

Core Corpus Skills for designing In-Sessional ESAP materials

Karin Whiteside, University of Reading
BAAL Corpus Linguistics SIG, 14th January 2021

My background

- **EAP-related corpus research**
 - Doctoral project – phraseologies around grammatical keywords in undergraduate History & PIR Essay texts (incl. sub-corpora from BAWE)
(Wordsmith & Wmatrix)
- **EAP practitioner use of corpora/corpus tools**
 - Edit your Thesis with Corpora (Charles, 2018*)
 - **Materials development for ESAP insessional**
(AntConc, AntCorGen)

*Charles, M. 2018. Corpus-assisted editing for doctoral students: More than just concordancing. *Journal of English for Academic Purposes*, 36, 15-25.

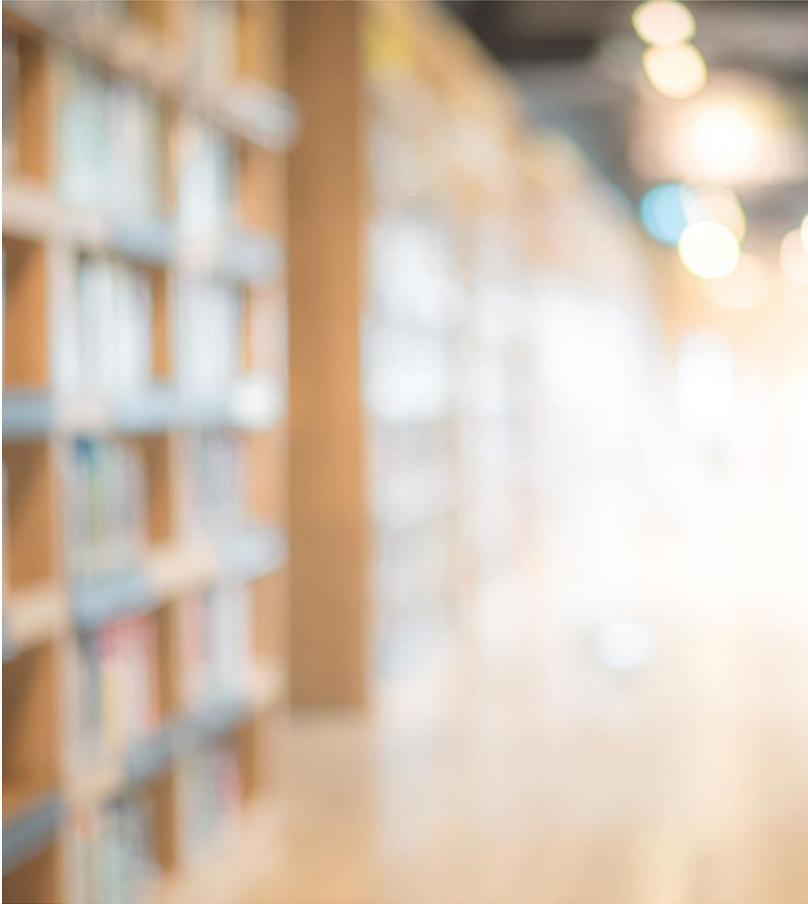
Some generalisations ...

ESAP corpus research

- principled corpus design
- careful/often v. time-consuming
- v. good at producing valuable ‘big picture’ insights ...
- *but* often less good at generating teachable bits of language in any kind of useful volume

Corpora/corpus tools in ESAP practice

- *as principled as possible* corpus design, but judicious compromises usually needed
- v. tight *time constraints* – shortcuts often needed
- need to generate as many useful, teachable bits of language as possible ...
- *but* this is time-consuming!



Teaching context

- Reading's **In-session ESAP** provision – 30+ different disciplinary cohorts
 - **Part 1 Psychology**
 - 2015/16 to 2019/20 **PY1SKE** (cf. PY1SK) **Skills for Psychology with Academic Language**
 - 2020/21 onwards
 - Non-credit-bearing with a smaller audience
 - **T&L funding to create an online resource for all students**

Writing for Psychology

Key Assignment Types: Content, Structure and Language

Introduction

A note on the approach used in this resource

Introduction

This set of online learning resources is designed to help you better understand how to write the key assignment types in Psychology - [Reports](#), [Essays](#), [Critical Reviews](#) and [Reflective Accounts](#).

[Click on the tabs at the top of the page to learn more about each assignment type.](#) There are three tabs for Reports, and one tab each for the other Psychology genres.

Created using Xerte open-source software
(<https://xerte.org.uk/index.php/en/>)

Reports: Introductions & Methods

Writing for Psychology

Introductions
Task 1 - Purposes of an Introduction
Task 2 - Stages of an Introduction
Content & Structure
Introductions - Key Language
Methods
Language Task 1
Methods - Examples
Language Task 2
Methods - Key Language

Introductions

Genre-focused tasks

The Introduction section is crucial for readers' understanding of the context and purpose of your study.

Task 1 and **Task 2** are designed to get you thinking about the content and structure of the Introduction section in a Psychology Report. Do these tasks then look at the [Content & Structure](#) and [Key Language](#) summaries below.

Language-focused tasks

Task 1 - Purposes of an Introduction

This Task focuses on the various purposes of an Introduction.

Reports: Introductions - Task 1

Below are **short extracts from the Introduction** of a Psychology Report on the subject of short-term memory entitled **The Effect of Semantically Related and Unrelated Background Speech on Immediate Recall of Printed Word Lists***.

Match each extract with the main purpose it is fulfilling within the Introduction, by **dragging** purposes from the top right corner.

Example texts used with permission from BAWE

Hypotheses

Key Language sheets created using corpora and corpus tools

Introductions - Key Language

- Introductions
- Task 1 - Purposes of an Introduction
- Task 2 - Stages of an Introduction
- Content & Structure
- Introductions - Key Language
- Methods
- Language Task 1
- Methods - Examples
- Language Task 2
- Methods - Key Language

The document below gives you **key vocabulary and phrases** for the different parts of an Introduction.

It has been hypothesized that deaf people may explore and see the visual world differently from hearing people because of ...

Recently,	there has been	(increasing)/(an)
In recent years		increased/growing/significant) interest in ...
In the past few years		a growing recognition of the importance of ...
In the past (fifteen) years		mounting evidence suggesting ...
In the last decade		etc.
To date		

e.g.

In recent years, there has been a growing interest in such individual variation, assessed using ...

[X]

has been found to be
has been shown to be

an important variable in ...
an important determinant of ...
a strong/reliable predictor/indicator of ...
a major factor (affecting) ...
associated with ...
(closely/positively) **related to ...**
(positively) **correlated with ...**
linked with .../(negatively) linked to ...
.../sensitive to ...

[Y]

e.g.

Recently, **competition** **has been found to be a major factor affecting** trust and trustworthiness ...

Attitude **has been shown to be a reliable predictor of** behavioural intention to make certain food choices ...

research has shown that

e.g.

However, previous **research has shown that** this effect is highly variable across studies

as comprehensive as possible, linking phraseologies & semi-fixed sequences to genre functions

Introductions

Task 1 - Purposes of an Introduction

Task 2 - Stages of an Introduction

Content & Structure

Introductions - Key Language

Methods

Language Task 1

Methods - Examples

Language Task 2

Methods - Key Language

The Method is like a recipe. It should be clear and easy to replicate. The **passive voice** is used to describe procedures simply and clearly. **Language which helps with precision and specificity** is also important.

Reports – Method – Key Language

Writing for Psychology

THE PASSIVE VOICE¹ TO DESCRIBE PROCEDURES

The **passive voice** - **VERB TO BE + PAST PARTICIPLE**, e.g. *participants were asked ...*, *the participant was asked ...* is a very important linguistic feature of a Report Methods section.

Participants	in all five experiments in Experiment 1 in ...	were	recruited	by ... from ... via ...
---------------------	--	-------------	------------------	--

e.g.

Participants were recruited from a secondary school in Belgium on a voluntary basis.

Participants were recruited via email and social networking sites and directed to the web address of the experiment.

participants they	were randomly	assigned to	one of (three) groups the (experimental) group and the (control) group. two different groups.
		divided into	a (control) group and an (experimental) group.

e.g.

Sixteen participants were randomly assigned to the low-threat group and 15 **were randomly assigned to the** high-threat group.

In Experiment 2, **participants were randomly assigned to two different** groups.

participants/ they	were presented	with	a/ the/ 3/ either ...
each/the participant	was presented		
<i>conditions/cues/ digits/ images/ instructions/ pictures/ stimuli/ words etc.</i>	were presented	in/ on/ for/ one by one/ repeatedly etc.	
<i>the stimulus/ the task/ each</i>	was presented		

- The Key Language sheets are available within the resource for the four main sections of Reports (Introductions, Methods, Results, Discussions) and for Critical Reviews (so far)
- The rest of this presentation will look at how the corpora and corpus tools (in AntConc) were used to create the Key Language sheets, and what corpus knowledge/skills were involved ...



The corpora

For Practical/Lab Reports

22 BAWE texts (levels 1-4) – 10 x methodology recount + 12 x research report – chopped into the target sections

- 0. Abstracts
- 1. Introductions
- 2. Methods
- 3. Results
- 4. Discussion
- Original txt files
- BAWE_Info re Psych Reports & Meth Recs

- 0011c
- 0011e
- 0011f N.B. No Abstract
- 0014a
- 0014b N.B. No sections
- 0016c
- 0022b
- 0033a
- 0033c
- 0037a
- 0037b
- 0057d N.B. No Abstract
- 0088a Only abstract intro & meth
- 0088c
- 0104a No Abstract or Results (Disc. re Method)
- 0131b No Abstract
- 0238a
- 0238b No Abstract
- 0238e
- 0240c
- 0262c
- 0262f

For Practical/Lab Reports

Research Article
texts from
AntCorGen
(Anthony, 2019)

- Psychology (37629)
 - All (37629)
 - Addiction (1112)
 - Applied psychology (37)
 - Behavior (18133)
 - Cognitive psychology (19151)
 - Collective human behavior (662)
 - Developmental psychology (672)
 - Emotions (5272)
 - Experimental psychology (291)
 - Gender identity (46)
 - Instinct (26)
 - Neuropsychology (771)
 - Pain psychology (39)
 - Personality (1241)
 - Psychological adjustment (76)
 - Psychological attitudes (1876)
 - Psychological defense mechanisms (19)
 - Psychological stress (1380)
 - Psychometrics (1618)
 - Psychophysics (433)
 - Relaxation (psychology) (97)
 - Sexual identity (48)
 - Social psychology (1530)

- 00. AntCorGen_PSYCH (1,393 RAs)
- 00. AntCorGen_PSYCH_Exp Psych only (281 RAs)

- abstract
- abstract_primary_display
- conclusions
- id
- introduction
- materials_and_methods
- results_and_discussion
- title

For Practical/Lab Reports

Research Article
texts from
AntCorGen
(Anthony, 2019)

Psychology (37629)
All (37629)
Addiction (1112)



00. AntCorGen_PSYCH (1,393 RAs)
00. AntCorGen_PSYCH_Exp Psych only (281 RAs)

Rationale:

- Lab Reports & RAs in Psychology have a lot of shared features
 - from my own observations of genre similarities
 - backed up by disciplinary specialists
 - tested via running word lists, keywords lists & n-grams using AntConc
- practically, a larger corpus which will yield a lot more useable language patterns for Report-writing

For the other genres ...

Essays – 56 texts: BAWE levels 1-3

Critical Reviews – 3 texts:

1 departmental UG student example

2 out of the 8 BAWE Psychology ‘critiques’ (both level 4)

Reflective Writing – 1 text: a lecturer-written example

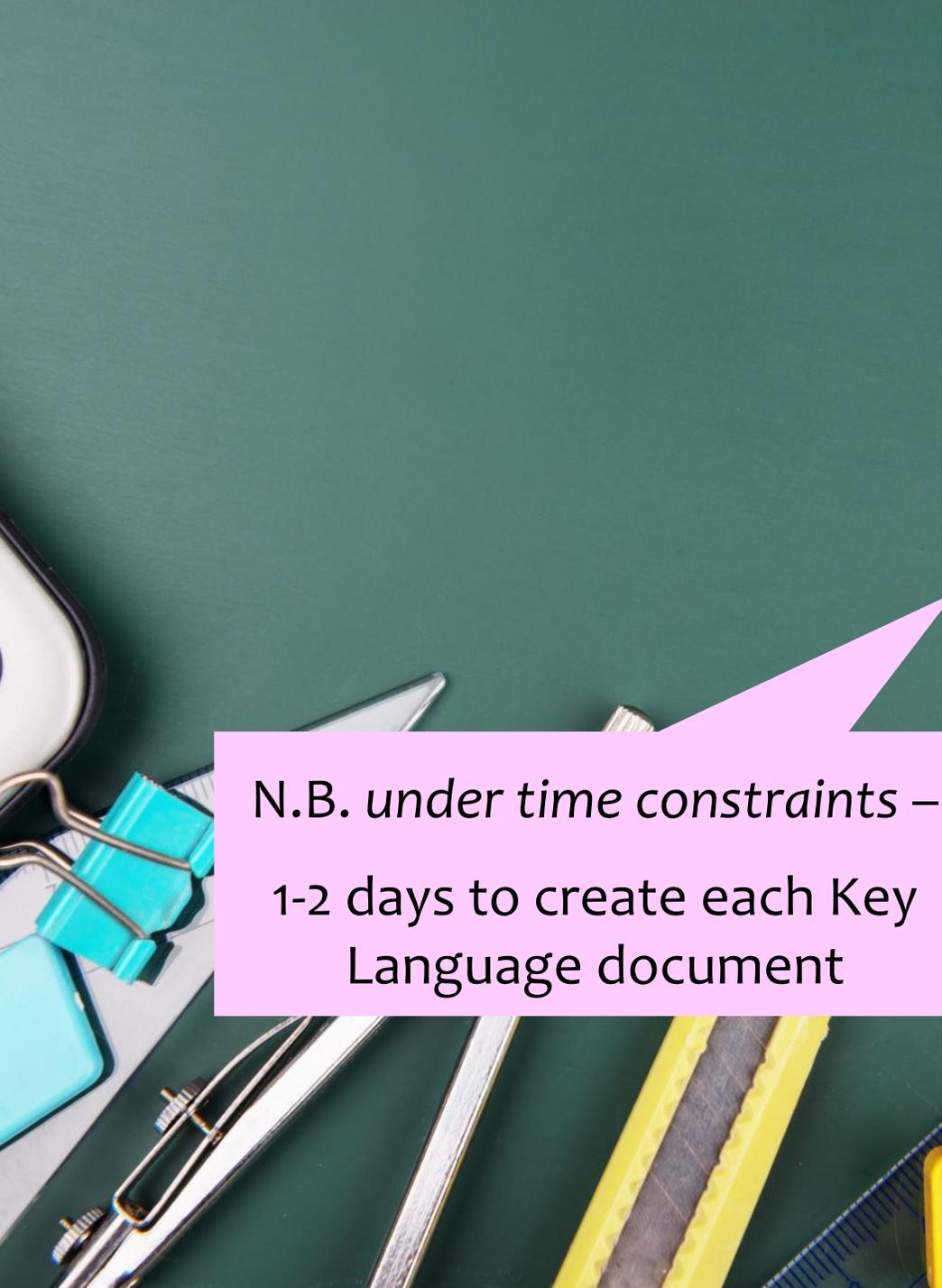
N.B. A comparison between these and the department-supplied example show they are similar enough that level difference doesn't matter



id	title
0016b	Write a review of Blackmore (1999) The Meme Machine
0031b	Is Bowlby's concept of attachment theory still relevant for studying social development?
0082b	Describe How Contemporary Developmental Research Has Challenged One of the Main Devel
0082h	Write a Review of Blackmore (1999) The Meme Machine.
0088b	Helping strangers around the globe
0104b	A Comment on Gee, Coventry, and Birkenhead
0171b	A Comment on Straver and Johnston (2001)
0384b	In light of our current understanding of anxiety disorders, evaluate Freud's interpretation and hand

Knowledge/skills used in corpus assembly

- knowing what relevant corpora are available
- manipulating corpus size
- linking genre knowledge to corpus knowledge:
 - Psychology's **'Practical Reports'** \leftrightarrow the BAWE **Methodology Recounts and Research Reports**
 - **RAs** \leftrightarrow **'Practical Report'-type texts**
 - BAWE **Critiques** \leftrightarrow the type of **'Critical Review'** that is being taught
- checking relevance/suitability of corpus texts
- for this purpose, use of corpus tools in combination with discourse analysis
- knowing when its OK to make pragmatic compromises re: level and genre



Creating the 'Key Language' sheets for Reports – how corpus tools were used

N.B. *under time constraints* –
1-2 days to create each Key
Language document

Using corpus tools to find
patterns for pre-determined
language features

Corpus Files		Concordance	Concordance Plot	File View	Clusters/N-Grams	Collocates	Word List	Keyword List
		Total No. of Cluster Types 59		Total No. of Cluster Tokens 1072				
Rank	Freq	Range	Cluster					
1	163	89	found that					
2	72	49	found to					
3	58	40	been found					
4	55	40	was found					
5	45	30	found in					
6	36	23	been found to					
7	35	24	have found					
8	28	23	found to be					
9	25	23	found that the					
10	25	18	has been found					
11	20	17	be found					
12	19	16	and found					
13	19	18	have been found					
14	19	15	it was found					
15	19	16	were found					
16	18	18	they found					
17	17	12	has been found to					
18	17	14	was found that					
19	17	12	we found					
20	16	13	have found that					

Clusters search for **found** in the Introductions (RA/Experimental Psychology corpus)

Search Term <input checked="" type="checkbox"/> Words <input type="checkbox"/> Case <input type="checkbox"/> Regex <input type="checkbox"/> N-Grams	Cluster Size
<input type="text" value="found"/> <input type="button" value="Advanced"/>	Min. <input type="text" value="2"/> Max. <input type="text" value="6"/>
<input type="button" value="Start"/> <input type="button" value="Stop"/> <input type="button" value="Sort"/>	Min. Freq. <input type="text" value="5"/> Min. Range <input type="text" value="5"/>
Sort by <input type="checkbox"/> Invert Order Search Term Position	
<input type="text" value="Sort by Freq"/> <input type="checkbox"/> On Left <input type="checkbox"/> On Right	

DESCRIBING INDIVIDUAL STUDIES

Individual studies are often included as examples of a wider research trend and usually talked about **in the past simple tense**.

For example/instance, [Author(s)]
 In one study, it was
 Another study conducted in [Columbia] **found** that ...
 One [functional magnetic resonance imaging (fMRI)] study
 [Author(s)] investigated/examined [topic] and
 [Author(s)] used [method] and
 Using [method] they

For example, one study/
 For example, [Lykken's landmark] study/ **showed** that ...
 One recent [adoption] study/
 A recent [animal] study
 [Authors] investigated [Topic]. These authors/They/The results/Results

Clusters tool → manual search through results for patterns involving single studies/past simple tense → **Concordances tool**

e.g.

Potts, Dedmon and Halford (1996) investigated the strength of motives, and found that the strongest motives in television viewers were entertainment, relaxation, information, and passing the time, while ...

One recent adoption study showed that the adoptive family rearing environment can be

SHOWING THE RESEARCH GAP

A LOT OF RESEARCH ATTENTION

Even though **much** attention has been paid to...
Although research **consistently** highlights ...
Although **several** studies have ...
Although **a large body of** research has ...

Most research focuses on...
 Studies have shown ...
 Studies ... have **mainly** examined ...
Several studies have ...
 There is **extensive** research investigating ...
 Prior research has tended to focus on ...

p. 8.

CONTRASTED WITH LITTLE/NO RESEARCH ATTENTION

, existing research **mainly** explored ...
None of these studies ... **neither** did they ...
 we **lack** insight into ...
Much less attention has been devoted to ...
 , **scant** attention has been paid to ...
 ... **relatively little** work has concentrated on

while **less** attention has been paid to ...
 , yet **little** attention has been given to ...
 ; however, **little** attention has been given to ...
However, there has been **little** research into ...
However, there has been **much less** focus on ...

Concordances tool to search for patterns around words like *research*, *much*, *little*, *However* etc.

WHAT IS KNOWN/HAS BEEN DONE

Taken together, the previous research suggests that ...
 It is already known that ...
 [Authors] suggested that ...
 Some evidence suggests ...
 Previous research on [Topic] has mostly focused on ...

CONTRASTED WITH WHAT IS STILL UNKNOWN

However, it is **(yet) unknown** whether ...
However, it is **not yet known** if ...
However, it is **less clear** how ...
However, it is **unclear/not clear** whether ...
However, it is **still unclear** whether ...

Analysis of **Concordances** for *However* (and of the 2-gram *it is*) revealed this 'known → unknow/not yet clear' pattern

Using corpus tools to reveal
salient features (which you
might not know about)

Rank	Freq	Keyness	Effect	Keyword
1	397	+ 627.51	0.0022	moral
2	5055	+ 496.8	0.0268	is
3	1033	+ 447.2	0.0056	has
4	921	+ 311.58	0.005	been
5	2606	+ 275.82	0.014	are
6	10250	+ 261.4	0.0524	to
7	715	+ 247.05	0.0039	people
8	1392	+ 226.44	0.0075	have
9	470	+ 201.87	0.0026	cognitive
10	920	+ 201.75	0.005	studies
11	614	+ 159.58	0.0033	self
12	325	+ 151.38	0.0018	processing
13	861	+ 147.62	0.0047	such

Keyword* List for introductions

*research corpus = introductions;
reference corpus = materials & methods
+ results & discussion

Keyword* List for materials & methods

*research corpus = materials & methods;
reference corpus = introductions +
results & discussion

Rank	Freq	Keyness	Effect	Keyword
1	2608	+ 1899.67	0.0258	were
2	2381	+ 979.74	0.0234	was
3	1049	+ 783.85	0.0106	each
4	1542	+ 653.59	0.0154	participants
5	15655	+ 562.92	0.1176	the
6	192	+ 553.2	0.002	consent
7	147	+ 447.67	0.0015	θ
8	180	+ 410.63	0.0018	informed
9	259	+ 395.97	0.0026	screen
10	177	+ 355.96	0.0018	university
11	129	+ 331.67	0.0013	ethics
12	275	+ 323.29	0.0028	asked
13	788	+ 302.64	0.0079	all
14	187	+ 300.5	0.0019	metaphors

Rank	Freq	Keyness	Effect	Keyword
1	397	+ 627.51	0.0022	moral
2	5055	+ 496.8	0.0268	is
3	1033	+ 447.2	0.0056	has
4	921	+ 311.58	0.005	been
5	2606	+ 275.82	0.014	are
6	10250	+ 261.4	0.0524	to
7	715	+ 247.05	0.0039	people
8	1392	+ 226.44	0.0075	have
9	470	+ 201.87	0.0026	cognitive
10	920	+ 201.75	0.005	studies
11	614	+ 159.58	0.0033	self
12	325	+ 151.38	0.0018	processing
13	861	+ 147.62	0.0047	such

Keyword* List for introductions

*research corpus = introductions;
reference corpus = materials & methods
+ results & discussion

Keyword* List for materials & methods

*research corpus = materials & methods;
reference corpus = introductions +
results & discussion

Rank	Freq	Keyness	Effect	Keyword
1	2608	+ 1899.67	0.0258	were
2	2381	+ 979.74	0.0234	was
3	1049	+ 783.85	0.0106	each
4	1542	+ 653.59	0.0154	participants
5	15655	+ 562.92	0.1176	the
6	192	+ 553.2	0.002	consent
7	147	+ 447.67	0.0015	θ
8	180	+ 410.63	0.0018	informed
9	259	+ 395.97	0.0026	screen
10	177	+ 355.96	0.0018	university
11	129	+ 331.67	0.0013	ethics
12	275	+ 323.29	0.0028	asked
13	788	+ 302.64	0.0079	all
14	187	+ 300.5	0.0019	metaphors

Rank	Freq	Keyness	Effect	Keyword
1	397	+ 627.51	0.0022	moral
2	5055	+ 496.8	0.0268	is
3	1033	+ 447.2	0.0056	has
4	921	+ 311.58	0.005	been
5	2606	+ 275.82	0.014	are
6	10250	+ 261.4	0.0524	to

Keyword* List for introductions

*research corpus = introductions;
reference corpus = materials & methods
+ results & discussion

- keywords quickly checked out with Clusters/Conc → quick decisions made regarding usefulness
- Reasons for use or non-use of keywords:
 - frequency was a factor, but some v. frequent keywords (e.g. *is*) didn't seem to have clear-enough patterns for my purposes (at least in the time I have available)
 - what I think will be most useful for genre purposes
 - keywords that link thematically (in a functional way), e.g. *each, the, all* for methods
 - some keywords overlap – phrases for one are picked up through analysis of another, e.g. *participants* already picked up in analysis of *were*
- N.B. I have undoubtedly missed useful things** but, given time constraints, can't be completely comprehensive!

Rank	Freq	Keyness	Effect	Keyword
1	108	+ 1899.67	0.0258	were
2	81	+ 979.74	0.0234	was
3	49	+ 783.85	0.0106	each
4	42	+ 653.59	0.0154	participants
5	655	+ 562.92	0.1176	the
6	2	+ 553.2	0.002	consent
7	7	+ 447.67	0.0015	θ
8	0	+ 410.63	0.0018	informed
9	9	+ 395.97	0.0026	screen
10	7	+ 355.96	0.0018	university
11	9	+ 331.67	0.0013	ethics
12	5	+ 323.29	0.0028	asked
13	8	+ 302.64	0.0079	all
14	7	+ 300.5	0.0019	metaphors

Recently,
In recent years
In the past few years
In the past (fifteen) years
In the last decade
To date

there **has been** (increasing/(an) increased/growing/significant) interest in ...
a growing recognition of the importance of ...
mounting evidence suggesting ...
etc.

e.g.

In recent years, there has been a growing interest in such individual variation, assessed using ...

Keywords* (*has* (and/or *been!*)) → Clusters → Concordances

[X]

has been found to be
has been shown to be

an important variable in ...
an important determinant of ...
a strong/reliable predictor/indicator of ...
a major factor (affecting) ...
associated with ...
(closely/positively) related to ...
(positively) correlated with ...
linked with .../(negatively) linked to ...
.../sensitive to ...

[Y]

e.g.

Recently, competition has been found to be a major factor affecting trust and trustworthiness ...

Keyword* (*were*) → Clusters → Concordances

participants they volunteers respondents	were	asked to /instructed to	<i>answer/ choose/ copy/ describe/ estimate/ evaluate/ identify/ imagine/ indicate/ judge/ move/ perform/ rate/ remember/ respond to/ score ... etc.</i> fill in /fill out /complete a questionnaire
	each/the participant	was	asked whether... not to ... on how to ...
			instructed

e.g.

Using the same procedure as for Experiments 1 and 2, participants were asked to *evaluate* how easily they understood each metaphor.

Next, they were asked to *rate* the global intensity and the duration of the emotion episode.

Each participant was asked to *complete* a questionnaire in which they evaluated

Rank	Freq	Keyness	Effect	Keyword
1	2608	+ 1899.67	0.0258	were
2	2381	+ 979.74	0.0234	was
3	1049	+ 783.85	0.0106	each
4	1542	+ 653.59	0.0154	participants
5	15655	+ 562.92	0.1176	the
6	192	+ 553.2	0.002	consent
7	147	+ 447.67	0.0015	θ
8	180	+ 410.63	0.0018	informed
9	259	+ 395.97	0.0026	screen
10	177	+ 355.96	0.0018	university
11	129	+ 331.67	0.0013	ethics
12	275	+ 323.29	0.0028	asked
13	788	+ 302.64	0.0079	all
14	187	+ 300.5	0.0019	metaphors

BEING PRECISE & SPECIFIC

It is very important in a Method to **always be precise and specific about who you are talking about** (e.g. *all* or *some* or *one* or *one particular sub-group* of the participants) **and what you are talking about** (e.g. *which stage of the experiment?* *Which tool or task?* *Which condition?*).

For this reason, **three small but important words are unusually frequent** within a Method: **the**, **each** and **all**.

the

The determiner **the** is important because it identifies **a known or pre-specified group, condition, material etc. within the study**.

In **the control group** the participants completed **the questionnaire** and were then shown **the advert**.

Participants **in the 'mnemonic' condition** were then taught **the technique** using example cards similar to the test cards.

Subjects were given **the printed sheet** containing all three word lists and asked to start learning **the first list only**, trying to ignore the sound from **the headphones**, while **the appropriate sound file** was played at the same time.

This questionnaire was altered slightly **for the experimental group** to measure the effect **the new peer group** was having over self esteem levels at university.

The determiners **each** and **all** clarify **who** from the participants **or which** from the tools, conditions, tasks etc. **is involved in any given stage of the procedure**.

each

Questionnaires were given out to **each of the participants**.

The design was counterbalanced so that **each of the thirty participants** looked at the word lists and corresponding files in a different order ...

... the dependent variables were the scores obtained **by each participant** for the Big Five personality

Total No. of N-Gram Types 32195 Total No. of N-

Rank	Freq	Range	N-gram
1	2867	265	of the
2	2040	258	in the
3	1134	240	to the
4	828	222	on the
5	737	215	that the
6	720	211	in a
7	695	163	e g
8	665	203	to be
9	645	202	of a
10	633	199	and the
11	594	202	it is
12	543	186	for the
13	478	189	with the
14	460	179	can be
15	450	177	as a
16	423	156	the same
17	418	175	by the
18	417	136	i e
19	406	174	such as
20	397	178	has been
21	385	160	for example
22	384	158	in which
23	373	154	from the
24	363	167	to a
25	315	140	based on
26	309	121	number of
27	308	160	as the
28	308	159	have been
29	306	162	is a

WHAT IS KNOWN/HAS BEEN DONE

Taken together, the previous research suggests that ...
 It is already known that ...
 [Authors] suggested that ...
 Some evidence suggests ...
 Previous research on [Topic] has mostly focused on ...

CONTRASTED WITH WHAT IS STILL UNKNOWN

However, it is (yet) unknown whether ...
 However, it is not yet known if ...
 However, it is less clear how ...
 However, it is unclear/not clear whether ...

It has been hypothesized that deaf people may explore and see the visual world differently from hearing people because of ...

Recently, In recent years In the past few years In the past (fifteen) years In the last decade To date there has been (increasing/(an) increased/growing/significant) interest in ... a growing recognition of the importance of ... mounting evidence suggesting ... etc.

e.g. In recent years, there has been a growing interest in such individual variation, assessed using ...

[X] has been found to be has been shown to be an important variable in ... an important determinant of ... a strong/reliable predictor/indicator of ... a major factor (affecting) ... associated with ... (closely/positively) r (positively) correla linked with .../(negati ... /sensitive to ... [Y]

again, v. quickly!

TALKING ABOUT PREVIOUS RESEARCH

A/a great number of studies have addressed/applied/examined/ investigated/measured/tried to/use...
 A/A increasing large limited small e.g.²

e.g. demonstrated/found/highlighted/proposed reported/revealed/shown ...

A number of studies have tried to explain the effects of colors on the human body.

N-Grams tool → qualitative selection of which ones to investigate further → Clusters/Concordances tool to identify useful patterns

When to use which corpus

RA corpus (AntCorGen)

- For sections & parts of sections where *the same kind of thing* is happening linguistically in both corpora
- Where a greater number of useful language patterns can be generated by using the RA corpus

BAWE corpus (approx. 22 research report & methodology recount texts)

- Where the student corpus diverges from the RA corpus:
 - Separate 'Results' & 'Discussion' in student texts vs. 'Results & Discussion' in RA corpus
 - Statistical analyses in Results
 - The 'claiming centrality'-type Move in the Introduction – going to be more general and simpler (the bar is lower!)
- Where enough useful examples can be found in the student texts
 - e.g. Aims & Hypotheses in the Introduction
 - Statistical analyses in Results

again, quick decisions with not much time to reflect/re-think!

SAYING THE TOPIC IS IMPORTANT

e.g.¹

Use adjectives
that express
importance ...

The study of human memory **is of great importance in** learning and education, as well as for disorders such as amnesia, and to generally understand the functions and processes of memory.

The effect of the presence of others on individual performance, decision making and behaviour **has been a major focus of** research since the pioneering study of Norman Triplett in 1898.

Gratitude is a significant emotion in modern day society and is seen to be a commonly experienced affect.

... or describe
why the topic
is important in
the real world

The need to affiliate with others **is a central part of** human nature.

Researchers have argued that **road crossing skills** in children **depend on attentional skills** and ability to concentrate on a task during a distracting event.

Eating disorders are becoming **an increasing problem**.

General or summary statements regarding your **Results**

Near the beginning of your Discussion section ([linking to hypotheses/predictions](#)) ...

Results show that **all three hypotheses were supported** by the data gathered. ...

The results obtained **support the hypothesis** posited at the beginning of the study, which stated that ...

The results **partially confirm the original hypothesis** - there was

Such results **confirmed our hypothesis**, which stated that ...

The results showed that there was no significant relationship between ... and ...

Generally, the findings **were inconsistent with the experimental hypothesis** that The results showed that ...

From the results, we can see that ...

We predicted that **Our results support this prediction**, showing that ...

... and near the end of your Discussion section (implications)

Suggestions for future research

The findings of this study may not be :

Despite being specific to ..., **the result**

Our results therefore **confirm the exper** study and suggest that future research

A superior design of this study may be drawn up using a more in-depth and suitable questionnaire consisting of ...

Replications of this study **would be useful in order to** clarify ...

... therefore **further or repeated research would be advantageous** to rule out the possi of ...

An interesting advance to the research would be to see if ...

... **it would be more relevant to everyday situations to test** memory whilst actual conversations or television recordings are played as background noise, or to test long te memory by delayed recall. **Conditions could also be extended to include** another language white noise, music or timed beeps. **This would clarify** exactly what kind of noise is distra to memory; all noise or just speech, all speech or just speech of a viable language.

Future work **should** ...

... future research **could** investigate ...

Future studies **could** examine ...

Future research **is needed** to ...

Critiquing your method

Potential problem in design

... the purpose of the advert cannot have escaped anyone in the intervention groups, arguably to the extent that the obviousness of the attempt at psychological manipulation **may have** annoyed several participants.

The main problem here is that all three used a slightly different scale.

Other problems include the type of distraction - ...

Possible impact of this problem

This may have caused them to give more positive answers.

This may have caused confusion for participants, who might have used the wrong scales whilst answering questions.

It might be that cartoons appeal more to boys than girls ...

es , here the within subjects design **may have affected** the recorded response times ...

Examples from the **Key Language for Discussion sheets** – created mostly via Concordance analysis of pre-determined words, e.g. various modal verbs used in this section

statistical vs. qualitative
analyses (+ how to use a v.
small 'corpus'!)

Concession → Criticism (using **contrastive signalling**)

Whilst this was controlled to some extent through placing the single-task both before and after the dual-task, it could still lead to noisy results.

Whilst the results for this experiment discuss the miss rate, they do not show the correct rejection rate.

In these experiments milliseconds were a very reliable measurement to use to time reactions. **However**, the report does not state how the 'tracking error' in the second experiment was measured.

Although quasi-experimentation has the advantage of ..., the internal validity is threatened by ...

Cautious criticism

The authors **do not appear to** have support for ...

... **there is a question as to whether** these assumptions are valid ...

This is a sampling bias and **can** lead to incorrect ...

The design of ... shows **potential** problems ...

... this design **can** encourage guesswork which **could have** been a concurrent factor that affected results ...

... it **could** still lead to noise in the results ...

Implications of methodological choices

The control test **prevented** questions being asked about ...

To ensure that the reliability of experiments is not threatened by imprecise measurements, ...

... random assignment ... **minimizes** individual difference confounds ...

An incomplete design of counterbalancing was used, **ensuring that** individual confounds are minimised ...

A confederate blind to the condition the participant is in is used **therefore reducing the potential for** experimenter effects and error variance ...

Using such a method **would mean that** the design of the experiments would have to be changed to ... [N.B. implication of suggested improvement to the research design]

Ecological validity in this study **is better** than a study done in a laboratory **because** the results can be applied to a natural gambling phenomenon ...

... procedure ... **increasing** validity.

Suggesting improvements

A more accurate way to run this test **would** be to ...

A higher level of realism in the test **would** allow for ...

A solution **would** be to ...

A sample from a broader population **should have** been used ...

Key Language for **Critical Reviews**

- Completely qualitative analysis of 2 texts (I still need to add in the third)
- A focus on repeated functions within this genre with phraseologies illustrating these and showing some possible language options

Writing for Psychology - A Timeline for Reflecting ...

looking back

linking the past to now

Stages ¹	Before the event	During the event	After the event	Explaining the experience	Now & the future
Questions to ask yourself	<ul style="list-style-type: none"> What did you think would happen? Why did you have these expectations? How did you feel? 	<ul style="list-style-type: none"> What happened and in what order? (i.e. what were the key stages of the event?) What did you think? How did you feel? How did you respond to things that happened? 	<ul style="list-style-type: none"> Were your expectations from <i>before</i> the event correct or mistaken? What were your insights immediately after the event, when your feelings were still fresh? Did your insights/feelings change when you had more emotional distance from the event? 	<ul style="list-style-type: none"> Can any aspects of what you have learnt be explained by theory or research evidence? 	<ul style="list-style-type: none"> What have you learnt? How has the experienced changed you? What do you know now? What will you do in similar situations in the future?
Useful language*	<p>I experienced some frustration that I was now required to learn how to write essays again.</p> <p>Looking back, I think I had wanted to demonstrate that I already knew how to write essays well.</p>	<p>I made sure that I spent a lot of time writing it because it was my first essay at University and I wanted to get high marks. I didn't spend too much time checking the guidelines because I had already written so many essays before.</p> <p>When I received my feedback, my mark was much lower than I expected it to be which upset me.</p> <p>Once I discussed the feedback with my essay marker I understood that ...</p> <p>I felt a little bit cheated at having to re-learn skills which I felt I had already mastered.</p> <p>However, something that my Skills tutor mentioned in a feedback tutorial about how it is the assignment which is being marked and not the student made me re-interpret my essay grade, and feel more positive about it.</p>	<p><i>On reflection, I think</i> because I had wanted to demonstrate that I already knew how to write essays well, this lower mark had hit me particularly hard.</p> <p>The process of receiving feedback was quite challenging for me, and made me lose confidence in my abilities.</p> <p>I had per personal</p> <p>I also re of the ac too low, master t</p>	<p>This is related to positive educational and psychological outcomes, including higher self-esteem (Martin & Marsh, 2006).</p> <p>Students' realistic expectations of teaching and learning in higher education are associated with better achievement outcomes (Nicholson, Putwain, Connors & Horby-Atkinson, 2013).</p>	<p>I now understand that ...</p> <p>I can apply this knowledge to my work at Part 2 by ...</p> <p>I feel that this experience has helped develop my ability to ...</p> <p>In fact, I do feel better about myself and my</p>

mainly *description*

mainly *evaluation*

Key Language for Reflective Writing

- One text!
- Language foci: *you, your senses, perceptions, evaluating, and verb tense* for different stages within the reflection

When to use and when not
to use corpus input ...

I've not prioritized creation of a 'Key Language' sheet for Essays because ...

rightly or wrongly???

Lab/Practical Reports

- a lot of formulaic language
- formulaic language links to genre expectations
- so, Key Language sheets provide both language to use and help teach the genre

Essays

- less use of formulaic language (beyond 'This essay will ...' and 'In conclusion, ...')
- reading-to-write processes and examples of resulting 'product' more useful for understanding of genre expectations
- a Key Language sheet might risk constraining students & making their Essay-writing too formulaic??

Critical writing and effective source use

Source Use & Writer Voice - A Task

Source Use & Writer Voice - Analysis

Argumentation in Essays

A Model for Argumentation

Psychology Essay Types - Task

Psychology Essay Types & Argumentation

Introductions & Conclusions

Introductions & Conclusions Task 1

Introductions & Conclusions Task 2

however ...

Concordance		Concordance Plot	File View	Clusters/N-Grams	Collocates	Word List
Total No. of N-Gram Types			55949	Total No. of N-Gram Tokens		105757
Rank	Freq	Range	N-gram			
1	563	55	of the			
2	357	54	in the			
3	322	54	to be			
4	321	52	it is			
5	266	36	et al			
6	244	52	to the			
7	216	52	that the			
8	211	52	can be			
9	193	53	on the			
10	180	49	and the			
11	169	48	in a			
12	163	48	such as			
13	154	50	as a			
14	149	28	cited in			
15	142	50	is a			
16	137	47	this is			
17	122	47	of a			
18	120	35	found that			
19	120	43	has been			
20	117	44	there is			

Format	View	Help
lo. of Cluster Types: 159		
lo. of Cluster Tokens: 567		
19	15	likely to be
18	14	to be more
17	13	to be a
16	10	need to be
16	12	seem to be
15	9	found to be
14	11	seems to be
13	11	to be the
12	8	appear to be
12	6	tend to be
11	9	appears to be
9	9	have to be
9	7	to be able
9	7	to be able to
8	8	more likely to
8	8	more likely to be
7	4	has to be
7	7	needs to be
5	5	been found to
5	5	been found to be

Format	View	Help
lo. of Cluster Types: 117		
lo. of Cluster Tokens: 338		
13	10	found that the
9	8	is that the
9	8	suggests that the
8	6	and that the
7	7	that the majority
6	3	implies that the
6	6	that the majority of
5	5	claim that the
5	5	fact that the
5	4	say that the
5	5	states that the
5	5	suggested that the
5	5	the fact that
5	5	the fact that the
4	3	argued that the
4	4	meaning that the
4	3	show that the
4	4	shown that the
4	4	that the child
4	2	that the human

Format	View	Help
lo. of Cluster Types: 92		
lo. of Cluster Tokens: 369		
29	10	can be seen
22	9	it can be
13	11	can be used
11	9	can be explained
10	4	can be seen that
10	9	can be used to
10	4	it can be seen
10	4	it can be seen that
9	8	that can be
9	9	this can be
8	6	can be applied
8	6	can be seen in
8	8	which can be
6	6	can be a
6	5	can be applied to
5	3	research can be
4	4	and can be
4	2	can be argued
4	2	can be argued that
4	4	can be explained by

Format	View	Help
lo. of Cluster Types: 60		
lo. of Cluster Tokens: 263		
22	13	it was found
22	13	it was found that
22	13	was found that
13	10	found that the
10	8	been found that
8	5	also found that
8	6	found that when
7	5	has been found
7	5	has been found that
7	7	have found that
		it has been found
		it has been found that
		they found that
		found that children
		and found that
		found that a
		found that participants
		found that there
		on been found
		been found that

Potentially a Key Language for Essays sheet focusing on

- reporting evidence/arguments from research
- interpreting and arguing from this evidence

Corpus knowledge/skills used

- how to use a range of corpus tools to:
 - explore language features you already know about *
 - throw up language features you may not know about **

(cf. Charles, 2018 – ‘addressing known issues’ vs. ‘identifying unknown issues’)
- when to use which type of tool & how to use them together,
- how to ‘read’ (or ‘*speed read*’) a corpus search result to select items to do further work with
- when to be more quantitative and when to be more qualitative with corpora
- when corpus work is most useful, when it’s (possibly) less so?
- *in all the above*, keeping a dialogue going between corpus work and genre analysis
 - using the genre analysis to better search/use the corpus*
 - using the corpus work to further analyse/understand target genres**

3. Academic Discourse

An EAP teacher will have a high level of systemic language knowledge including knowledge of discourse analysis.

Knowledge & understanding of –	Ability to –	Possible indicators –
<p>discourse features and sub-technical vocabulary which would allow teachers to read and make sense of texts without being subject specialists:</p> <ul style="list-style-type: none"> • grammar and syntax at the level of phrase, clause and sentence • discourse features beyond the sentence • cohesion and coherence • semantics and pragmatics • approaches to text classification, e.g., theories of genre and text type 	<p>apply theories of text and discourse analysis to course organization, materials selection and development, and assessment</p>	<p>analyse examples of academic genres in terms of the features of systemic language knowledge listed</p> <p>show the teaching of features of systemic language knowledge as an integral part of a course</p> <p>provide analytical feedback, across the range of features of systemic language knowledge, which promotes accuracy in students' spoken and written outputs</p>

how to use corpora/corpus tools constructively *in combination with* theories of text and discourse organization

how to use corpora and corpus tools to identify and explore these

Some generalisations ...

ESAP corpus research

- principled corpus design
- careful/often v. time-consuming
- v. good at producing valuable 'big picture' insights ...
- *but* often less good at generating teachable bits of language in any kind of useful volume

Corpora/corpus tools in ESAP practice

- *as principled as possible* corpus design, but judicious compromises usually needed
- v. tight *time constraints* – *shortcuts* often needed
- need to generate as many useful, teachable bits of language as possible ...
- *but* this is time-consuming!

Some generalisations ...

ESAP corpus research

- principled corpus design

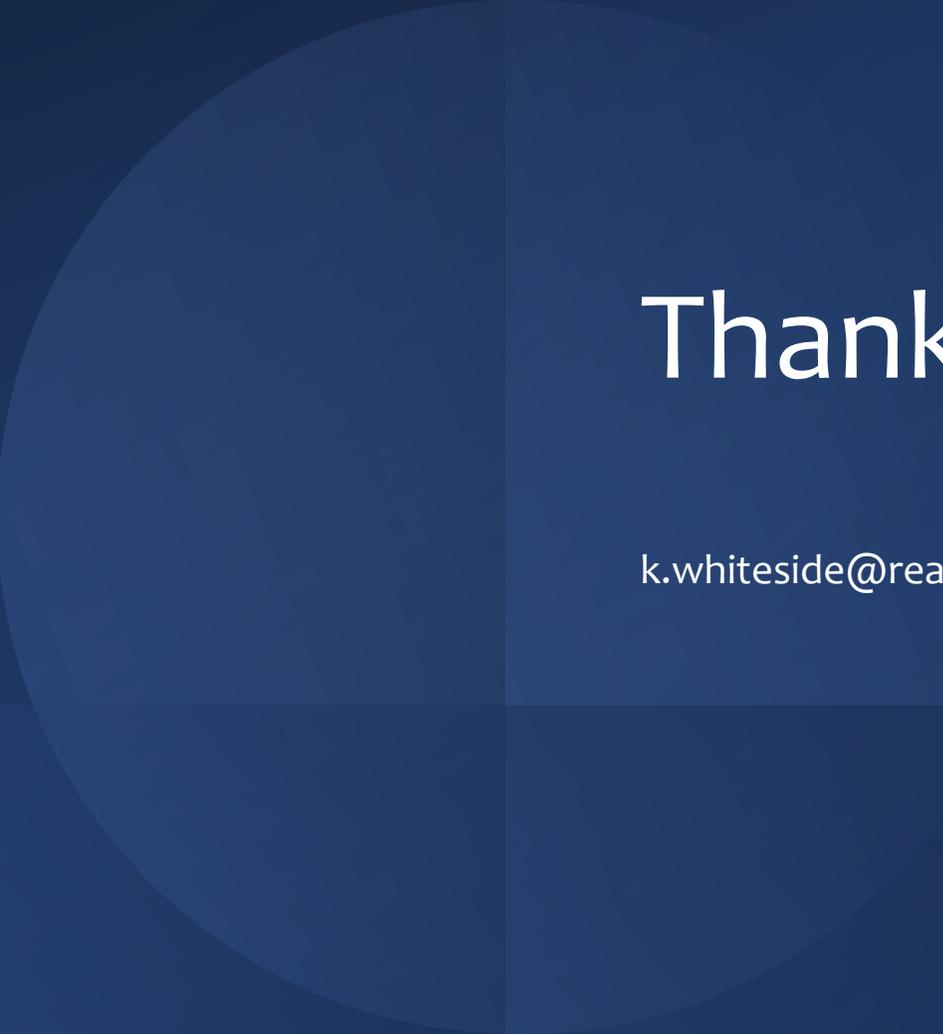
also, ...

- being pragmatic
- being able to judge
 - how much is possible
 - what to prioritise
 - when you've done enough!
- **KEY - making pragmatic judgements that facilitate good ESAP practice rather than leading to poor/questionable practice**

“your stopping condition is different”
(Kendon, 2021)

Corpora/corpus tools in ESAP practice

- *as principled as possible* corpus design, but judicious compromises usually needed
- v. tight *time constraints* – shortcuts often needed
- need to generate as many useful, teachable bits of language as possible ...
- *but this is time-consuming!*



Thank you for listening!

k.whiteside@reading.ac.uk