and Behavioral Sciences, 2*(2), 3700-3705.

Hwang, W. Y., Shadiev, R., Hsu, J. L., Huang, Y. M., Hsu, G. L., & Lin, Y. C. (2016). Effects of storytelling to facilitate EFL speaking using Web-based multimedia system. *Computer Assisted Language Learning, 29*(2), 215-241.

Hwang, W. Y, Shih, T. K., Yeh, S. C., Chou, K. C., Ma, Z. H., & Sommool, W. (2014). Recognition-based physical response to facilitate EFL learning. *Educational Technology & Society, 17*(4), 432–445.

Hwang, W. Y., Shadiev, R., & Huang, S. M. (2011). A study of a multimedia web annotation system and its effect on the EFL writing and speaking performance of junior high school students. *ReCALL, 23*(2), 160-180.

Iamsaard, P., & Kerdpol, S. (2015). A study of effect of dramatic activities on improving English communicative speaking skill of grade 11th students. *English Language Teaching, 8*(11), 69-78.

Iswardati, I. (2016). The implementation of group investigation to improve the students' speaking skill. *Dinamika Ilmu, 16*(2), 245-261.

Juriah, J. (2015). Implementing controlled composition to improve vocabulary mastery of EFL students. *Dinamika Ilmu, 15*(1), 137-162.

Lázaro, A., & García Mayo, M. d. P. (2012). L1 use and morphosyntactic development in the oral production of EFL learners in a CLIL context. *International Review of Applied Linguistics in Language Teaching (IRAL), 50*(2), 135-160.

Lai, Y. S., Hung-Hsu, T., & Yu, P. T. (2009). A multimedia English learning system using HMMs to improve phonemic awareness for English learning. *Journal of Educational Technology & Society, 12*(3), 266-281.

Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in second language acquisition, 32*(2), 265-302.

Ma'rof, A. M., Redzuan, M. R., Anderson, R. C., & Ma'rof, A. A. (2012). Improving young indigenous Malaysian children's incidental vocabulary acquisition and oral narrative skills through shared-book reading. *Life Science Journal, 9*(4), 841-848.

Martinsen, R., Montgomery, C., & Willardson, V. (2017). The effectiveness of video-based shadowing and tracking pronunciation exercises for foreign language learners. *Foreign Language Annals, 50*(4), 661-680. doi:10.1111/flan.12306

Marzuki, M., Prayogo, J. A., & Wahyudi, A. (2016). Improving the EFL learners' speaking ability through interactive storytelling. *Dinamika Ilmu, 16*(1), 15-34.

Miller, P. C. & Pan, W. (2012) Recasts in the L2 classroom: A meta-analytic review. *International Journal of Educational Research, 56,* 48-59.

Pinter, A. (2007). Some benefits of peer–peer interaction: 10-year-old children practising with a communication task. *Language Teaching Research, 11*(2), 189-207.

Rew, S., & Moon, Y. I. (2013). The effects of using English drama on the learning of target expressions for primary school students. *The Journal of Asia TEFL, 10*(4), 215-239.

Ruiz, R. S., & Pardo, S. L. (2016). Improving English oral skills through pair work and peer-peer dialogue. *Revista de Lenguas para Fines Específicos, 22*(1), 170-189.

Saito, K. (2012) Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *Tesol Quarterly, 46*(4), 842-854.

Saz, O., Rodríguez, V., Lleida, E., Rodriguez, W. R., & Vaquero, C. (2010). The use of multimodal tools for pronunciation training in second language learning of preadolescents. *Language Teaching: Techniques, Developments and Effectiveness.* Hauppauge (NY): Nova Science Publishers.

Shadiev, R., Hwang, W. Y., Yueh-Min, H., & Liu, T. Y. (2015). The impact of supported and annotated mobile learning on achievement and cognitive load. *Journal of Educational Technology & Society, 18*(4), 53-69.

Sharifi, A., Jahedizadeh, S., & Ghanizadeh, A. (2017). The effect of simulation on middle school students' perceptions of classroom activities and their foreign language achievement: a mixed-methods approach. *International Electronic Journal of Elementary Education, 9*(3), 667-680.

Tabatabaei, O., & Gahroei, F. R. (2011). The contribution of movie clips to idiom learning improvement of Iranian EFL learners. *Theory & Practice in Language Studies, 1*(8), 990-1000.

Tehrani, A. R., Barati, H., & Youhanaee, M. (2013). The effect of methodology on learning vocabulary and communication skills in Iranian young learners: a comparison between audiolingual method and natural approach. *Theory and Practice in Language Studies, 3*(6), 968-976.

Urmeneta, C. E., & Sola, A. S. (2009). Language learning through tasks in a content and language integrated learning (CLIL): Science classroom. *Porta Linguarum: revista internacional de didáctica de las lenguas extranjeras, 11*, 65-83.

Urrutia León, W., & Vega Cely, E. (2010). Encouraging teenagers to improve speaking skills through games in a Colombian public school. *Profile Issues in Teachers Professional Development, 12*(1), 11-31.

	Vaca Torres, A. M., & Gómez Rodríguez, L. F. (2017). Increasing EFL learners oral production at a public school through project-based learning. *Profile Issues in Teachers Professional Development, 19*(2), 57-71. van de Guchte, M., M. Braaksma, G. Rijlaarsdam & P. Bimmel (2016) Focus on Form through task repetition in TBLT. *Language Teaching Research, 20*, 300-320. Yang, P.-l., F. Tong, B. J. Irby, R. Lara-Alecio, N. Ramos & M. Nava-Walichowski (2016) The effect of storytelling and retelling and higher order thinking on oral performance of elementary students in English as Foreign Language (EFL) program: A pilot study in Mainland China and Taiwan. *Asian Journal of Education and Training, 2*, 23-33. Zhao, J. 2014. Project-based instruction in teaching Chinese as a foreign language. In *Contemporary Approaches to Activity Theory: Interdisciplinary Perspectives on Human Behavior,* 108-127. IGI Global.
Not sourced	Cheng, Y.-L. (2012). Teaching phonics to Chinese L1 EFL pupils: Pathway to the future. *Computer-Assisted Language Learning and Teaching, 2,* 76-94. Črčinović, J. & Rozman, M. B. (2004). Teaching English language with songs. *Informatologia, 37*, 217-220. Planas, S. R. (2011). Kidscast: Podcasts and videoconferences for a 'global' ESL student exchange. In *INTED2011 Proceedings* (pp. 1543-1546). IATED. Saadati, Z. (2011). The impact of fill-in-the-map and construct-a-map techniques on EFL learners' use of tenses in oral accuracy. *European Journal of Social Sciences, 22,* 299-308. Ulusoy, G. & Demirbilek, M. (2013). The effects of using situation comedy video on English speaking. *Anthropologist, 16,* 351-361.
Listening	
Included	Cabrera & Martínez (2001) Griva & Semoglou (2012) Harris (2007) Li, Wu & Lin (2017) Pérez Cañado & Lancaster (2017) Rouhi, Nabavi & Mohebbi (2014) Tišma (2016) Verdugo & Belmonte (2007)

	Vulchanova, Aurstad, Kvitnes & Eshuis (2015)
Excluded	Alvarez-Marinelli, H., Blanco, M., Lara-Alecio, R., Irby, B. J., Tong, F., Stanley, K., & Fan, Y. (2016). Computer assisted English language learning in Costa Rican elementary schools: an experimental study. *Computer Assisted Language Learning, 29*(1), 103-126. Chou, M. H. (2013). A content-based approach to teaching and testing listening skills to grade 5 EFL learners. *International Journal of Listening, 27*(3), 172-185. González, N. I. P. (2010). Teaching English through stories: A meaningful and fun way for children to learn the language. *PROFILE, 12,* 95-106. Hong, Z. W., Huang, Y. M., Hsu, M., & Shen, W. W. (2016). Authoring robot-assisted instructional materials for improving learning performance and motivation in EFL classrooms. *Educational Technology & Society, 1*9(1), 337-349. Hwang, W.Y, Shih, T. K., Yeh, S.C., Chou, K.C., Ma, Z.H., & Sommool, W. (2014). Recognition-based physical response to facilitate EFL learning. *Educational Technology & Society, 17* (4), 432–445. Samawiyah, Z., & Saifuddin, M. (2016). Phonetic symbols through Audiolingual Method to improve the students' listening skill. *Dinamika Ilmu, 16*(1), 35-46.
Not sourced	Chateaureynaud, M. A. (2014). New technologies and minority languages learning: How to use interactive technological tools to develop Occitan language. In L. G. Chova, A. L. Martínez & I. C. Torres (Eds.), *Inted2014: 8th International Technology, Education and Development Conference*, 7624-7629. Shintani, N. & Wallace, P. M. (2014). Effects of listening support in second language classroom: A meta-analysis. *English Teaching and Learning 38*(3) 71-101. Uchiyama, T. (2011). Reading versus telling of stories in the development of English vocabulary and comprehension in young second language learners. *Reading Improvement, 48*(4), 168-179. Zhaku-Kondri, B. 2014. Using song lyrics in the classroom: assessing the utility of song lyrics for the acquisition of a foreign language. *The 1st International Conference on Research and Educatıon – Challenges Toward the Future* (ICRAE2013), Shkodra, Albania.
General language	

Included	Artieda, Roquet & Nicolás-Conesa (2017) Bauckham (2016) Boulton & Cobb (2017) Cable, Driscoll, Mitchell, Sing, Cremin, Earl & Heins (2010) Campfield & Murphy (2017) Collins, Halter, Lightbown & Spada (1999) Collins & White (2011) Coral, Lleixà & Ventura (2016) Dalton-Puffer (2011) Davis (2017) de Diezmas (2016) Driscoll, Jones, Martin, Graham-Matheson, Dismore & Sykes (2004) Edelenbos, Johnstone & Kubanek (2007) Elgün-Gündüz, Akcan & Bayyurt (2012) Grgurović, Chapelle & Shelley (2013) Harris & Ó Duibhir (2011) Lo Bianco (2009) Mitchell & Myles (Forthcoming 2018) Ofsted (2011) Spada & Lightbown (2008)
Excluded	Al Harrasi, K. T. S. (2014). Using "Total physical response" with young learners in Oman. *Childhood Education, 90*(1), 36-42. Al-Elaimat, A. R. (2013). The effect of using computer games on lower basic stage students' achievement in English at Al–Salt schools. *International Education Studies, 6*(2), 160-171. Almekhlafi, A. G. (2006). The effect of computer assisted language learning (CALL) on United Arab Emirates English as a foreign language (EFL) school students' achievement and attitude. *Journal of Interactive Learning Research, 17*(2), 121-142. Alpar, M. (2013). The importance of games in teaching foreign languages to children. *Educational Research and Reviews, 8*(15), 1247-1255. Bataineh, R. F., & Hani, N. A. B. (2011). The effect of a CALL program on Jordanian sixth-grade students' achievement. *Teaching English with Technology, 11*(3), 3-24. Baş, G., & Beyhab, Ö. (2017). Effects of multiple intelligences supported project-based learning on students' achievement levels and attitudes towards English lesson. *International*

Electronic Journal of Elementary Education, 2(3), 365-386.

Breidbach, S., & Viebrock, B. (2012). CLIL in Germany: Results from recent research in a contested field of education. *International CLIL Research Journal, 1*(4), 5-16.

Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System, 39*(4), 523-532.

Chang, C. W., & Chen, G. D. (2010). Using a humanoid robot to develop a dialogue-based interactive learning environment for elementary foreign language classrooms. Journal of Interactive Learning Research, 21(2), 215-235.

Chiu, Y. H. (2013). Computer- assisted second language vocabulary instruction: A meta- analysis. British Journal of Educational Technology, 44(2), E52-E56.

Chiu, Y. H., Kao, C. W., & Reynolds, B. L. (2012). The relative effectiveness of digital game- based learning types in English as a foreign language setting: A meta- analysis. British Journal of Educational Technology, 43(4), E104-E107.

Cloud, N., Lakin, J., & Leininger, E. (2011). Learner-centered teaching: the core of effective practices for adolescent English language learners. *TESOL Journal, 2*(2), 132-155.

Connolly, T. M., Stansfield, M., Josephson, J., Lázaro, N., Rubio, G., Ortiz, C. R., ... & Tsvetanova, S. (2008, March). Using alternate reality games to support language learning. In *Proceedings of the Seventh IASTED International Conference on Web-based Education* (pp. 76-81). ACTA Press.

Dalton-Puffer, C., Nikula, T. & Smit, U. (2010). Language use and language learning in CLIL: Current findings and contentious issues. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.) *Language use and language learning in CLIL classrooms* (pp. 279-292). Amsterdam: John Benjamins Publishing.

de Graaff, R., Koopman, G. J., & Westhoff, G. (2007). Identifying effective L2 pedagogy in content and language integrated learning (CLIL). *Vienna English Working Papers, 16*(3), 12-19.

Dev, M. (2016). Constructivist approach enhances the learning: a search of reality. *Journal of Education and Practice, 7*(25), 59-62.

Dourda, K., Bratitsis, T., Griva, E., & Papadopoulou, P. (2013). Combining game based learning with content and language integrated learning approaches: A research proposal utilizing QR codes and Google Earth in a geography-based game. In *Proceedings of the 6th European Conference on Games Based Learning*, (pp. 130-139).

Drew, I. (2009). Using the early years literacy programme in primary EFL Norwegian classrooms. In M. Niklov (Ed.) *Early learning of modern foreign languages: processes and outcomes,* (pp. 108-120). Bristol: Multilingual Matters.

Eren, Z. (2004). The use of short-story in teaching English to the students of public high schools. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 26,* 41-47.

Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning. In P. Rebuschat (Ed.) *Implicit and explicit learning of languages*, (pp. 443-482). Amsterdam: John Benjamins.

Goorhuis-Brouwer, S., & de Bot, K. (2010). Impact of early English language teaching on L1 and L2 development in children in Dutch schools. *International Journal of Bilingualism, 14*(3), 289-302.

Haggan, M. (2010). Text messaging, pragmatic competence, and affective facilitation in the EFL context: A pilot study. *Canadian Journal of Applied Linguistics/Revue canadienne de linguistique appliquée, 13*(2), 150-164.

Hamamorad, A. (2016). Integrating CALL into language teaching: implementing WBLL technique to teach English language to EFL learners in a secondary school in Kurdistan region. *Journal of Education and Practice, 7*(1), 38-47.

Hussain, I., Hamdani, S. N. H., Quraishi, U., & Zeeshan, M. (2010). Effect of direct teaching method on the academic achievement of high and low achievers in the subject of English at the secondary level. *Journal of College Teaching and Learning, 7*(7), 45-50.

Jarvis, S. (2013). How effective is it to teach a foreign language in the Foundation Stage through songs and rhymes? *Education 3-13, 41*(1), 47-54.

Korosidou, E., & Griva, E. (2016). " It's the same world through different eyes": a content and language integrated learning project for young EFL learners. *Colombian Applied Linguistics Journal, 18*(1), 116-132.

Kristiawan, M. (2013). The Implementation of cooperative learning in English class of favorite school of secondary high school 5 Batusangkar, West Sumatera. *International Journal of Educational Administration and Policy Studies, 5*(6), 85-90.

Kwok, V. H. (2015). Robot vs. human teacher: Instruction in the digital age for ESL learners. *English Language Teaching, 8*(7), 157-163.

Lantz-Andersson, A. (2018). Language Play in a Second Language: Social Media as Contexts for Emerging Sociopragmatic Competence. *Education and Information Technologies, 23*(2), 705-724.

Lasagabaster, D. (2000). The effects of three bilingual education models on linguistic creativity. *IRAL-International Review of Applied Linguistics in Language Teaching, 38*(3-4), 213-228.

Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied*

Linguistics Journal, 1(1), 30-41.

Legg, R. (2009). Using music to accelerate language learning: An experimental study. *Research in Education, 82*(1), 1-12.

Lightbown, P. M. (2014). Making the minutes count in L2 teaching. *Language Awareness, 23*(1-2), 3-23.

Lin, H. (2015) A meta-synthesis of empirical research on the effectiveness of computer-mediated communication (CMC) in SLA. *Language Learning & Technology, 19*(2), 85-117.

Llinares, A., & Pastrana, A. (2013). CLIL students' communicative functions across activities and educational levels. *Journal of pragmatics, 59,* 81-92.

Mei-Ju, C., & Ching-Chi, C. (2014). Let's scaffold elementary G1 students playing with English. *Procedia-Social and Behavioral Sciences, 143,* 542-549.

Merisuo-Storm, T. (2009). Content and language integrated learning (CLIL): Providing children with good first and foreign language skills. *Journal of Education Research, 3*(1-2), 69-85.

Mhathúna, M. M. (2008). Supporting children's participation in second-language stories in an Irish-language preschool. *Early Years, 28*(3), 299-309.

Mukundan, J., Mahvelati, E. H., & Nimehchisalem, V. (2012). The effect of an intensive English program on Malaysian secondary school students' language proficiency. *English Language Teaching, 5*(11), 1-7.

Myers, J. L., & Chang, S. F. (2009). A multiple-strategy-based approach to word and collocation acquisition. *IRAL-International Review of Applied Linguistics in Language Teaching, 47*(2), 179-207.

Oghyanous, P. A. (2017). The effect of brain-based teaching on young EFL learners' self-efficacy. *English Language Teaching, 10*(5), 158-166.

Pérez Cañado, M. L. (2006). The effects of explicit spelling instruction in the Spanish EFL classroom: Diagnosis, development and durability. *Language Awareness, 15*(1), 20-37.

Philominraj, A., Jeyabalan, D., & Vidal-Silva, C. (2017). Visual learning: a learner centered approach to enhance English language teaching. *English Language Teaching, 10*(3), 54-62.

Pisarenko, V. (2017). Teaching a foreign language using videos. *Social Sciences, 6*(4), 125.

Plonsky, L. (2011). The effectiveness of second language strategy instruction: a meta-analysis. *Language learning, 61*(4), 993-1038.

Pritchard, A., Hunt, M., & Barnes, A. (2010). Case study investigation of a videoconferencing experiment in primary schools, teaching modern foreign languages. *Language*

Learning Journal, 38(2), 209-220.

Shahamat, A., & Mede, E. (2016). Integration of collaborative learning in Grade K-5 EFL classrooms. *Education 3-13, 44*(6), 682-697.

Sundqvist, P. (2016) Gaming and young language learners. In F. Farr & L. Murray (Eds.), *The Routledge Handbook of Language Learning and Technology* (pp. 446-458). London: Routledge.

Tinsley, T., & Board, K. (2016). *Language trends 2015/16: The state of language learning in primary and secondary schools in England.* Reading: Education Development Trust. Retrieved from https://www.britishcouncil.org/sites/default/files/language_trends_sur vey_2016_0.pdf

Trypa, A., & Mastrogiannis, A. (2010). Involving drama in the foreign language learning process: teaching English to Greek primary school students with the aid of drama activities. In *ICERI2010 Proceedings* (pp. 5742-5748). IATED.

Tsou, W. (2005). The effects of cultural instruction on foreign language learning. *RELC journal, 36*(1), 39-57.

Tsou, W., Wang, W., & Tzeng, Y. (2003, July). Applying computer multimedia storytelling Website in foreign language learning. In *Advanced Learning Technologies, 2003. Proceedings. The 3rd IEEE International Conference* (pp. 262-263). IEEE.

Wannagat, U. (2007). Learning through L2–Content and language integrated learning (CLIL) and English as medium of instruction (EMI). *International Journal of Bilingual Education and Bilingualism, 10*(5), 663-682.

West, L., & Verspoor, M. (2016). An impression of foreign language teaching approaches in the Netherlands. *Levende Talen Tijdschrift, 17*(4), 26-36.

Yang, J. C., Chen, C. H., & Jeng, M. C. (2010). Integrating video-capture virtual reality technology into a physically interactive learning environment for English learning. *Computers & Education, 55*(3), 1346-1356.

Yang, J. C., Quadir, B., & Chen, N. S. (2016). Effects of the badge mechanism on self-efficacy and learning performance in a game-based english learning environment. *Journal of Educational Computing Research, 54*(3), 371-394.

Yunus, M. M., Nordin, N., Salehi, H., Sun, C. H., & Embi, M. A. (2013). Pros and cons of using ICT in teaching ESL reading and writing. *International education studies, 6*(7), 119-130.

Ziegler, N. (2016). Synchronous computer-mediated communication and interaction: A meta-analysis. *Studies in Second Language Acquisition, 38*(3), 553-586.

Not sourced	Blondin, C., Candelier, M., Edelenbos, P., Johnstone, R., Kubaneck, A., & Taeschner, T. (1998). *Foreign languages in primary and pre-school education: A review of recent research within the European Union.* London: CILT. Dalton-Puffer, C. (2008). Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe. In W. Delanoy & L. Volkmann (Eds.), *Future perspectives for English language teaching* (pp. 139-158) Heidelberg: Carl Winter. Grigoropoulos, J. E. & Arnold, H. (2016). The i2Flex instructional methodology implemented in middle school classes for young EFL and foreign language learners. In M. D. Avgerinou & S. P. Gialamas (Eds.) *Revolutionizing K-12 Blended Learning through the i2Flex Classroom Model* (pp. 208-242). IGI Global. Lewis, W. G. (2010). *Raising standards in Welsh second language at ages 5-16*. National Assembly for Wales. Lewis, W. G. (2010). *Welsh second language in the National Curriculum: The case of teaching Welsh in English-medium schools.* National Assembly for Wales. Li, G., & Ni, X. (2013). Effects of a technology-enriched, task-based language teaching curriculum on Chinese elementary students' achievement in English as a foreign language. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT), 3*(1), 33-49. Lightbown, P. M. (2014). CBLT: What we know now. In P. M. Lightbown (Ed.) *Focus on Content-Based Language Teaching - Oxford Key Concepts for the Language Classroom.* Oxford: Oxford University Press. Low, L., Brown, S., Johnstone, R., & Pirrie, A. (1995). *Foreign languages in Scottish primary schools–Evaluation of the Scottish pilot projects.* Final Report to Scottish Office. Mármol, G. A. (2012). Systematic introduction of vocabulary and its effect on acquisition: Primary Education students of English as a Foreign Language. *Porta Linguarum: revista internacional de didáctica de las lenguas extranjeras, 18,* 149-160. Medina, L. M. 2003. *Bilingual education*. San Diego, CA: Greenhaven Press. Pishghadam, R., Khodadady, E., & Rad, N. D. (2011). The effect of form versus meaning-focused tasks on the development of collocations among Iranian intermediate EFL learners. *English Language Teaching, 4,* 180-189.